Darmstadt Discussion Papers in Economics

Endogenous (Re-)Distributive Policies and Economic Growth: A Comparative Static Analysis

Günther Rehme

Nr. 185

Arbeitspapiere
des Instituts für Volkswirtschaftslehre
Technische Universität Darmstadt

ISSN: 1438-2733

Economic Theory

Endogenous (Re-)Distributive Policies and Economic Growth: A Comparative Static Analysis

Günther Rehme*

Technische Universität Darmstadt†

May 21, 2007

Abstract

This paper analyzes the interplay of growth, (re-)distribution and policies when the latter are set exogenously or when the latter depend on economically important fundamentals. A redistribution policy generally causes lower growth, but less so when there is technological progress. The model implies that high (endogenous) tax rates may not necessarily imply low growth. The paper shows that the long-run cross-country relationship between growth and endogenous policy is generally not clear-cut. But this relies on conditions that can be used for identification in empirical research. The paper also argues that workers benefit more from technical progress than capital owners, even though inequality might and growth would rise.

KEYWORDS: Growth, Distribution, Endogenous Policy

JEL Classification: O4, D3, H2

---

*This paper is based on research that I have conducted at the European University Institute in Florence. I would like to thank Tony Atkinson, James Mirrlees, Robert Waldmann and Robert K. von Weizsäcker for helpful advice and useful suggestions. 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

It is often shown that policies which are optimal for the accumulated factor of production maximize growth and that high (re-)distributive taxes slow down long-run growth. See, for instance, Perotti (1993), Alesina and Rodrik (1994), Bertola (1993) or Persson and Tabellini (1994).

However, when analyzing the effects of policy on growth empirically, policy is mostly viewed as exogenously determined and in this context it turns out that - at least across countries - these theoretical predictions do not appear to command strong empirical support. See, for example, Barro (1991), Easterly and Rebelo (1993), Perotti (1994), and Sala-i-Martin (1996).[1]

In this paper I add to these contributions by distinguishing between exogenous and endogenous policy. The latter is given when policy is set optimally and, thus, takes account of fundamental economic variables. Endogeneity of policy may help explain why we observe policy-growth relationships that are sometimes at odds with theory. To make this point I focus on theory as one step in uncovering what differences exogenous or endogenous policy may imply for the policy-growth nexus.

For the analysis I concentrate on two policy instruments as metaphors for wider policy packages that may be analyzed in more general frameworks. One instrument is a tax rate that may cause a disincentive to accumulate. The other is an indicator for direct redistribution from the accumulated to the non-accumulated factor of production. A simplified version of the widely known model of Alesina and Rodrik (1994) that incorporates features shared by many other models provides the theoretical "lens" through which existing results are interpreted.

In the model taxes can be used to finance productive inputs in production as in Barro (1990) or to grant direct transfers to the non-accumulated

[1]Recent discussions of those issues can be found in Bénabou (1996), Bertola (1999), Temple (1999), Aghion, Caroli, and García-Peñalosa (1999), and Jovanovic (2000). It should be noted that some authors argue that growth is invariant to (some) policy (measures). See, for example, Stokey and Rebelo (1995). On the whole, though, a lot of historical evidence suggests that growth and development do in fact react to (fundamental) policy changes, including tax and redistribution policies. See, for instance, Landes (1998). In this paper I follow the latter evidence.

factor of production. As is common the accumulated factor of production is identified with capital and the non-accumulated factor of production with (unskilled) labour.

First, the model predicts that in equilibrium an inverted U-shaped relationship between taxes and growth holds, when taxes are set exogenously. Taxes higher than those which are optimal for the capital owners imply lower growth. Furthermore, higher taxes imply higher redistribution from capital to labour. These results are in line with many other theoretical contributions.

Second, we introduce optimizing governments. By assumption governments are either only concerned about the workers or only about the capital owners.[2] The optimal policy of an entirely pro-capital government is tantamount to a growth maximizing policy in the model.[3] In contrast, an entirely pro-labour government chooses higher taxes and, thus, lower growth.

In the model all optimal policies depend on three fundamental economic variables: the rate of time preference, an index of the state of technology and the (pre-tax) share of capital (income in total income). Thus, policy is economically endogenous.

Acknowledging that all these factors play a potential role, we concentrate on the state of technology (aggregate efficiency) as the prime mover of policy and fix the other determinants for the analysis. This is rationalized by the importance that aggregate efficiency is usually accorded to in explanations of long-run changes to the economic structure of a country. See, for example, Prescott (1998).

Analyzing the consequences of changes in aggregate efficiency the follow-

[2]This simplifying assumption captures that political preferences may be structurally fixed one way or the other for a long period of time. For evidence on this see, for example, Garrett (1998). Furthermore, the qualitative results would not change if instead governments attached different social weights on the workers' or capital owners' welfare.

[3]This need not necessarily be the case. For instance, Bertola (1993) shows that indirect taxes may produce different results. In this context, Rehme (2002a) uses an equivalence result for direct and indirect taxation to show that under certain conditions the optimal policy of the non-accumulated factor of production may also maximize growth and imply higher growth than under the optimal policy for the accumulated factor of production. However, here we follow the more conventional notion that the owners of the accumulated factor of production would generally opt for higher growth than the owners of the non-accumulated factor of production.

ing results then emerge: When fixing policy at some arbitrary (including some optimal) level, higher efficiency implies higher growth, but lower redistribution. The first result corresponds to conventional wisdom. A better technology allows a better use of resources in the accumulation process and that is reflected in a higher growth rate. The second result is not so straightforward. In the analysis redistribution is measured in terms of an arbitrary policy generating a particular after-tax-and-subsidy factor income distribution relative to the after-tax-and-subsidy factor income distribution generated under a growth maximizing policy.[4]

The latter distribution is independent of aggregate efficiency.[5] Thus, the after-tax capital income rises relatively more, when policy is fixed and efficiency rises. Hence, efficiency gains accrue relatively more to the capital owners and so redistribution from labour to capital is lower in this case.

These results imply a tradeoff: For given policy, higher efficiency entails lower redistribution, but higher growth. For given efficiency, taxes higher than those which are optimal for growth imply lower growth, but more redistribution. This suggests that governments could tax the beneficial effects of higher efficiency away and redistribute more. However, this only holds if policy is set exogenously.

In the third stage of the analysis, it is acknowledged that policy is economically endogenous. Thus, changes in efficiency will have a direct and an indirect effect on growth and redistribution.

The model then implies that for given political preferences the observed

[4]Many models on the growth-distribution nexus focus on the factor income distribution. As has been pointed out by Perotti (1996), these models are difficult to bring to the data. In defence of concentrating on the factor income distribution - in a theoretical model - we use the argument that the factor income distribution is of considerable importance for the distribution of personal incomes. See, for example, Atkinson (1983) and Atkinson and Bourguignon (2000). Furthermore, this approach has the advantage of seeing clearly, how policy affects the personal income distribution through the factor income distribution. That is not always so clear when concentrating on the personal income distribution directly.

[5]The ratio of pre-tax capital and labour income is independent of efficiency A. However, the ratio of after-tax incomes depends on the ratio of the tax rate to the wage rate. For fixed taxes, higher A implies higher wages. Thus, that ratio decreases. Since after-tax capital income is the pre-tax return on capital minus the (given) tax rate, it follows that after-tax capital income rises relative to the wage rate. Consequently, redistribution towards labour decreases in the model.

association between growth and taxes would be positive within countries. Thus, even optimizing, entirely pro-labour governments respect the beneficial effects of higher efficiency by not increasing taxes too much.[6] The same is true for redistribution. That implies that a negative relationship between redistribution and growth should be observed. This holds if one views changes in aggregate efficiency over time as capturing the development process in a particular country with optimizing governments.

The predictions are less clear-cut, when the analysis is applied to a cross-section of countries. In fact, the observed tax-growth as well as the redistribution-growth relationships should generally be ambiguous when the distribution of aggregate efficiency takes on more general forms. It would still be true that the observed tax-growth relationship is positive, but only if all countries in a sample would have the same political preferences. That, of course, is quite unlikely. Thus, no clear prediction on this relationship is in general possible.[7]

From the latter result negative implications for cross-country would seem to be inevitable. For instance, Rodrik (2005) has recently argued that we learn "nothing from regressing economic growth on policies." However, the present paper allows for a more constructive message. We may simply argue that there is a need to disentangle more precisely the relationship between policy and growth, by taking account of the influence of deep variables like aggregate efficiency, which (possibly) bear on both policy and growth. Theory may provide an important guiding tool in this respect.[8]

In a last step the welfare implication of efficiency changes are analyzed. Again policy is taken to be endogenous. Interestingly, I find that the individ-

[6]The intuition for this is that in the long-run the workers benefit more from the intertemporal gain induced by a higher growth rate with relatively lower taxes than by higher redistribution with relatively higher taxes.

[7]Furthermore, we might observe that countries operating under pro-labour policies may have higher growth than those under pro-capital policies. But for this to be the case, the former countries need to be sufficiently efficient to support such a regime. When one finds that redistribution and growth are positively associated, then the model attributes this to sufficient efficiency advantages of pro-labour vis-a-vis pro-capital countries.

[8]One way to constructively cope with this issue in relation to existing results may, for instance, be to analyze the problems associated with interpreting estimates in growth regressions when using theory. For such an approach see, for example, Rehme (2002b).

ual worker as well as a pro-labour government would never benefit less from efficiency increases than a capital owner or a pro-capital government. More efficiency is in the interest of all agents in the model, but - interestingly - the workers would prefer it relatively more.

The paper is organized as follows: Sections 2 and D.3 set up the model and derive the equilibrium. Sections 4 to 6 presents an analysis of exogenous policy, introduces optimizing governments and relates policy to economic fundamentals. Section 7 relates growth to endogenous policy. Section 8 presents the welfare analysis. Section 9 provides concluding remarks.

2 The Model

The economy is populated by two types of many, price-taking and infinitely lived individuals who are all equally patient. One group of agents, the capitalists, owns wealth equally and does not work. The other group is made up of workers who own (raw) labour equally, but no capital.[9] Population is stationary and consists of l workers and n capitalists of whom there are less, that is, $l > n$. Each individual derives logarithmic utility from the consumption of a homogeneous, malleable good.[10] Aggregate output is produced

[9]The assumption uses a short-cut of a result in Bertola (1993). He has shown in an endogenous growth model that for utility maximizing, infinitely lived agents who do not own initial capital, it is not optimal to save/invest out of wage income along a *long-run*, i.e. steady state, balanced growth path. Similarly, it is not optimal to work for those who only own capital initially. Thus, the set-up is reminiscent of Kaldor (1956), where different proportions of profits and wages are saved. However, in Kaldorian models growth determines factor share incomes, whereas in endogenous growth models the direction is rather from factor shares to growth.

[10]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

according to

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

where Y_t denotes aggregate output, K_t is the real capital stock, L_t is labour supplied, and G_t are public inputs to production.[11] Capital is broadly defined and by assumption human capital is strictly complementary to physical capital.

Thus, in the model capitalists who, for instance, own computers know how to operate them as well. This eliminates a separate treatment of how human capital is accumulated and entails that the return on human capital services equals that of physical capital services in a perfectly competitive economy. For a justification of such an approach in a different context see Mankiw, Romer, and Weil (1992).

The constant efficiency index A reflects the economy's state of technology. It depends on cultural, institutional and technological development and captures long-run, exogenous factors that play a role in production.

Each worker inelastically supplies $\frac{1}{l}$ units of (raw) labour at each point in time. As there are l workers in the economy, $L_t = 1$ so that the total labour endowment equals unity. Furthermore, the model abstracts from problems arising from the depreciation of the capital stock so that output and factor returns are really defined in net terms. This has no consequences for the price-taking, market clearing logic of the model.

2.1 The Public Sector

Following Alesina and Rodrik (1994) we analyze a wealth tax scheme which is meant to serve as a metaphor capturing the essential features of many

instantaneous utility.

[11]Like Barro (1990) one may assume that the government owns no capital and that it buys a flow of output from the private sector and makes it available to the *individual* firm. Then public inputs to production would be rival. Alternatively one may assume that *total* government expenditure affects private production in a non-rival way. This empirically relevant distinction does not matter analytically in the model. Note that in the absence of a government due to civil war or other forms of unrest, the economy would break down and the agents would starve.

different sets of (re-)distributive policies.[12] The government taxes wealth at the constant rate τ and transfers a constant share λ of its tax revenues to the workers. The tax on capital should be viewed as a tax on all resources that are accumulated, including human capital.

In line with most of the literature on capital taxation the paper abstracts from the taxation of raw labour. That allows one to focus on the distributional conflicts between accumulated and non-accumulated factors of production.

The government runs a balanced budget under the constraint,

$$\tau K_t = G_t + \lambda \tau K_t. \tag{2}$$

Of the tax revenues τK_t the workers receive $\lambda \tau K_t$ as transfers and G_t is spent on public inputs to production.

Thus, the government has two potential instruments for redistribution at its disposal. First, it can use the tax rate to generate revenue to provide public resources in production. This will have an effect on the factor returns in equilibrium and allows it to change the factor income distribution. Thus, the tax rate is redistributive in an indirect, and productive sense. Secondly, the government can raise revenue and directly redistribute these resources to the non-accumulated factor of production, i.e. workers, when setting λ.[13] These transfers are a direct way of redistributing resources to the workers. The effects of both instruments are analyzed, because both are used in reality.

[12] As tax schemes differ widely across countries due to historical, institutional or political differences an answer to the question why a society chooses a particular scheme has to remain outside of this model. For similar arguments and example what redistributive mechanisms the wealth tax scheme may capture see Alesina and Rodrik's paper. Furthermore, in the same framework they 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.

[13] As human and physical capital are strict complements by assumption this is a strong from of redistribution. It implies that if a capital good is given to the workers the corresponding services necessary to operate that good are also given to them. As a one good economy is contemplated, giving the capital good to the workers for consumption does not cause a problem.

2.2 The Private Sector

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 G_t as given, and rent capital and labour in spot markets in each period. The price of output serves as numéraire and is set equal to one. Profit maximization entails that firms pay each factor of production its marginal product

$$r_t = \frac{\partial Y_t}{\partial K_t} = \alpha A \left(\frac{G_t}{K_t}\right)^{1-\alpha} \quad , \quad w_t = \frac{\partial Y_t}{\partial L_t} = (1-\alpha) A \left(\frac{G_t}{K_t}\right)^{1-\alpha} K_t,$$

where we have used the normalization $L_t = 1, \forall t$. Substituting for G_t implies

$$r_t = \alpha A[(1-\lambda)\tau]^{1-\alpha} \tag{3}$$

$$w_t \equiv \eta(\tau,\lambda) K_t = (1-\alpha) A[(1-\lambda)\tau]^{1-\alpha} K_t. \tag{4}$$

Because of the productive role of government services policy has a bearing on the marginal products. The return on capital is constant over time, while the wages grow with the capital stock. Notice that higher direct transfers to the workers (higher λ) lower r and η, while higher taxes raise them. This is because higher taxes mean that more public resources can be channelled into production, raising the return to private capital and the wage rate. In contrast, granting more direct transfers implies that less resources are available for public inputs in production and, thus, the marginal products would be lower.

The total wage and transfer income is $\eta(\tau,\lambda)K_t + \lambda\tau K_t$. Each worker receives an equal share of it and derives utility from consuming his entire income. The representative worker's intertemporal welfare is given by

$$\int_0^\infty \ln c_t^W \, e^{-\rho t} dt \quad \text{where} \quad c_t^W = (\eta(\tau,\lambda) + \lambda\tau)\, \tilde{k}_t, \text{ and } \tilde{k}_t \equiv \frac{K_t}{l}. \tag{5}$$

Thus, the owners of the non-accumulated factor of production do not invest and are not taxed by assumption.

The capitalists choose how much to consume or invest. They have perfect foresight about the price and tax rate paths, which they take as given. The representative capital owner maximizes his intertemporal utility according to

$$\max_{c_t^k} \int_0^\infty \ln c_t^k \, e^{-\rho t} dt \tag{6}$$

$$s.t. \qquad \dot{k}_t = (r - \tau) k_t - c_t^k \tag{7}$$

$$k(0) = \overline{k_0}, \quad k(\infty) = \text{free}, \tag{8}$$

where $k_t \equiv \frac{K_t}{n}$. Equation (7) is the dynamic budget constraint of the capitalist which depends on his after-tax income. The growth rate of consumption and wealth can be calculated in a standard way (see Appendix A) and is given by

$$\gamma \equiv \frac{\dot{c}_t^k}{c_t^k} = \frac{\dot{k}_t}{k_t} = (r - \tau) - \rho. \tag{9}$$

Growth is increasing in the after-tax return on capital and constant over time. Furthermore, from equations (9) and (7), and the usual transversality condition one verifies that $c_t^k = \rho k_t$ is the capitalist's optimal level of consumption.[14]

3 Market Equilibrium.

Constant policies imply constant r and hence constant γ. The economy's overall resource constraint implies

$$I_t = \dot{K}_t = (r - \tau) K_t + (\eta + \lambda \tau) K_t - C_t^k - C_t^W. \tag{10}$$

As the workers' consumption is $C_t^W = (\eta + \lambda \tau) K_t$ in the aggregate, this constraint is binding, simplifying (10) to $\dot{K}_t = (r - \tau) K_t - C_t^k$. The capitalists' consumption $C_t^k = n c_t^k$ and wealth $K_t = n k_t$ grow at the constant rate γ.

[14]The derivation of this result is standard for one sector endogenous growth models as presented here. See e.g. Barro and Sala–i–Martin (1995), ch. 4.

Substitute $G_t = (1-\lambda)\tau K_t$ in (1). Recalling $L_t = 1$ and taking logarithms and time derivatives yields $\frac{\dot{Y}_t}{Y_t} = \frac{\dot{K}_t}{K_t} = \frac{\dot{G}_t}{G_t}$. Hence, the economy is characterized by *balanced growth* with $\gamma = \frac{\dot{c}^k_t}{c^k_t} = \frac{\dot{k}_t}{k_t} = \frac{\dot{c}^W_t}{c^W_t} = \frac{\dot{C}^k_t}{C^k_t} = \frac{\dot{K}_t}{K_t} = \frac{\dot{C}^W_t}{C^W_t} = \frac{\dot{G}_t}{G_t} = \frac{\dot{Y}_t}{Y_t}$ where $\gamma = (r-\tau) - \rho$ and $r = \alpha A[(1-\lambda)\tau]^{1-\alpha}$.

3.1 (Re-)Distribution

We follow Alesina and Rodrik and call "redistribution" any policy that distributes *income* to the non-accumulated factor of production while reducing the incentive to accumulate. Thus, they assess income redistribution relative to growth maximizing policies.[15] Such a benchmark for assessing income distributions may have its virtues when taking into account that people appear to have difficulties disentangling the relationship between utility enhancing growth and the distribution of income. See e.g. Amiel and Cowell (1999).

For this paper's purposes we measure the income dispersion in terms of the ratio of the factor income of the representative capital owner to the factor income of the representative worker. For pre-tax (and pre-transfer) incomes it is given by $F^g = \frac{r\frac{K}{n}}{\eta\frac{K}{l}} = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{l}{n}\right)$. Notice that F^g is independent of policy and the level of technology, A.

For after-tax incomes one gets

$$F = \frac{(r-\tau)\frac{K}{n}}{[\eta + \lambda\tau]\frac{K}{l}} = \left(\frac{r-\tau}{\eta+\lambda\tau}\right)\left(\frac{l}{n}\right). \tag{11}$$

Under a growth maximizing policy $\hat{\tau} = [\alpha(1-\alpha)A]^{\frac{1}{\alpha}}$ and $\lambda = 0$ so that $\hat{r} - \hat{\tau} = \hat{\tau}\left(\frac{\alpha}{1-\alpha}\right)$. Furthermore, $\eta = (1-\alpha)A\hat{\tau}^{1-\alpha}$. Thus, the post-tax factor

[15] In terms of personal income distributions one may ask why those policies should serve as a benchmark. For example, it may well be the case that moving from a growth maximizing to some other policy may increase income inequality in personal incomes and decrease growth. Most people would assess such a redistributing policy shift with reference to a policy that grants equal incomes per person. This is done in the working paper version of this paper. See Rehme (2000). However, here *redistribution* is defined in terms of factor incomes relative to growth maximizing policies. In particular, it is defined as taking real resources (*wealth*) from the owners of the accumulated factor of production by giving them to the owners of the non-accumulated factor of production in comparison of a situation where factor incomes are induced by a growth maximizing policy..

income distribution under that policy is

$$\hat{F} = \left(\frac{\hat{\tau} \left(\frac{\alpha}{1-\alpha} \right)}{(1-\alpha) A \hat{\tau}^{1-\alpha}} \right) \left(\frac{l}{n} \right) = \left(\frac{\hat{\tau}^{\alpha} \left(\frac{\alpha}{1-\alpha} \right)}{(1-\alpha) A} \right) \left(\frac{l}{n} \right) = \left(\frac{\alpha^2}{1-\alpha} \right) \left(\frac{l}{n} \right),$$

which is independent of A. Given this we define "redistribution" as the difference between the ratio of the income dispersion F generated under some policy (τ, λ) in comparison to the income dispersion $\hat{F}$, generated under growth maximization, in terms of $\hat{F}$. Thus,

$$\Pi \equiv \frac{\hat{F} - F}{\hat{F}} = 1 - \frac{F}{\hat{F}} = 1 - \left(\frac{r - \tau}{\eta + \lambda \tau} \right) \left(\frac{1-\alpha}{\alpha^2} \right). \tag{12}$$

We concentrate on cases where $\tau \geq \hat{\tau}$ with possibly $\lambda > 0$. Then for any such τ and λ with $\tau(1-\lambda) > \hat{\tau}$ we have $(r-\tau) \leq (\hat{r} - \hat{\tau})$ and $\eta(\tau(1-\lambda)) + \lambda\tau > \hat{\eta}$. Thus, $\frac{F}{\hat{F}} < 1$ and so $\Pi > 0$ under such policies, which are the ones we find in equilibrium. Hence, any increase (decrease) in Π from a situation with $\Pi \geq 0$ means that redistribution increases (decreases). Intuitively, an increase in redistribution is associated with relatively more income going to the workers relative to the income going to them under the growth maximizing policy.

4 Exogenous Policy

In this section we analyze the effects of changes in policy on growth and redistribution when policy is set exogenously, that is, when potentially important economic determinants of policy are ignored when choosing policy.

For the growth rate we have

$$\gamma = (r - \tau) - \rho = (\alpha A[(1-\lambda)\tau]^{1-\alpha} - \tau) - \rho. \tag{13}$$

Lemma 1 *Given everything else, growth is first increasing and then decreasing in τ and maximized when $\lambda = 0$ and $\tau = \hat{\tau} \equiv [\alpha(1-\alpha)A]^{\frac{1}{\alpha}}$. Direct transfers ($\lambda > 0$) to the workers for given taxes imply lower growth.*

Thus, if relatively high taxes - for example, for granting direct transfers to the workers - are levied, then growth is lower[16] when $\tau \gtrless \hat{\tau}$. For given taxes, transferring resources to the workers (higher λ) lowers growth, because these resources will not be used productively and will reduce the marginal products.

For exogenous policy changes one obtains that the reaction of redistribution Π in equation (12) depends on the sign of $\frac{\partial F}{\partial \lambda}_{|\tau}$ which hinges on[17]

$$\begin{aligned} & \frac{l}{n}\left[(r_\lambda - 1)(\eta + \lambda\tau) - (\eta_\lambda + \lambda)(r - \tau)\right] \\ = \; & r_\lambda \cdot \eta + r_\lambda \cdot \lambda \cdot \tau - \eta_\lambda \cdot r + \eta_\lambda \cdot \tau - \tau \cdot (r - \tau). \end{aligned}$$

As $r_\lambda \eta = \eta_\lambda r$ and $r_\lambda, \eta_\lambda < 0$ it follows that the expression is negative and, hence, $\frac{\partial F}{\partial \tau}_{|\lambda} < 0$ and so $\frac{\partial \Pi}{\partial \tau}_{|\lambda} > 0$. Similarly, for the sign of $\frac{\partial F}{\partial \lambda}_{|\tau}$ one obtains

$$\frac{l}{n}\left[(r_\tau - 1)(\eta + \lambda\tau) - (\eta_\tau + \lambda)(r - \tau)\right]$$

which is also negative.[18] Hence, $\frac{\partial \Pi}{\partial \lambda}_{|\tau} > 0$. But then redistribution Π satisfies $\frac{\partial \Pi}{\partial \tau}_{|\lambda} > 0$ and $\frac{\partial \Pi}{\partial \lambda}_{|\tau} > 0$.

Lemma 2 *Given everything else, redistribution* Π *increases, when the direct transfers* λ *and/or the tax rate* τ *are increased.*

Thus, there is a monotonic relationship between taxes and redistribution in the model. In that sense higher tax rates than $\hat{\tau}$ always redistribute income towards the non-accumulated factor of production.

4.1 Implications

If we ignore important determinants of economic policy, the model yields the following predictions about observable associations between redistribution

[16] This is because less G_t is available to raise the marginal products when the costs of financing G_t rise too much. Thus, the optimal G is attained when $\tau = \hat{\tau}$ and, in terms of output, the marginal cost of providing public inputs equals their marginal benefit.

[17] We denote partial derivatives by a subscript from now on. Thus, x_A denotes $\frac{\partial x}{\partial A}$.

[18] Notice that $[(r_\tau - 1)(\eta + \lambda\tau) - (\eta_\tau + \lambda)(r - \tau)] = r_\tau \cdot \eta + r_\tau \cdot \lambda \cdot \tau - \eta - \lambda \cdot \tau - \eta_\tau \cdot r + \eta_\tau \cdot \tau - \lambda \cdot r + \lambda \cdot r$ where $r_\tau \cdot \eta = \eta_\tau \cdot r$. Furthermore, $r_\tau \cdot \lambda \cdot \tau < \lambda \cdot r$ and $\eta_\tau \cdot \tau < \eta$. Thus, the expression for $\frac{\partial F}{\partial \lambda}_{|\tau}$ is negative.

and growth on the one hand, and tax rates and growth on the other. In order to relate to the next sections we focus on tax rates such that $\tau > \hat{\tau}$ which capture the conventional wisdom of a negative association between growth and redistribution, resp. tax rates. The graphs below visualize the simple relationships and follow straightforwardly from lemma (1) and (2).

Figure 1: Growth and Taxes

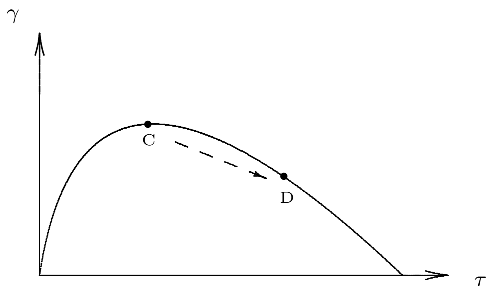
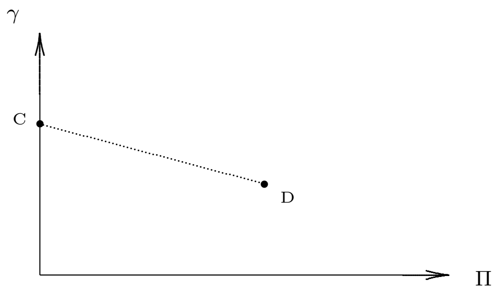


Figure 2: Growth and Redistribution



Point C in figure (1) and (2) depicts the growth maximizing policy with $\tau = \hat{\tau}$. According to our definition of redistribution point C in figure 2 then implies a situation of no redistribution, that is, $\Pi = 0$. An increase in taxes (or higher λ) means that we move towards point D with lower growth, but higher redistribution. Thus, the predicted relationship under the assumption that everything else is given also entails that one should observe a negative relationship between redistribution and growth. Consequently, we should observe a negative tradeoff between redistribution and growth. For any empirical analysis for a cross-section of countries this prediction would hold under the maintained assumption that the countries are structurally homogeneous and technologically similar.

5 Optimizing Governments

Next, we put structure on policy by introducing optimal behaviour of governments. That allows us to identify which policy is chosen by a government having particular preferences. Concentrating on simple policy determination to make the link, we will find that some governments will prefer $\hat{\tau}$ as their optimal choice, while others may prefer higher taxes, $\tau > \hat{\tau}$.

To this end consider governments that represent a representative worker or capital owner.[19] The intertemporal welfare of an entirely pro-capital, V^r, resp. entirely pro-labour government, V^l, is then given by (see Appendix B)

$$V^r(c_t^k) = \frac{\ln(\rho k_0)}{\rho} + \frac{\gamma}{\rho^2} \quad \text{and} \quad V^l(c_t^W) = \frac{\ln\left[(\eta(\tau,\lambda) + \lambda\tau)\tilde{k_0}\right]}{\rho} + \frac{\gamma}{\rho^2}. \quad (14)$$

The governments respect the right of private property and maximize the welfare of their clientele under the condition $\lambda \geq 0$. That restricts the governments in that even a pro-capital government does not tax workers, because a negative λ would effectively amount to a tax on wages.

This assumption are invoked for the following reasons: One easily verifies that an entirely pro-capital government would optimally choose λ such that the workers would receive zero after-tax-and-transfer income. We rule out such 'enslavement' of raw labour, because it is not a viable political option in most countries - at least not anymore. Furthermore, we rule out expropriation of capital for the governments. Although a command optimum in the model would involve expropriation of capital even for a government maximizing the welfare of the capital owners, it is ignored as a policy option because it is not very common in the real world.

The optimal *pro-labour* policy is derived in Appendix C and is given by

$$\text{If} \quad \rho \geq [(1-\alpha)A]^{\frac{1}{\alpha}} \quad \text{then:} \quad \tau = \rho, \qquad \lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\rho}. \quad (15)$$

[19]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) and Roemer (2006).

If $\rho < [(1-\alpha)A]^{\frac{1}{\alpha}}$ then:

$$\tau[1-\alpha(1-\alpha)A\tau^{-\alpha}] = \rho(1-\alpha), \qquad \lambda = 0\,. \tag{16}$$

Let $\check{\tau}$ solve these equations. For a wide range of parameter values the pro-labour government chooses not to grant direct (unproductive) transfers to the workers in the optimum. In particular, there will be no direct transfers when the workers are sufficiently patient as in (16). In that case the beneficial effects of a higher growth rate coupled with relatively lower taxes outweigh the intertemporal welfare obtained by higher taxes and lower growth when the workers are relatively impatient. Thus, the agents may trade off higher consumption now against higher growth and so higher consumption in the future.

Furthermore, the conditions on the parameters imply that if the government grants direct transfers and so chooses a positive λ, then equation (15) implies $(1-\lambda)\check{\tau} = [(1-\alpha)A]^{\frac{1}{\alpha}}$. If $\Theta \equiv (1-\alpha)A$, then

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

But $\check{\tau} = \rho \geq \Theta^{\frac{1}{\alpha}}$ in (15), and $\gamma > 0$ requires $r - \check{\tau} - \rho > 0$. So $\check{\tau}$ has to satisfy

$$\check{\tau} > \Theta^{\frac{1}{\alpha}} \;\wedge\; \left(\frac{\alpha}{1-\alpha}\right)\Theta^{\frac{1}{\alpha}} > 2\check{\tau} \;\;\Leftrightarrow\;\; \check{\tau}\;\left(\frac{\alpha}{1-\alpha}\right)\Theta^{\frac{1}{\alpha}} > \Theta^{\frac{1}{\alpha}}\;2\check{\tau} \;\Leftrightarrow \alpha > \frac{2}{3}.$$

Thus, direct redistribution through transfers to the workers is only optimal when the share of capital α is sufficiently higher than that of public inputs or labour, i.e., when capital receives a relatively high pre-tax share in total income and consequently factor income inequality is relatively high.

Lemma 3 *Positive growth under a policy with direct transfers to the workers, $\lambda > 0$, requires $\alpha > \frac{2}{3}$.*

Thus, the model captures that direct transfers are optimally called for when the factor income distribution is very unequal.[20] This would correspond

[20]Notice that we contemplate a model with a broad definition of capital. Then it may well be the case that α is bigger than 0.5. See e.g. Mankiw, Romer, and Weil (1992).

to the conventional wisdom that redistribution policies that are in the interest of lesser privileged groups have that feature.

In contrast, the *pro-capital* government chooses

$$\tau = \hat{\tau} \equiv [\alpha(1-\alpha)A]^{\frac{1}{\alpha}} \quad \text{and} \quad \lambda = 0. \tag{17}$$

Thus, the pro-capital government does not directly transfer resources to the workers and acts growth maximizing in the model by granting the maximum after-tax return on capital.[21]

6 Economic Fundamentals and Policy

All the optimal policies derived above (see equations (15), (16), and (17)) depend on three fundamental variables, namely, the state of technology A, the rate of time preference, ρ, and the share of capital, α. The rate of time preference is not considered any further, because it is considered a variable that is hard to measure.[22] Most researchers find that there is not much variability in the capital share, α, over time and across countries.[23] For that reason we will not focus on it either. However, the efficiency index A is of prime interest in many models of growth and distribution. In this paper we also take A to be a prime determinant of policy, especially when comparing

[21] In Appendix C it is also shown that any government that attaches more social weight on the capitalists' welfare would choose taxes closer to $\hat{\tau}$ leading to higher growth. All the subsequent results would then hold in relative terms.

[22] It should be noted, however, that the model would imply an interesting result as regards patience. Political preferences alone do not rule out the possibility of choosing a high growth policy. Trivially, but importantly a pro-labour government would mimic a growth maximizing policy if the workers are very patient, that is, if ρ goes to zero in equation (16). This is a special, but interesting case. It means that a government which places maximal weight on the non-accumulated factor of production may act like a growth maximizer if it puts almost equal weight on the welfare of future generations (low ρ). The two policies coincide only if $\rho \to 0$ which causes problems for the convergence of the utility indices. For any empirical observation, however, it suffices that ρ be very low while the utility indices still converge. In this case the *measured* tax and growth rates under an optimal pro-labour, pro-capital or a growth maximizing policy would be *observationally* indistinguishable.

[23] See e.g. Barro and Sala–i–Martin (1995), Mankiw, Romer, and Weil (1992), p. 341, Sachs (1979), Table 3, or more recently Gollin (2002).

countries. For that reason we will focus on the role of effects of A on policy and growth in the rest of the paper. Thus, we take policy to be endogenous by depending on the important economic fundamental A.

This merits a clarifying remark: Many people argue that A is essentially unobservable. Others find that A is very hard to measure. This is almost common knowledge in applied work. Without further, sometimes strong assumptions it is hard to pin down A empirically. Here I will *not* make these assumptions. I will simply assume that empirical indicators (or proxies) of A will produce the same predictions as in this paper, if they are monotonically related to A and not (directly) under the control of any decision maker. That various factors may bear on A is precisely what calls for a theory about A. See, for example, Prescott (1998). But if the theoretical "lens" provided by this paper's model is used, any empirical indicator that is monotonically related to this model's A would suffice to support the model's predictions and, thus, render them testable. For that reason an analysis of the effects of (perhaps hard to measure, but conventionally thought to be important) A on policy and growth should be of sufficient interest.[24]

6.1 Exogenous Policy and Economic Efficiency

Next, we look at the effects of a change in aggregate efficiency A on growth and distribution for an arbitrary and a growth maximizing policy. To this end we fix the policy choices τ and $\hat{\tau}$ and look what happens when aggregate efficiency changes. The arbitrary tax rate may coincide with $\tau = \check{\tau}$, that is, the optimal policy of an entirely pro-labour government but the results apply to more general policy choices.

That exercise yields the following: The growth rate in (13) implies that for given policy an increase in A raises growth and implies an upward shift of the concave relationship between taxes and growth.

[24] Notice that in terms of conventional production-function-based explanations the paper implies a simple theory about one determinant of aggregate efficiency, namely that captured by the influence of productive inputs in production G. The indicator A then represents the unexplained part of (total) aggregate efficiency in a production function framework as in, for instance, Hall and Jones (1999). However, formulating a detailed theory about A is beyond the scope of this paper.

Lemma 4 *An increase in efficiency raises growth for given policy.*

The result that a more efficient economy has higher growth corresponds to common economic intuition. The important point here is that this result holds for given policy. Since higher taxes, when set exogenously, imply lower growth, there is an interesting trade-off between the effect of taxes or efficiency on growth.

Next, we look at redistribution. Clearly, under the growth maximizing policy $\Pi = 0$ and is independent of A. For other policies we obtain that the sign of $\frac{\partial \Pi}{\partial A}$ depends on the sign of $\frac{\partial F}{\partial A}$ and, thus, on

$$r_A \cdot (\eta + \lambda\tau) - \eta_A \cdot (r - \tau) = r_A \cdot \eta + r_A \cdot \lambda \cdot \tau - \eta_A \cdot r + \eta_A \cdot \tau. \qquad (18)$$

One verifies that $\eta_A r = \eta_A r$ and $r_A, \eta_A > 0$ for given policy (τ, λ) so that the expression in (18) is positive. Thus,

Lemma 5 *If there is redistribution,* $\Pi > 0$*, an increase in efficiency* A *for given policy reduces redistribution, i.e.* $\frac{\partial \Pi}{\partial A} < 0$*. Under a growth maximizing policy,* $\Pi = 0$*, a change in efficiency does not have any effect on redistribution.*

The reason for this interesting, rather surprising result is the following: Higher efficiency raises both the *pre-tax* return on capital and pre-tax wages, but the relative change would be the same for both types of income and so $rk/\eta k$ does not change at all. In turn, the relative change would be different for the *after-tax* incomes. The relative tax burden borne by the accumulated factor of production, $\tau k/rk$, would fall for a given policy. As $r = \alpha/(1-\alpha)\eta$, the relative tax burden in terms of the wage rate would also fall. Then the relative income gain induced by higher efficiency accrues relatively more to the capital owners. Thus, when holding policy constant, the capitalists' after-tax income rises by relatively more than the workers' post-tax income. Hence, for given policy relatively more income would be distributed towards capital when aggregate efficiency rises.

6.2 Implications

Lemma (1), (2), (4) and (5) imply for given policies with $\tau > \hat{\tau}$ that there is an interesting tradeoff between between policy and economic efficiency. For a given level of efficiency higher taxes reduce growth and imply more redistribution, whereas for given policy an increase in efficiency raises growth and implies lower redistribution.

This suggests that for empirical analyses that across countries the association between taxes and growth may not be clear-cut. For example, we may observe countries with high tax rates, but coupled with high efficiency these countries may also feature relatively high growth.

In contrast, in a cross-section of countries the association between redistribution and growth is likely to be observed to be negative. However, this will crucially depend on the levels of the tax rates and economic efficiency.

7 Endogenous Policy and Growth

Now we relax the previous assumption that $\check{\tau}$ and $\hat{\tau}$ are fixed and expressly allow the optimal tax rates to vary in response to changes in aggregate efficiency A. To this end we consider the policy choices for entirely pro-labour and entirely pro-capital policies as given in equations (15), (16), and (17). The relationship between tax rates and growth as well as between redistribution and growth under the optimal pro-capital or pro-labour policies will be analyzed in turn.

7.1 Endogenous taxes and growth

For the optimal pro-capital policy one easily verifies from equation (17) that $\frac{d\hat{\tau}}{dA} > 0$. Thus, the optimal tax rate would higher under that policy.

Furthermore, the maximum after-tax return, $\hat{r} - \hat{\tau} = \hat{\tau}\left(\frac{\alpha}{1-\alpha}\right)$, is increasing in A since $\frac{d(\hat{r}-\hat{\tau})}{dA} = \left(\frac{\alpha}{1-\alpha}\right)\frac{d\hat{\tau}}{dA}$. Hence, $\frac{d\hat{\gamma}}{dA} > 0$ as well. Thus, an increase in efficiency raises the maximum after-tax return, the growth maximizing tax rate and maximum growth under the optimal pro-capital policy.

Next, turn to a pro-labour government that does not directly transfer resources to the workers, $\lambda = 0$. The effect of an increase in A on optimal taxes in (16) is

$$\left(1 - \alpha(1-\alpha)^2 A\tau^{-\alpha}\right) d\tau - \left(\alpha(1-\alpha)\tau^{1-\alpha}\right) dA = 0$$
$$\frac{d\tau}{dA} = \alpha(1-\alpha)\tau\left(\tau^{\alpha} - \alpha(1-\alpha)^2 A\right)^{-1}. \tag{19}$$

As $\check{\tau} > \hat{\tau}$, the expression is positive.[25] Hence, an increase in efficiency makes a non-redistributing, pro-labour government increase its optimal tax rate. Next, $\frac{d\gamma}{dA} = r_A + (r_\tau - 1)\frac{d\tau}{dA} > 0$ if

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

which is equivalent to $1 > 1 - \alpha$ and true. Thus, $\frac{d\check{\gamma}}{dA} > 0$ if $\lambda = 0$ in (16).

Suppose the government chooses positive direct transfers to the workers. Then equation (15) implies $\frac{d\check{\tau}}{dA} = 0$ and $\frac{d\check{\lambda}}{dA} < 0$. Then $\frac{d\gamma}{dA} = r_A + r_\lambda \frac{d\lambda}{dA} > 0$ since $r_\lambda < 0$ and $r_A > 0$.

Proposition 1 *The optimal policies of a pro-capital or pro-labour government imply that higher efficiency leads them to choose either higher taxes when $\lambda = 0$ or the same taxes and lower transfers. An increase in efficiency leads to higher growth under the optimal pro-capital and pro-labour policy.*

Hence the beneficial effect of public resources is enhanced when economies are more efficient. This is taken account of when government choose their optimal policies. Thus, the beneficial effect of higher efficiency is *not* counteracted by choosing taxes so that growth is reduced. In that sense there is a tradeoff between taxes and efficiency which is solved in the optimum so that the beneficial effects of higher efficiency are maintained. Observationally it would imply that in a cross-section of countries we would observe that

[25] To see this notice that $\frac{d\tau}{dA} > 0$ requires $\tau^{\alpha} > \alpha(1-\alpha)^2 A$ which is equivalent to $\tau > \hat{\tau}(1-\alpha)^{\frac{1}{\alpha}}$ and always satisfied since $\check{\tau} > \hat{\tau}$ and $(1-\alpha)^{\frac{1}{\alpha}} < 1$.

more efficient counties would choose higher taxes, but would also have higher growth.

7.2 Endogenous redistribution and growth

Clearly, under the pro-capital policy $\Pi = 0$ which is independent of aggregate efficiency A. Under the pro-labour policy, in contrast, we obtain the following expressions for redistribution, $\Pi = 1 - \frac{\check{F}}{\hat{F}}$:[26]

$$\Pi_{|\lambda=0} = 1 - \frac{1}{\alpha}\left(1 - \frac{\check{\tau}^{\alpha}}{\alpha A}\right) \quad \text{and} \quad \Pi_{|\lambda>0} = 1 - \frac{1}{\alpha}\left(\frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\rho} - \frac{1-\alpha}{\alpha}\right).$$

Clearly, $\frac{d\Pi_{|\lambda>0}}{dA} < 0$. For $\frac{d\Pi_{|\lambda=0}}{dA}$ it suffices to check whether $\frac{\check{\tau}^{\alpha}}{\alpha A}$ is increasing or decreasing in A. The reaction of this term depends on the sign of

$$\alpha \check{\tau}^{\alpha-1} \frac{d\check{\tau}}{dA} \alpha A - \alpha \check{\tau}^{\alpha}$$

which is negative from equation (19) since $\frac{d\check{\tau}}{dA} < \frac{\check{\tau}}{\alpha A}$ is equivalent to

$$\alpha(1-\alpha)\check{\tau}\left(\check{\tau}^{\alpha} - \alpha(1-\alpha)^2 A\right)^{-1} < \frac{\check{\tau}}{\alpha A} \quad \Leftrightarrow \quad \alpha(1-\alpha)A < \check{\tau}^{\alpha}$$

and true, because $\hat{\tau}^{\alpha} = \alpha(1-\alpha)A$ and $\check{\tau} > \hat{\tau}$. Hence, $\frac{d\Pi_{|\lambda=0}}{dA} < 0$.

Proposition 2 *When policy is endogenous and political preferences are fixed, an increase in efficiency leads to higher taxes but generally does not imply more and sometimes implies less redistribution.*

Thus, the two counteracting effects of taxes and efficiency are such that the effect of efficiency dominates are regards redistribution. That means that according to the model no optimizing governments - not even optimizing,

[26] For this notice that $\frac{\check{F}}{\hat{F}}$ when $\lambda = 0$ boils down to $\frac{\check{F}}{\hat{F}} = \left(\frac{\check{r}-\check{\tau}}{\check{\eta}}\right)\left(\frac{1-\alpha}{\alpha^2}\right)$ which reduces to $\Pi_{|\lambda=0}$ since $\check{r} = \alpha A \check{\tau}^{1-\alpha}$ and $\check{\eta} = (1-\alpha)A\check{\tau}^{1-\alpha}$. For $\frac{\check{F}}{\hat{F}}$ when $\lambda > 0$ notice that $\check{F} = \left(\frac{\check{r}-\check{\tau}}{\check{\eta}+\check{\lambda}\check{\tau}}\right)$. But $\check{\eta} + \check{\lambda}\check{\tau} = (1-\alpha)A[\check{\tau}(1-\check{\lambda})]^{1-\alpha} + \check{\lambda}\check{\tau}$ as well as $\check{\tau} = \rho$ and $\lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\rho}$. Thus, $\check{\eta} + \check{\lambda}\check{\tau} = (1-\alpha)A[(1-\alpha)A]^{\frac{1}{\alpha}-1} + \rho - [(1-\alpha)A]^{\frac{1}{\alpha}} = \rho$. Substituting in $1 - \frac{\check{F}}{\hat{F}}$ yields the expression for $\Pi_{|\lambda>0}$.

strictly pro-labour governments - would ever choose a policy that counteracts the positive effect of higher efficiency on growth. Even governments that like redistribution would not sacrifice the beneficial effect of higher growth on intertemporal welfare by increasing taxes too much, when efficiency increases.

7.3 Implications

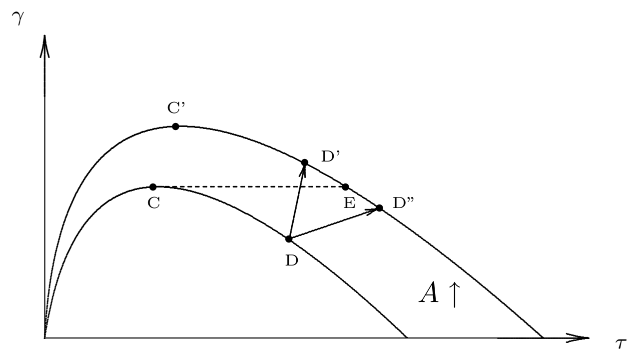
Proposition (1) and (2) have important consequences for empirical work. Suppose we have data on taxes, redistribution and growth for a cross-section of countries. The efficiency index A is very hard to observe and many assumptions have to be made to measure it. Thus, suppose we do not have data on A. But we know that A has an impact on growth and policy. Then the model implies the following for observable policy variables:

The possible combinations of (observable) tax rate and growth combinations are depicted in figure (3). For countries that experience an increase in efficiency and given the optimal policy of a pro-capital or pro-labour government we see that for each of them separately we would observe a positive association between taxes and growth. For instance, comparing pro-capital policies under C or C' yields that the C' country would feature higher taxes and higher growth in comparison to the C country. The same holds for the optimal pro-labour policies when comparing D with D' or D". Thus, when political preferences are structurally fixed for long periods of time[27] and, in the course of development, A increases over time, then - in contrast to the assumption that policy is set exogenously - we would observe that *within* countries the association between tax rates and growth is positive.[28]

[27] Evidence for this for OECD countries can be found in, for example, Garrett (1998).

[28] This is in contrast to the prediction that higher taxes generally entail lower growth. Of course, that prediction was derived under the assumption that all countries have the same technology and policy is set exogenously.

Figure 3: Growth and Endogenous Taxes

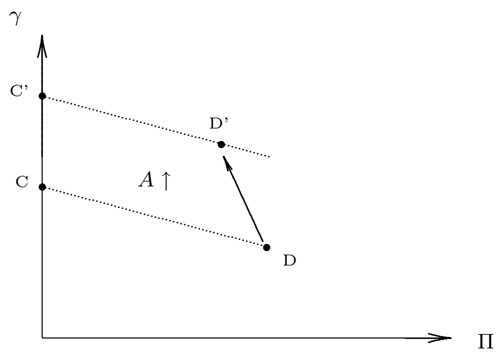


However, in a sample with countries where pro-labour or pro-capital policies are pursued this would not hold anymore. Suppose we take each point as reflecting some countries that are technologically similar. For instance, D may represent a set of countries that pursue pro-labour policies all sharing the same efficiency, let us say, A_1. When comparing these countries with those pursuing pro-capital policies at C' with A_2 and $A_2 > A_1$, then the observed relationship is likely to be negative. Thus, in a cross-section one would have to know the distribution of A and of the political preferences. Otherwise, no clear prediction would be possible.[29]

Furthermore, no clear relationship between redistribution and economic growth can be ascertained. This can be gleaned from figure (4).

[29] Clearer predictions are possible when additional assumptions are made. For instance, in Rehme (2006b), or Rehme (2006a) it is found that when countries are exposed to the danger of capital outflows in a world where capital is internationally highly mobile, then pro-labour governments would not find it optimal to choose a point such a D''. Thus, when a positive association between high, redistributive tax rates and growth is observed across countries, then it is likely that we have no countries in C', some countries that are in C and some other ones that are in D'.

Figure 4: Growth and Endogenous Redistribution



For countries with the same political preferences either no relationship between growth and redistribution can be established when e.g. comparing C and C' for pro-capital policies, or the observed relationship would be negative for pro-labour policies, when comparing D and D'.

However, we may also observe a positive relationship, when one has a sample of countries, where one set belongs to C and the other one to D'. In this case the redistributing countries at D' would have to be sufficiently more efficient than the pro-capital countries at C.[30]

Again it would be necessary to know the distribution of efficiency and of political preferences across countries and/or over time.

Proposition 3 *When growth and policy both depend on economically important fundamentals so that policy is economically endogenous, no clear relationship between policy and growth can in general be found by empirical analyses unless information on the distribution of these fundamentals across countries or over time is known.*

The application of this result will apply to almost any model that relates policy and growth to economic fundamentals when introducing optimizing behaviour of governments. This would seem like a rather negative result for

[30]This may indeed be the situation that is found in some empirical studies which find a positive association between growth and redistribution.

empirical cross-section research.[31]

However, the ambiguity of the policy-growth relationships resulting from differences in aggregate efficiency can in principle be resolved, once endogeneity of policy is explicitly acknowledged, and analyzed in detail, and the corresponding results are used to investigate the growth-policy nexus. Thus, the result may rather be a constructive one.

If one constructs good measures of A and for policy preferences, we can use theory to improve on existing empirical research on the policy-growth nexus. This should be doable. Thus, potential improvements should essentially be possible. In that sense the paper may be taken to convey a "good" message for empirical research.

8 Welfare

As a better technology affects long-run growth and policy under the optimal policies, this will also have an impact on welfare.[32] Thus, it is an interesting question what the relative, overall welfare implications of changes in A in a model with distributional conflict are.

From (14) one verifies that $0 < \frac{dV^r}{dA}_{|\hat{\tau}} < \frac{dV^l}{dA}_{|\tilde{\tau}}$. See Appendix D. Thus, in the model an advance in technology would benefit a pro-labour government relatively *more* in the long run than a pro-capital government.

Proposition 4 *Governments that represent the non-accumulated factor of production only and that wish to redistribute resources from the accumulated to the non-accumulated factor of production have a* <u>*relatively*</u> *greater incentive in the long run to have an economy with a superior technology than*

[31]For instance, it would lend some support to the recent arguments presented by Rodrik (2005), who shows that by means of conventional growth regression no clear relationship between policy and growth can in principle to uncovered. It would also support the argument put forth in, for example, Rehme (2002b) that the estimated coefficients in growth regressions relating growth and policy variables would in general be biased. Although the biases can be signed from theory, it would be difficult, but it would in principle be still possible to interpret existing empirical results or correct for the biases in applied work.

[32]Most of this section's results generalize to utility functions of the constant intertemporal elasticity of substitution type (CIES). A proof for this can be found in the addendum at the end of this paper.

governments representing the accumulated factor of production only.

The result suggests interesting long-run consequences of the effects of e.g. institutional reform on growth and welfare. Of course, things may be different in the short run when workers might have to learn new technologies or there is resistance to reform. For models studying these issues see e.g. Fernandez and Rodrik (1991), Helpman and Rangel (1999) or Canton, de Groot, and Nahuis (1999).

It is important to notice that Proposition 4 applies to governments. For *given* optimal policies the worker's or capital owner's welfare may react differently to changes in A. In this context Appendix D also shows the following

Proposition 5 *For given, optimal pro-capital or pro-labour policies the workers never benefit less from technical progress in the long run than the capital owners. Unless the pro-labour policy redistributes wealth, the workers benefit relatively more than the capital owners.*

That result allows for various interpretations. For instance, if the workers benefit relatively more from technical progress in the long run than the capital owners, they should be relatively more interested in innovations and should be willing to pay a higher (shadow) price for it. Such prices may, for instance, have to be paid for short-run (in the model pre-t_0) phenomena such as the pain to learn new technologies, short-run unemployment or any - perhaps - adverse effects on the income distribution.

9 Concluding Remarks

This paper presents a theoretical analysis to investigate the long-run relationship between public policies and growth. Within a framework that shares common features with many models it is shown that optimizing governments take account of fundamental economic variables when making their decisions. Thus, the paper focuses on endogenous policy.

It is shown that changes of fundamental economic variables have interesting effects on policies and through the latter on the growth rate and the income distribution. Several findings of the paper are noteworthy.

The effects of policy on growth are often different when policy is taken to be exogenous or endogenous. In the model there is a trade-off between higher, exogenously set taxes and economic efficiency. Given efficiency, higher taxes may reduce growth, but imply more redistribution. For given taxes, higher efficiency implies higher growth, but less redistribution.

When governments optimize and, thus, policy becomes endogenous, it turns out that no optimizing government would tax the beneficial effects of higher efficiency away. For given political preferences, though, the observed relationship between taxes and growth should be positive, and that between redistribution and growth should be negative.

However, the model's predictions have ambiguous implications for any cross-section analysis. The paper implies that it would be necessary - probably based on theory - to disentangle in more detail the relationship between policy and economic fundamentals as well as growth. This paper has made a move in that direction extending.

Another noteworthy result of the paper concerns welfare. An increase in technological efficiency generally raises taxes and growth, given particular political preferences. But higher efficiency also raises the agents' welfare under the optimal policies considered. Interestingly, the relative welfare gains are found to be often higher for the workers and always higher for a pro-labour government. Thus, workers may have a greater incentive that their economies be efficient.

Several caveats apply. Obviously, economic growth is influenced by many things. This paper simply argues that analyzing the long-run interplay of fundamental economic variables and public policy may provide useful insights about differences in growth and distribution experiences within or across countries. In this context, an empirical analysis of the present paper's predictions appears to be called for, but has been beyond the scope of this paper, and will be taken up in future research.

A The Capital Owners' Optimum

The necessary first order conditions for the maximization problem are given by equations (7), (8) and

$$\frac{1}{c_t^k} - \mu_t = 0 \tag{A1a}$$

$$\dot{\mu_t} = \mu_t \rho - \mu_t \, (r - \tau) \tag{A1b}$$

$$\lim_{t \to \infty} \; k_t \mu_t e^{-\rho t} = 0. \tag{A1c}$$

where μ_t is a positive co-state variable. From equation (A1a), (A1b) it follows that $\frac{\dot{c_t^k}}{c_t^k} = (r - \tau) - \rho$. Furthermore, for constant τ and from the transversality condition (A1c) and the budget constraint (7) it follows that $\frac{\dot{c_t^k}}{c_t^k} = \frac{\dot{k_t}}{k_t}$. Thus, $\gamma = \frac{\dot{c_t^k}}{c_t^k} = \frac{\dot{k_t}}{k_t}$.

B Welfare

The agents' welfare is $V^r = \int_0^t \ln c_t^k \; e^{-\rho t}$ and $V^l = \int_0^t \ln c_t^W \; e^{-\rho t}$. Let $t \to \infty$ and use integration by parts. For this define $v_2 = \ln c_t^j$, $dv_1 = e^{-\rho t} dt$ where $j = k, W$. Recall that $dv_2 = \frac{\dot{c_t^j}}{c_t^j} = \gamma$ for $j = k, W$ and constant in steady state. Then $v_1 = -\frac{1}{\rho} \; e^{-\rho t}$ so that

$$\begin{aligned} \int_0^\infty \ln c_t^j \; e^{-\rho t} \; dt &= \frac{1}{\rho} \Big[- \; \ln c_t^j \; e^{-\rho t} \Big]_0^\infty + \frac{1}{\rho} \int_0^\infty \gamma \; e^{-\rho t} \; dt \\ &= \frac{\ln c_0^j}{\rho} - \frac{1}{\rho^2} \Big[\gamma \; e^{-\rho t} \Big]_0^\infty \end{aligned}$$

where $c_0^k = \rho k_0$ and $c_0^W = (\eta + \lambda \tau) \tilde{k_0}$. Evaluation at the particular limits yields the expressions in (14).

C Optimal Policies

The government solves: $\max_{\tau, \lambda} \; (1 - \beta) \, V^r + \beta \, V^l$ s.t. $\lambda \geq 0$ where β is the social weight attached to each group's 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

$$\frac{\eta_\tau + \lambda}{\eta_\lambda + \tau} = \frac{\gamma_\tau}{\gamma_\lambda}. \tag{C1}$$

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

$$\tau = \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{1-\lambda}. \tag{C2}$$

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. (C2) 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,

$$\tau = \beta\rho \quad \text{and} \quad \lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\beta\rho}. \tag{C3}$$

which is equation (15) 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

$$(1-\alpha)\beta\rho = \tau\left[1 - \alpha(1-\alpha)A\tau^{-\alpha}\right] \tag{C4}$$

which holds only if $\beta\rho < [(1-\alpha)A]^{\frac{1}{\alpha}}$. For $\beta = 1$ this is equation (16) in the text.

If $\beta = 0$ it follows that $\lambda = 0$, $\gamma_\tau = r_\tau - 1 = 0$ and $\tau = \hat{\tau}$. Thus, the pro-capital government acts growth maximizing in the model.

Lemma $\gamma(\tau)$ is inversely related to β.

Proof: $\gamma_\tau < 0$ for $\check{\tau} > \hat{\tau}$ in (15) and (16). Also $\gamma(\tau) = \alpha A\left((1-\lambda)\tau\right)^{1-\alpha} - \tau - \rho$. Clearly, if $\lambda > 0$, then $\frac{d\check{\tau}}{d\beta} > 0$ in (C3), and $(1-\lambda)\tau = [(1-\alpha)A]^{\frac{1}{\alpha}}$. Thus, $\frac{d\gamma}{d\beta} < 0$.

Suppose $\beta > 0$ and $\lambda = 0$. Then $\check{\tau}$ is given as in (C4) so that by the implicit function theorem $\frac{d\check{\tau}}{d\beta} > 0$. Thus, $\frac{d\gamma}{d\beta} = \gamma_\tau \frac{d\check{\tau}}{d\beta} < 0$ which proves the lemma.

D Technology Effects on Welfare

Under the *optimal* policies the welfare in (14) is given by $V^i(A, \tau(A), \lambda(A))$ where $i = l, r$. An increase in A changes welfare by

$$dV^i = \frac{\partial V^i}{\partial A}\, dA + \frac{\partial V^i}{\partial \tau}\, \frac{\partial \tau}{\partial A}\, dA + \frac{\partial V^i}{\partial \lambda}\, \frac{\partial \lambda}{\partial A}\, dA.$$

By the *envelope theorem* $\frac{\partial V^r}{\partial \tau} = 0$ under the optimal pro-capital policy and $\frac{\partial V^l}{\partial \tau} = \frac{\partial V^l}{\partial \lambda} = 0$ under the optimal pro-labour policy. Thus,

$$\frac{dV^r}{dA}_{|\hat{\tau}} = \frac{\partial \hat{\gamma}}{\partial A_{|\hat{\tau}}}\left(\frac{1}{\rho^2}\right) \quad \text{and} \quad \frac{dV^l}{dA}_{|\check{\tau},\lambda} = \frac{\partial(\eta + \lambda\check{\tau})}{\partial A}_{|\check{\tau},\lambda}\left(\frac{1}{(\eta + \lambda\check{\tau})\rho}\right) + \frac{\partial \check{\gamma}}{\partial A_{|\check{\tau},\lambda}}\left(\frac{1}{\rho^2}\right).$$

Notice that under the optimal pro-labour policy $(\eta + \lambda\check{\tau}) = \rho$ when $\lambda > 0$ and $\frac{\partial \eta}{\partial A_{|\check{\tau},\lambda=0}}\left(\frac{1}{\eta\rho}\right) > 0$ when $\lambda = 0$. But $\frac{\partial \hat{\gamma}}{\partial A_{|\hat{\tau}}} = \alpha\hat{\tau}^{1-\alpha} < \frac{\partial \check{\gamma}}{\partial A_{|\check{\tau},\lambda}} = \alpha\check{\tau}^{1-\alpha}$ because $\check{\tau} > \hat{\tau}$. Hence, $0 < \frac{dV^r}{dA}_{|\hat{\tau}} < \frac{dV^l}{dA}_{|\check{\tau},\lambda}$ so that a *government* representing the average worker benefits relatively more than a *government* representing the average capital owner.

Quite another question is how each *individual*'s welfare is affected by a change in A *given* the *optimal* policies. For instance, under the optimal *pro-capital* policy an increase in A implies $\frac{dV^r}{dA}_{|\hat{\tau}} = \frac{\partial\gamma}{\partial A}\frac{1}{\rho^2}$ for a capital owner. For the worker it implies

$$\frac{dV^l}{dA}_{|\hat{\tau}} = \frac{\partial V^l}{\partial A} + \frac{\partial V^l}{\partial\tau}\frac{\partial\tau}{\partial A} = \left(\frac{\partial\eta}{\partial A} + \frac{\partial\eta}{\partial\tau}\frac{\partial\tau}{\partial A}\right)\frac{1}{\eta\rho} + \left(\frac{\partial\gamma}{\partial A} + \frac{\partial\gamma}{\partial\tau}\frac{\partial\tau}{\partial A}\right)\frac{1}{\rho^2}.$$

When $\tau = \hat{\tau}$ none of these derivatives is negative so that $\frac{dV^r}{dA}_{|\hat{\tau}} < \frac{dV^l}{dA}_{|\hat{\tau}}$ and a worker benefits more from technical progress than a capital owner under a pro-capital policy.

Under the optimal $\lambda = 0$, pro-labour policy and using the envelope theorem the welfare changes are

$$\frac{dV^r}{dA}_{|\check{\tau}} = \left(\frac{\partial\gamma}{\partial A} + \frac{\partial\gamma}{\partial\tau}\frac{\partial\tau}{\partial A}\right)\frac{1}{\rho^2} \quad \text{and} \quad \frac{dV^l}{dA}_{|\check{\tau}} = \frac{\partial\eta}{\partial A}\frac{1}{\eta\rho} + \frac{\partial\gamma}{\partial A}\frac{1}{\rho^2}.$$

As $\frac{\partial\gamma}{\partial\tau} < 0$ and $\frac{\partial\tau}{\partial A} > 0$ when $\tau = \check{\tau}$ it follows that $\frac{dV^r}{dA}_{|\check{\tau}} < \frac{dV^l}{dA}_{|\check{\tau}}$.

Under a $\lambda > 0$, pro-capital policy $(\eta + \lambda\check{\tau}) = \rho$ so that changes in A only affect γ in V^i in (14). But then $\frac{dV^r}{dA}_{|\check{\tau},\lambda} = \frac{dV^l}{dA}_{|\check{\tau},\lambda}$ and the workers and capital owners would benefit equally.

References

Aghion, P., E. Caroli, and C. García-Peñalosa (1999): "Inequality and Economic Growth: The Perspective of the New Growth Theories," *Journal of Economic Literature*, 37, 1615–1660.

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

Amiel, Y., and F. A. Cowell (1999): *Thinking about Inequality.* Cambridge University Press, Cambridge, UK.

Atkinson, A. B. (1983): *The Economics of Inequality.* Clarendon Press, Oxford, 2nd edn.

Atkinson, A. B., and F. Bourguignon (2000): "Introduction: Income Distribution and Economics," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 1, pp. 1–58. Elsevier, Amsterdam.

Barro, R. J. (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.

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.

Bénabou, R. (1996): "Inequality and Growth," in *NBER Macroeconomics Annual 1996*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 11–73. MIT Press, Cambridge and London.

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

——— (1999): "Macroeconomics of Distribution and Growth," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon. Elsevier Science, North-Holland.

Canton, E. J. F., H. L. F. de Groot, and R. Nahuis (1999): "Vested Interests and Resistance to Technology Adoption," Discussion Paper 99106, CentER, Tilburg.

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

Fernandez, R., and D. Rodrik (1991): "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty," *American Economic Review*, 81, 1146–1155.

Garrett, G. (1998): *Partisan Politics in the Global Economy.* Cambridge University Press, Cambridge, U.K.

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

Hall, R. E. (1988): “Intertemporal Substitution in Consumption,” *Journal of Political Economy*, 96, 339–357.

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.

Helpman, E., and A. Rangel (1999): “Adjusting to a New Technology: Experience and Training,” *Journal of Economic Growth*, 4, 359–383.

Jovanovic, B. (2000): “Growth Theory,” Working Paper 7468, NBER.

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

Landes, D. S. (1998): *The Wealth and Poverty of Nations. Why Some are so Rich and Some so Poor*. W. W. Norton, New York.

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.

Perotti, R. (1993): “Political Equilibrium, Income Distribution, and Growth,” *Review of Economic Studies*, 60, 755–776.

——— (1994): “Income Distribution and Investment,” *European Economic Review*, 38, 827–835.

——— (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.

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

Rehme, G. (2000): “Economic Growth and (Re-)Distributive Policies: A Comparative Dynamic Analysis,” Working Paper ECO 00/13, 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): “Economic Growth and (Re-)Distributive Policies in a Non-Cooperative World,” *Journal of Economics*, ??, forthcoming.

——— (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.

——— (2006): "Democracy, Education, and Equality," in *Graz Schumpeter Lectures*, Econometric Society Monographs; No. 40, Cambridge, U.K. Cambridge University Press.

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.

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.

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

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

Addendum

In this addendum I show that the welfare and the most important comparative static results of the paper generally carry over when employing the general class of constant intertemporal elasticity of substitution (CIES) utility functions.

The set-up is same as in the paper. In particular, everything up to the first part of section 5.2 is identical.

D.1 Capitalists

The representative capital owner maximizes his intertemporal utility according to

$$\max_{c_t^k} \int_0^\infty \frac{(c_t^k)^{1-\sigma} - 1}{1-\sigma}\, e^{-\rho t} dt, \quad \sigma > 0 \tag{1}$$

$$s.t. \qquad \dot{k_t} = (r-\tau)k_t - c_t^k \tag{2}$$

$$k(0) = \overline{k_0}, \quad k(\infty) = \text{free}, \tag{3}$$

where $k_t \equiv \frac{K_t}{n}$ denotes the capital stock per capital owner and $\nu \equiv \frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution. The growth rate of consumption and wealth is given by

$$\gamma \equiv \frac{\dot{c_t^k}}{c_t^k} = \frac{\dot{k_t}}{k_t} = \frac{(r-\tau) - \rho}{\sigma}. \tag{4}$$

Furthermore, from equations (9) and (7), and the usual transversality condition one verifies that $c_t^k = (r-\gamma)k_t$ with $c_t^k = c_0^k e^{\gamma t}$ is the capitalist's optimal level of consumption.

D.2 Workers

The representative worker's intertemporal welfare is given by

$$\int_0^\infty \frac{(c_t^W)^{1-\sigma} - 1}{1-\sigma}\, e^{-\rho t} dt \quad \text{where} \quad \sigma > 0, \tag{5}$$

$$c_t^W = (\eta(\tau,\lambda) + \lambda\tau)\,\tilde{k_t}, \quad \text{and} \quad \tilde{k_t} \equiv \frac{K_t}{l}$$

denotes the capital stock per worker and $\nu \equiv \frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution.

D.3 Market Equilibrium.

The market equilibrium is the same as in the paper with *balanced growth* at $\gamma = \frac{\dot{c_t^k}}{c_t^k} = \frac{\dot{k_t}}{k_t} = \frac{\dot{c_t^W}}{c_t^W} = \frac{\dot{C_t^k}}{C_t^k} = \frac{\dot{K_t}}{K_t} = \frac{\dot{C_t^W}}{C_t^W} = \frac{\dot{G_t}}{G_t} = \frac{\dot{Y_t}}{Y_t}$ where $\gamma = \frac{(r-\tau)-\rho}{\sigma}$ and $r = \alpha A[(1-\lambda)\tau]^{1-\alpha}$.

D.4 Optimizing Governments

Again assume that the governments are entirely pro-labour or pro-capital. The welfare function of a worker and capitalist in (1) and (5) in equilibrium satisfies

$$W^i(0) = \int_0^\infty \frac{(c_t^i)^{1-\sigma} - 1}{1-\sigma} e^{-\rho t} dt \quad \text{where} \quad i = k, W \tag{6}$$

$$c_t = c_0^i e^{\gamma t} \quad \text{and} \quad \gamma = \frac{(r-\tau)-\rho}{\sigma}.$$

Let $(r-\tau) \equiv R$ denote the after tax return for the rest of the analysis. Then one readily verifies that $W^i(0)$ integrates to[33]

$$W^i(0) = \frac{(c_0^i)^{1-\sigma}}{(1-\sigma)(\rho-(1-\sigma)\gamma)} - \frac{1}{\rho(1-\sigma)}, \quad \text{where} \quad i = k, W \tag{7}$$

and we assume that $\rho > (1-\sigma)\gamma$ for the integral to converge. (If $\sigma = 1$, then $W^i(0) = \frac{\ln c_0^i}{\rho} + \frac{\gamma}{\rho^2}$.)

Now for c_0^i with $i = W, k$ we have for the workers that

$$c_0^W = \eta \tilde{k}_0, \quad \text{where} \quad \tilde{k}_0 \equiv \frac{K_0}{l}. \tag{8}$$

For the capitalists we get[34]

$$c_0^k \;=\; (R-\gamma)k_0 = \left(R - \frac{R-\rho}{\sigma}\right)k_0 = \left(\frac{\sigma-1}{\sigma} R + \frac{\rho}{\sigma}\right)k_0 \tag{9}$$

where $k_0 \equiv \frac{K_0}{n}$.

Now let the governments optimize. For simplicity I will ignore the choice of λ. Thus, I set it equal to zero. Optimality for the choice of τ requires that

$$\frac{(1-\sigma)^2(c_0^i)^{-\sigma}\left(\frac{\partial c_0^i}{\partial \tau}\right)(\rho-(1-\sigma)\gamma) + (1-\sigma)^2\left(\frac{\partial \gamma}{\partial \tau}\right)(c_0^i)^{1-\sigma}}{(1-\sigma)^2(\rho-(1-\sigma)\gamma)^2} \;=\; 0$$

[33]To see this notice that $c_t^i = c_0^i e^{\gamma t}$ and $\gamma = \frac{R-\rho}{\sigma}$ imply

$$\int_0^\infty \frac{(c_t^i)^{1-\sigma}-1}{1-\sigma} e^{-\rho t}dt \;=\; \int_0^\infty \frac{(c_0^i\, e^{\gamma t})^{1-\sigma}-1}{1-\sigma} e^{-\rho t}dt$$
$$= \frac{(c_0^i)^{1-\sigma}}{1-\sigma}\int_0^\infty e^{(1-\sigma)\gamma t}\, e^{-\rho t}dt - \int_0^\infty \frac{1}{1-\sigma} e^{-\rho t}dt$$

Thus, we obtain

$$W^i(0) = \frac{(c_0^i)^{1-\sigma}}{1-\sigma}\int_0^\infty e^{[(1-\sigma)\gamma-\rho]t}dt - \frac{1}{\rho(1-\sigma)}.$$

Integrating and evaluating at the particular limits yields equation (7).

[34]Note that $\rho > (1-\sigma)\gamma \Leftrightarrow \rho > (1-\sigma)R$, ensuring $c_0^k > 0$.

$$\frac{(c_0^i)^{1-\sigma}\left[\left(\frac{\partial c_0^i}{\partial \tau}\frac{1}{c_0^i}\right)(\rho-(1-\sigma)\gamma)+\left(\frac{\partial \gamma}{\partial \tau}\right)\right]}{(\rho-(1-\sigma)\gamma)^2} = 0 \quad (10)$$

$$\left(\frac{\partial c_0^i}{\partial \tau}\frac{1}{c_0^i}\right)(\rho-(1-\sigma)\gamma)+\frac{\partial \gamma}{\partial \tau} = 0. \quad (11)$$

Now for the two groups this implies

$$\frac{\partial c_0^W}{\partial \tau} = \eta_\tau \tilde{k_0} = (1-\alpha)^2 A\tau^{-\alpha}\tilde{k_0} \quad \text{and} \quad \frac{\partial c_0^k}{\partial \tau} = (R_\tau - \gamma_\tau)k_0 = \left(\frac{\sigma-1}{\sigma}\right)R_\tau k_0.$$

We will concentrate on optimal tax rates that induce non-negative after-tax returns. So we rule out tax rates such that $(r-\tau)<0$.

The Pro-Capital Government

As $R_\tau = r_\tau - 1$, one easily verifies that the pro-capital government will set τ so that $R_\tau = 0$ and $\gamma_\tau = 0$, that is, $\tau = \hat{\tau}$. Thus, it chooses the growth maximizing tax rate $\hat{\tau} = [\alpha(1-\alpha)A]^{\frac{1}{\alpha}}$. We know that $\frac{d\hat{\tau}}{dA} = \frac{\hat{\tau}}{\alpha A} > 0$.

The Pro-Labour Government

As $\frac{\partial c_0^W}{\partial \tau} > 0$ for all $\tau > 0$, I find that γ_τ must be negative in the second term of the numerator in (11). But then the optimal policy of a pro-labour government is such that $\tau > \hat{\tau}$. In fact, the solution satisfies

$$\frac{\eta_\tau}{\eta}(\rho-(1-\sigma)\gamma)+\gamma_\tau = 0. \quad (12)$$

As $\frac{\eta_\tau}{\eta} = \frac{1-\alpha}{\tau}$ and $\gamma_\tau = \frac{r_\tau - 1}{\sigma}$ we can multiply through by τ to obtain

$$(1-\alpha)(\rho-(1-\sigma)\gamma)+\frac{r_\tau\cdot\tau-\tau}{\sigma} = 0$$

Noticing that $r = \alpha A\tau^{1-\alpha}$ we get $r_\tau \cdot \tau = (1-\alpha)\alpha A\tau^{-\alpha}\cdot\tau = (1-\alpha)r$. Then

$$\begin{aligned}
\rho+\frac{r}{\sigma} &= (1-\sigma)\left(\frac{r-\tau-\rho}{\sigma}\right)+\frac{\tau}{\sigma(1-\alpha)} \\
\rho+\frac{r}{\sigma} &= \frac{(1-\sigma)r}{\sigma}-\frac{(1-\sigma)\tau}{\sigma}-\frac{(1-\sigma)\rho}{\sigma}+\frac{\tau}{\sigma(1-\alpha)} \\
\rho\sigma+r &= (1-\sigma)r-(1-\sigma)\tau-(1-\sigma)\rho+\frac{\tau}{1-\alpha} \\
\rho(1-\alpha)+\sigma(1-\alpha)r &= [1-(1-\sigma)(1-\alpha)]\,\tau. \quad (13)
\end{aligned}$$

This can be written more compactly as

$$\rho(1-\alpha) = [[1-(1-\sigma)(1-\alpha)]-\sigma(1-\alpha)\alpha A\tau^a]\,\tau$$

and yields the same solution as in the log case when $\sigma = 1$. Let $\check{\tau}$ denote the solution to this equation.

For the comparative static results we rearrange (13) to get

$$\begin{aligned}\rho(1-\alpha)+\sigma(1-\alpha)[r-\tau] &= (1-(1-\alpha))\tau \\ \rho(1-\alpha)+\sigma(1-\alpha)[r-\tau] &= \alpha\ \tau. \end{aligned} \tag{14}$$

For the change in τ due to a change in A we implicitly differentiate the last expression

$$\begin{aligned}\sigma(1-\alpha)[r_\tau-1]d\tau+\sigma(1-\alpha)r_A dA &= \alpha\ d\tau \\ [\sigma(1-\alpha)[r_\tau-1]-\alpha]\,d\tau+\sigma(1-\alpha)r_A dA &= 0.\end{aligned}$$

Thus, we obtain

$$\frac{d\check{\tau}}{dA}=-\frac{\sigma(1-\alpha)r_A}{\sigma(1-\alpha)[r_\tau-1]-\alpha}>0 \tag{15}$$

since the denominator is negative evaluated at $\check{\tau}>\hat{\tau}$ and $r_\tau<1$ when $\tau=\check{\tau}$.

Furthermore, one easily verifies that a change in σ raises $\check{\tau}$, in particular, we have

$$\frac{d\check{\tau}}{d\sigma}=-\frac{(1-\alpha)(r-\tau)}{\sigma(1-\alpha)[r_\tau-1]-\alpha}>0. \tag{16}$$

Thus, a lower elasticity of substitution (higher σ) implies higher taxes under the optimal pro-labour policy.

E Welfare Analysis

We denote by $\hat{\tau}$ the optimal policy for the pro-capital government and $\check{\tau}$ for the pro-labour government. In any optimum we have

$$\hat{\tau}=[\alpha(1-\alpha)A]^{\frac{1}{\alpha}} \quad \text{and} \quad \check{\tau}=\psi(\alpha,A,\rho,\sigma) \tag{1}$$

where $\check{\tau}>\hat{\tau}$ and $\frac{\hat{\tau}}{dA}>0$ and $\frac{\check{\tau}}{dA}>0$.

Thus, policy is endogenous and depends on economic fundamentals. I will not analyze the exact properties of the optimal pro-labour policies any further here. As before we only concentrate on the relative welfare effects of aggregate efficiency. To this end, we fix α, ρ, and σ.

> *We ask what will happen to the welfare of a worker relative to that of a capitalist if aggregate efficiency rises.*

Given the optimal policies and their effect on the marginal products and returns, we now evaluate the welfare functions at the optimal policies. Thus,

$$W^{i^*}(0)=W^i(\alpha,A,\rho,\sigma)=\frac{(c_0^{i^*})^{1-\sigma}}{(1-\sigma)(\rho-(1-\sigma)\gamma^*)}-\frac{1}{\rho(1-\sigma)}, \tag{2}$$

where $i=k,W$ and the superscript $*$ denotes variables that are evaluated at the optimal policies. These, as we recall, are either pro-labour or pro-capital.

As the last term in the welfare function in (2) is the same constant for capitalists and workers, any comparison between welfare is not affected if we define a transformed welfare function

$$V^{i^*} = W^{i^*} + \frac{1}{\rho(1-\sigma)} = \frac{(c_0^{i^*})^{1-\sigma}}{(1-\sigma)(\rho-(1-\sigma)\gamma^*)} \tag{3}$$

with $i = k, W$. V^{i^*} ranks welfare in the same way as W^{i^*} except that the origin of comparison for both groups of agents has been moved by the same number. Notice that the constant term $\frac{1}{\rho(1-\sigma)}$ does not play a role when finding the optimal policies. Thus, it does not feature in them. Hence, choosing policies using V^i or W^i yields the same solutions.

Next, we establish conditions under which the welfare of a worker is less than that of a capitalist. To this end we look at the difference in welfare, denoted by Δ, and obtain by equation (3)

$$\Delta \equiv V^{W*} - V^{k*} = \frac{(c_0^W)^{1-\sigma} - (c_0^k)^{1-\sigma}}{(1-\sigma)(\rho-(1-\sigma)\gamma)} \tag{4}$$

The condition for $\Delta < 0$ boils down to $c_0^{W*} < c_0^{k*}$ and is true[35], if

$$\eta^* \tilde{k_0} < (R^* - \gamma^*)k_0 \quad \Leftrightarrow \quad \frac{n}{l} < \frac{R^* - \gamma^*}{\eta^*}.$$

The first inequality says that the condition is satisfied if the wage income of a (low skilled) worker is less than the income that the capitalist uses for his consumption. This is made up of the after tax income $R^* k_0$ minus his investment outlays γk_0. If the inequality did not hold, there would be an incentive compatibility problem, because the capitalists might just sell their entire wealth and become a worker and be better off. We rule this out since it appears implausible. The second inequality implies that the capitalists are relatively better off, if the number of (low skilled) workers is sufficiently large. This seems a rather mild assumption to make and might reflect the situation in, for instance, not so developed countries. Thus, I will make that assumption for what is to follow.

Assumption 1 *There is a sufficiently large number of workers in the economy, that is,* $\frac{n}{l} < \frac{R^* - \gamma^*}{\eta^*}$.

[35] Notice that V^{i*} is a positive number if $\sigma < 1$, and it is a negative number when $\sigma > 1$. If $\sigma < 1$, then $\Delta < 0$ follows immediately when $c_0^k > c_0^W$. If $\sigma > 0$, then $\Delta < 0$ is satisfied when

$$\frac{(c_0^W)^{1-\sigma} - (c_0^k)^{1-\sigma}}{(1-\sigma)(\rho-(1-\sigma)\gamma)} \quad \Leftrightarrow \quad (c_0^W)^{1-\sigma} - (c_0^k)^{1-\sigma} > 0$$

since $(1-\sigma)$ is a negative number in this case. Thus, $\Delta < 0$ requires

$$(c_0^W)^{1-\sigma} > (c_0^k)^{1-\sigma} \quad \Leftrightarrow \quad 1 = \left(\frac{c_0^W}{c_0^k}\right)^{\sigma-1}$$

which is true when $c_0^k > c_0^W$. Thus, it does not matter for the condition $\Delta < 0$ whether $\sigma \gtrless 1$.

Lemma 6 *If* $\dfrac{n}{l} < \dfrac{R^* - \gamma^*}{\eta^*}$, *then* $c_0^{W*} < c_0^{k*}$ *and* $V^{W*} < V^{k*}$.

The relative welfare change of the two kinds of agents when efficiency changes and policy is endogenous is given by

$$\frac{dV^i}{dA}_{|\tau^*} = \frac{\partial V^i}{\partial A}_{|\tau^*} + \frac{\partial V^i}{\partial \tau}_{|\tau^*} \frac{d\tau}{dA} \quad \text{where} \quad i = W, k \tag{5}$$

and τ^* denotes the policy chosen in the optimum. For welfare comparisons these derivatives must be checked under the optimal pro-capital policy $\hat{\tau}$ and under the optimal policy $\check{\tau}$. I will concentrate on the optimal pro-capital policy first.

E.1 Welfare Comparison under the Optimal Pro-Capital Policy

The optimal pro-capital policy maximizes growth and is given by $\hat{\tau}$. Notice that $\frac{d\hat{\tau}}{dA} = \frac{1}{\alpha}[\alpha(1-\alpha)A]^{\frac{1}{\alpha}-1}\alpha(1-\alpha) = \frac{\hat{\tau}}{\alpha A}$. Furthermore, under that policy we have $R_\tau = r_\tau - 1 = 0$ and $\gamma_\tau = 0$.

I want to show that $\frac{dV^W}{dA}_{|\hat{\tau}} > \frac{dV^k}{dA}_{|\hat{\tau}}$ *or* $\frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} > \frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k}$.

Given that we contemplate the optimal pro-capital policy we can use the envelope theorem on $\frac{dV^k}{dA}$ to get $\frac{dV^k}{dA} = \frac{\partial V^k}{\partial A}$ since $\frac{\partial V^k}{\partial \tau} = 0$ in the optimum. Thus, $\frac{dV^W}{dA} > \frac{dV^k}{dA}$ holds if

$$\frac{\partial V^W}{\partial A}_{|\hat{\tau}} + \frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA} > \frac{\partial V^k}{\partial A}_{|\hat{\tau}}. \tag{6}$$

We know that

$$\begin{aligned}
\frac{\partial V^i}{\partial A}_{|\hat{\tau}} &= \frac{(1-\sigma)^2 (c_0^i)^{-\sigma}\left(\frac{\partial c_0^i}{\partial A}\right)(\rho - (1-\sigma)\gamma) + (1-\sigma)^2\left(\frac{\partial \gamma}{\partial A}\right)(c_0^i)^{1-\sigma}}{(1-\sigma)^2(\rho - (1-\sigma)\gamma)^2} \\
&= \frac{(c_0^i)^{1-\sigma}\left[\left(\frac{\partial c_0^i}{\partial A}\frac{1}{c_0^i}\right)(\rho - (1-\sigma)\gamma) + \left(\frac{\partial \gamma}{\partial A}\right)\right]}{(\rho - (1-\sigma)\gamma)^2}.
\end{aligned}$$

Notice that $\frac{\partial V^i}{\partial A} > 0$ under our assumption that $(\rho - (1-\sigma)\gamma) > 0$. For the term $\frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA}$ I obtain

$$\begin{aligned}
\frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA} &= \frac{(1-\sigma)^2 (c_0^W)^{-\sigma}\left(\frac{\partial c_0^W}{\partial \tau}\right)(\rho - (1-\sigma)\gamma) + (1-\sigma)^2\left(\frac{\partial \gamma}{\partial \tau}\right)(c_0^W)^{1-\sigma}}{(1-\sigma)^2(\rho - (1-\sigma)\gamma)^2} \cdot \frac{d\hat{\tau}}{dA} \\
&= \frac{(c_0^W)^{1-\sigma}\left[\left(\frac{\partial c_0^W}{\partial \tau}\frac{1}{c_0^W}\right)(\rho - (1-\sigma)\gamma) + \left(\frac{\partial \gamma}{\partial \tau}\right)\right]}{(\rho - (1-\sigma)\gamma)^2} \cdot \frac{d\hat{\tau}}{dA}.
\end{aligned}$$

We contemplate the optimal pro-capital policy, so $\gamma_\tau = 0$. Thus,

$$\frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA} = \frac{(c_0^W)^{1-\sigma}\left[\left(\frac{\partial c_0^W}{\partial \tau}\frac{1}{c_0^W}\frac{d\hat{\tau}}{dA}\right)(\rho - (1-\sigma)\gamma)\right]}{(\rho - (1-\sigma)\gamma)^2}.$$

Hence,

$$\frac{\partial V^W}{\partial A}_{|\hat{\tau}} + \frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA} = \frac{(c_0^W)^{1-\sigma}\left[\Lambda\ (\rho-(1-\sigma)\gamma)+\left(\frac{\partial\gamma}{\partial A}\right)\right]}{(\rho-(1-\sigma)\gamma)^2}.$$

where $\Lambda = \left(\frac{\partial c_0^W}{\partial A}\ \frac{1}{c_0^W}\right) + \left(\frac{\partial c_0^W}{\partial \tau}\ \frac{1}{c_0^W}\ \frac{d\hat{\tau}}{dA}\right)$. As $c_0^W = \eta\tilde{k}_0$ we have $\Lambda = \frac{\eta_A}{\eta} + \frac{\eta_\tau}{\eta}\ \frac{d\hat{\tau}}{dA}$ where subscripts denote the partial derivatives.[36] As $\eta = (1-\alpha)A\hat{\tau}^{1-\alpha}$ and $\frac{d\hat{\tau}}{dA} = \frac{\hat{\tau}}{\alpha A}$ we get

$$\Lambda = \frac{1}{A} + \frac{(1-\alpha)^2 A\hat{\tau}^{-\alpha}}{(1-\alpha)A\hat{\tau}^{1-\alpha}}\ \frac{\hat{\tau}}{\alpha A} = \frac{1}{A} + \frac{1-\alpha}{\alpha\ A} = \frac{1}{\alpha A} \tag{7}$$

Furthermore, for the capitalists' welfare we let $\Pi \equiv \frac{\partial c_0^k}{c_0^k} = \frac{R_A - \gamma_A}{R-\gamma}$.

I will now distinguish the cases $\sigma > 1$ and $\sigma < 1$.

The Case $\sigma > 1$.

When $\sigma > 1$ the intertemporal elasticity of substitution is less than unity.[37]

Now for the claimed welfare effects it must then be that

$$\begin{aligned}
\frac{\partial V^W}{\partial A}_{|\hat{\tau}} + \frac{\partial V^W}{\partial \tau}_{|\hat{\tau}} \cdot \frac{d\hat{\tau}}{dA} &> \frac{\partial V^k}{\partial A} \\
\frac{(c_0^W)^{1-\sigma}\left[\Lambda\ (\rho-(1-\sigma)\gamma)+\left(\frac{\partial\gamma}{\partial A}\right)\right]}{(\rho-(1-\sigma)\gamma)^2} &> \frac{(c_0^k)^{1-\sigma}\left[\Pi\ (\rho-(1-\sigma)\gamma)+\left(\frac{\partial\gamma}{\partial A}\right)\right]}{(\rho-(1-\sigma)\gamma)^2} \\
\frac{\left[\Lambda\ (\rho-(1-\sigma)\gamma)+\left(\frac{\partial\gamma}{\partial A}\right)\right]}{\left[\Pi\ (\rho-(1-\sigma)\gamma)+\left(\frac{\partial\gamma}{\partial A}\right)\right]} &> \left(\frac{c_0^k}{c_0^W}\right)^{1-\sigma}
\end{aligned} \tag{8}$$

As $c_0^k > c_0^W$ and $\sigma > 1$ so that $(1-\sigma) < 0$, the RHS of the last inequality is less than 1. Thus, a sufficient condition for the overall inequality to hold is that the LHS be bigger than one. One easily verifies that that is the case if $\Lambda > \Pi$. Now

$$\begin{aligned}
\Lambda &> \Pi \\
\frac{1}{\alpha A} &> \frac{r_A - \gamma_A}{R-\gamma} \\
\frac{1}{\alpha}((r-\hat{\tau})-\gamma) &> r_A \cdot A - \gamma_A \cdot A \\
((r-\hat{\tau})-\gamma) &> \alpha\ r\left(\frac{\sigma-1}{\sigma}\right) \\
(1-\alpha)\ r\left(\frac{\sigma-1}{\sigma}\right) + \frac{\rho}{\sigma} &> \hat{\tau}\left(\frac{\sigma-1}{\sigma}\right)
\end{aligned} \tag{9}$$

[36]For example, η_τ denotes $\frac{\partial\eta}{\partial\tau}$.

[37]This is conventionally argued to be the empirically relevant case. See, for example, Hall (1988). However, recent results suggest that the intertemporal elasticity of substitution is rather close to unity. See, for instance, Beaudry and van Wincoop (1996).

But at $\hat{\tau}$ we have that $(1-\alpha)r = \hat{\tau}$. This is because $r = \alpha A \hat{\tau}^{1-\alpha}$ and $\hat{\tau} = [\alpha(1-\alpha)A]^{\frac{1}{\alpha}}$. As $\rho/\sigma > 0$ for $\sigma < \infty$ we have $\Lambda > \Pi$. Hence, when $\sigma > 1$, the workers benefit more from higher A than the capital owners.

Lemma 7 *If* $\sigma > 1$, *then* $\frac{dV^W}{dA}_{|\hat{\tau}} > \frac{dV^k}{dA}_{|\hat{\tau}}$.

Intuitively, Lemma 7 says that the marginal change in welfare due to a change in A is greater for the workers than for the capital owners. Thus, no matter what level welfare of the two agents we start from, the change is bigger for the workers.

The case: $\sigma < 1$

If $\sigma < 1$, the intertemporal elasticity of substitution is higher than one. People react relatively more to differences in r and ρ, and may be quite willing to shift consumption over time. This case is not found to represent actual behaviour very well.

To deal with the case $\sigma < 1$ under the optimal pro-capital policy we will focus on the relative welfare changes $\frac{dV^i}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^i}$. This yields for a worker

$$\begin{aligned} \frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} &= \frac{(c_0^W)^{1-\sigma}\Lambda\ (\rho-(1-\sigma)\gamma) + \left(\frac{\partial\gamma}{\partial A}\right)(c_o^W)^{1-\sigma}}{(\rho-(1-\sigma)\gamma)^2} \\ &\cdot \frac{(1-\sigma)(\rho-(1-\sigma)\gamma)}{(c_0^W)^{1-\sigma}} \\ &= (1-\sigma)\Lambda + \frac{(1-\sigma)\frac{\partial\gamma}{\partial A}}{(\rho-(1-\sigma)\gamma)} \end{aligned}$$

For a capitalist we similarly obtain

$$\frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k} = (1-\sigma)\Pi + \frac{(1-\sigma)\frac{\partial\gamma}{\partial A}}{(\rho-(1-\sigma)\gamma)}$$

Thus, if $\sigma < 1$ and $(1-\sigma) > 0$, the change in welfare is greater for a worker than a capitalist when

$$\begin{aligned} \frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} &> \frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k} \\ \Lambda &> \Pi. \end{aligned}$$

But this inequality holds by the same arguments which led to (9). Thus,

Lemma 8 *If* $\sigma < 1$, *then*

$$\frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} > \frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k}$$

Hence, we can summarize our findings in

Proposition 6 *If period utility is of the constant intertemporal elasticity of substitution type and $\sigma \gtrless 1$, where $\nu \equiv \frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution, then a worker benefits relatively more from technical progress (higher A) than a capitalist under a pro-capital policy, that is,*

$$\frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} > \frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k} \quad \textit{if} \quad \sigma < 1,$$
$$\frac{dV^W}{dA}_{|\hat{\tau}} > \frac{dV^k}{dA}_{|\hat{\tau}} \quad \textit{if} \quad \sigma > 1.$$

E.2 Welfare Comparison under the Optimal Pro-Labour Policy

Under the optimal pro-labour policy a worker benefits relatively more from technical progress if

$$\frac{dV^W}{dA}_{|\check{\tau}} > \frac{dV^k}{dA}_{|\check{\tau}} \quad \text{or} \quad \frac{dV^W}{dA}_{|\check{\tau}} \left(\frac{1}{V^W}\right) > \frac{dV^k}{dA}_{|\check{\tau}} \left(\frac{1}{V^k}\right).$$

We know that $V^W < V^k$ by Lemma 6.

We first check whether $\frac{dV^W}{dA}_{|\check{\tau}} > \frac{dV^k}{dA}_{|\check{\tau}}$. Under the optimal policy we can again employ the envelope theorem so that the condition boils down to

$$\frac{\partial V^W}{\partial A}_{|\check{\tau}} > \frac{\partial V^k}{\partial A}_{|\check{\tau}} + \frac{\partial V^k}{\partial \tau}_{|\check{\tau}} \cdot \frac{d\check{\tau}}{dA}$$

For the worker we have

$$\begin{aligned}
\frac{\partial V^W}{\partial A}_{|\check{\tau}} &= \frac{(1-\sigma)^2 (c_0^W)^{-\sigma} \left(\frac{\partial c_0^W}{\partial A}\right) (\rho - (1-\sigma)\gamma) + (1-\sigma)^2 \left(\frac{\partial \gamma}{\partial A}\right) (c_0^W)^{1-\sigma}}{(1-\sigma)^2 (\rho - (1-\sigma)\gamma)^2} \\
&= \frac{(c_0^W)^{1-\sigma} \left(\frac{\partial c_0^W}{\partial A} \cdot \frac{1}{c_0^W}\right) (\rho - (1-\sigma)\gamma) + \left(\frac{\partial \gamma}{\partial A}\right) (c_0^W)^{1-\sigma}}{(\rho - (1-\sigma)\gamma)^2}
\end{aligned}$$

and $\frac{\partial V^W}{\partial \tau}_{|\check{\tau}} \cdot \frac{d\check{\tau}}{dA} = 0$ by the envelope theorem. For a capital owner we similarly obtain

$$\begin{aligned}
\frac{\partial V^k}{\partial A}_{|\check{\tau}} + \frac{\partial V^k}{\partial A}_{|\check{\tau}} \cdot \frac{d\check{\tau}}{dA} &= \frac{(c_0^k)^{1-\sigma} \left(\frac{\partial c_0^k}{\partial A} \cdot \frac{1}{c_0^k}\right) (\rho - (1-\sigma)\gamma) + \left(\frac{\partial \gamma}{\partial A}\right) (c_0^k)^{1-\sigma}}{(\rho - (1-\sigma)\gamma)^2} \\
&+ \frac{(c_0^k)^{1-\sigma} \left(\frac{\partial c_0^k}{\partial \tau} \cdot \frac{1}{c_0^k}\right) (\rho - (1-\sigma)\gamma) + \left(\frac{\partial \gamma}{\partial \tau}\right) (c_0^k)^{1-\sigma}}{(\rho - (1-\sigma)\gamma)^2} \cdot \frac{d\check{\tau}}{dA}
\end{aligned}$$

The case: $\sigma > 1$

As $\left(\frac{c_0^k}{c_0^W}\right)^{1-\sigma} < 1$ when $\sigma > 1$, one verifies that the welfare change for a worker is bigger than for a capital owner if

$$\begin{aligned} \frac{c_A^W}{c_0^W} B + \gamma_A &> \frac{c_A^k}{c_0^k} B + \gamma_A + \frac{c_\tau^k}{c_0^k} B \frac{d\check{\tau}}{dA} + \gamma_\tau \frac{d\check{\tau}}{dA} \\ \frac{c_A^W}{c_0^W} &> \frac{c_A^k}{c_0^k} + \frac{c_\tau^k}{c_0^k} \cdot \frac{d\check{\tau}}{dA} + \frac{\gamma_\tau}{B} \frac{d\check{\tau}}{dA} \end{aligned} \tag{10}$$

where $B = (\rho - (1-\sigma)\gamma)$. Now

$$\frac{c_A^W}{c_0^W} = \frac{\eta_A}{\eta} = \frac{1}{A}, \quad \frac{c_\tau^k}{c_0^k} = \frac{R_\tau - \gamma_\tau}{R - \gamma}, \quad \text{and } \frac{c_A^k}{c_0^k} = \frac{R_A - \gamma_A}{R - \gamma},$$

where $R = (r - \tau)$. Thus, the inequality in (10) becomes

$$\frac{1}{A} > \frac{R_A - \gamma_A}{R - \gamma} + \frac{R_\tau - \gamma_\tau}{R - \gamma} \cdot \frac{d\check{\tau}}{dA} + \frac{\gamma_\tau}{B} \frac{d\check{\tau}}{dA}. \tag{11}$$

$R_A = r_A$ and $R_\tau = (r_\tau - 1) < 0$ evaluated at $\check{\tau}$. Furthermore, $r_A \cdot A = r$, and one easily verifies that $B = R - \gamma$. Thus, it must be that

$$\begin{aligned} \frac{1}{A} &> \frac{R_A - \gamma_A}{R - \gamma} + \frac{R_\tau}{R - \gamma} \cdot \frac{d\check{\tau}}{dA} & (12) \\ 1 &> \frac{r - \frac{r}{\sigma} + R_\tau \cdot \frac{d\check{\tau}}{dA} \cdot A}{R - \gamma} \\ \sigma(R - \gamma) &> (\sigma - 1)r + \sigma \cdot R_\tau \cdot \frac{d\check{\tau}}{dA} \cdot A \\ \sigma\left[(r - \check{\tau}) - \frac{(r - \check{\tau}) - \rho}{\sigma}\right] &> (\sigma - 1)r + \sigma \cdot R_\tau \cdot \frac{d\check{\tau}}{dA} \cdot A \\ (\sigma - 1)r - (\sigma - 1)\check{\tau} + \rho &> (\sigma - 1)r + \sigma \cdot R_\tau \cdot \frac{d\check{\tau}}{dA} \cdot A \\ \rho &> (\sigma - 1)\check{\tau} + \sigma \cdot R_\tau \cdot \frac{d\check{\tau}}{dA} \cdot A & (13) \end{aligned}$$

Notice that $R_\tau < 0$ and $\frac{d\check{\tau}}{dA} \cdot A > 0$ so that the second term on the RHS of the inequality is negative. However, the first term on the RHS is positive. We can then use (15) to obtain

$$\rho > (\sigma - 1)\check{\tau} + \sigma \cdot R_\tau \cdot \frac{\sigma(1-\alpha) r_A}{\alpha - \sigma(1-\alpha)(r_\tau - 1)} \cdot A.$$

As $r_A \cdot A = r$ we get

$$\rho > (\sigma - 1)\check{\tau} + \sigma \cdot R_\tau \cdot \frac{\sigma(1-\alpha) r}{\alpha - \sigma(1-\alpha)(r_\tau - 1)}. \tag{14}$$

This inequality is not easy to analyze. I have evaluated the expressions on the RHS further, but I have not found definite conditions on the parameters which ensure that the inequality holds when $\sigma > 1$. Furthermore, numerical simulations have revealed that for certain (plausible, i.e. often used by calibration exercises) constellations of (A, α, ρ) the inequality might hold if σ is larger but not too much larger than one. Hence, for larger σ the inequality does *not* hold. Thus,

Lemma 9 *If $\sigma^+ > \sigma > 1$, where σ^+ denotes some threshold σ, and for a variety of parameter values of (A, α, ρ), the workers benefit relatively more than the capital owners under a pro-labour policy when there is technical progress and the intertemporal elasticity of substitution ν is lower than unity.*

If $\sigma > \sigma^+ > 1$, then the capital owners will benefit relatively more than the workers under a pro-labour policy, when there is technical progress and the intertemporal elasticity of substitution ν is sufficiently lower than unity.

The case: $\sigma < 1$

For this case we check whether $\frac{dV^W}{dA}_{|\check{\tau}} \left(\frac{1}{V^W}\right) > \frac{dV^k}{dA}_{|\check{\tau}} \left(\frac{1}{V^k}\right)$. Now we can go through the same first steps as for the case with $\sigma > 1$. Given the expressions for each derivative and V^i one readily verifies that for the workers to benefit relatively more it must be that (10) must be satisfied. But that boils down to the conditions stated in (12) and so in (13). Form the latter we immediately infer that the inequality condition is satisfied for $\sigma < 1$. Thus,

Lemma 10 *If $\sigma < 1$, then the workers will benefit relatively more than the capital owners under a pro-labour policy, when there is technical progress and the intertemporal elasticity of substitution ν is larger than unity.*

Summarizing we have

Proposition 7 *If period utility is of the constant intertemporal elasticity of substitution type and $\sigma < 1 < \sigma^*$, where $\nu \equiv \frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution and σ^* denotes some σ not too much larger than unity, then a worker benefits relatively more from technical progress (higher A) than a capitalist under a pro-labour policy, that is,*

$$\begin{aligned} \frac{dV^W}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^W} &> \frac{dV^k}{dA}_{|\hat{\tau}} \cdot \frac{1}{V^k} \quad \textit{if} \quad \sigma < 1, \\ \frac{dV^W}{dA}_{|\hat{\tau}} &> \frac{dV^k}{dA}_{|\hat{\tau}} \quad \textit{if} \quad \sigma^* > \sigma > 1, \end{aligned}$$

If period utility is of the constant intertemporal elasticity of substitution type and $\sigma > \sigma^* > 1$, where $\nu \equiv \frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution and σ^* denotes some σ not too much larger than unity, a worker benefits relatively less from technical progress (higher A) than a capitalist under a pro-labour policy, that is,

$$\frac{dV^W}{dA}_{|\hat{\tau}} < \frac{dV^k}{dA}_{|\hat{\tau}} \quad \textit{if} \quad \sigma > \sigma^* > 1,$$

F Comparative Statics

As the optimal pro-capital policy is identical the one derived under logarithmic utility, the comparative static results are the same as in the main text.

For the optimal pro-labour policy two conditions must be checked. The first one concerns the effect of growth on a change in A if policy is endogenous. See p. 25 in the resubmitted paper. For $\frac{d\gamma}{dA} > 0$ we must have $r_A + (r_\tau - 1)\frac{d\check{\tau}}{dA} > 0$. Now $\frac{d\check{\tau}}{dA} = \frac{\sigma(1-\alpha)r_A}{\alpha-\sigma(1-\alpha)(r_\tau-1)}$. Thus,

$$r_A + (r_\tau - 1)\frac{d\check{\tau}}{dA} > 0 \quad \Leftrightarrow \quad r_A + \frac{(r_\tau - 1)\sigma(1-\alpha)r_A}{\alpha - \sigma(1-\alpha)(r_\tau - 1)} > 0.$$

The denominator of the second term on the LHS of the inequality is positive (see equation (15)), whereas the numerator is negative as $(r_\tau - 1) < 0$ at $\tau = \check{\tau}$. Thus, the inequality holds if

$$\begin{aligned} \alpha\, r_A - \sigma(1-\alpha)(r_\tau - 1)r_A + (r_\tau - 1)\sigma(1-\alpha)r_A &> 0 \\ \alpha\, r_A &> 0 \end{aligned}$$

which is true. Hence, $\frac{d\gamma}{dA} > 0$ under the optimal pro-labour policy and utility is of the CIES type.

The second point in relation to the results in the paper on p. 26, section 7.2, is the question whether $\frac{d\Pi}{dA}_{|\lambda=0} < 0$. As in the paper the condition for this is that

$$\alpha\check{\tau}^{\alpha-1}\frac{d\check{\tau}}{dA}\alpha A - \alpha\check{\tau}^{\alpha} < 0.$$

The inequality is satisfied if $\frac{d\check{\tau}}{dA} < \frac{\check{\tau}}{\alpha A}$. The latter inequality boils down to

$$\begin{aligned} \sigma(1-\alpha)r_A &< \frac{\check{\tau}}{\alpha A}\left[\alpha - \sigma(1-\alpha)(r_\tau - 1)\right] \\ \sigma(1-\alpha)\alpha\, r &< \alpha\,\check{\tau} - \sigma(1-\alpha)(r_\tau - 1)\check{\tau} \\ \sigma(1-\alpha)\alpha\, r &< \alpha\,\check{\tau} - \sigma(1-\alpha)((1-\alpha)r - \check{\tau}) \\ \sigma(1-\alpha)\alpha\, r &< \alpha\,\check{\tau} - \sigma(1-\alpha)r + \sigma\alpha(1-\alpha)r + \sigma(1-a)\check{\tau} \\ 0 &< \alpha\,\check{\tau} - \sigma(1-\alpha)(r - \check{\tau}). \end{aligned} \tag{1}$$

But $\check{\tau}$ satisfies (14) and so

$$\alpha\,\check{\tau} = \rho(1-\alpha) + \sigma(1-\alpha)[r - \check{\tau}].$$

Substituting for $\alpha\,\check{\tau}$ yields for (1)

$$\begin{aligned} 0 &< \rho(1-\alpha) + \sigma(1-\alpha)[r - \check{\tau}] - \sigma(1-\alpha)(r - \check{\tau}) \\ 0 &< \rho(1-\alpha). \end{aligned}$$

which is true. Hence, $\frac{d\Pi}{dA}_{|\lambda=0} < 0$ under the optimal pro-labour policy.