

# Policies of Different Governments: Persistence and Interactions

by

Frantisek Ricka

A.B. Economics, Princeton University (2003)

Submitted to the Department of Economics  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

© 2008 Frantisek Ricka. All rights reserved.

The author hereby grants to Massachusetts Institute of Technology permission to  
reproduce and  
to distribute copies of this thesis document in whole or in part.

Signature of Author .....

Department of Economics  
5 August 2008

Certified by .....

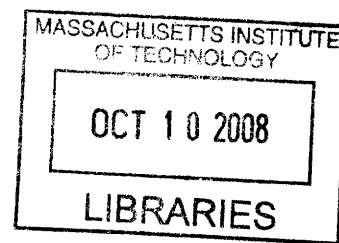
*[Signature]*  
Pentti J. K. Kouri Career Development Assistant Professor of Economics  
Thesis Supervisor

Certified by .....

*[Signature]*  
Francesco Giavazzi  
Visiting Professor of Economics  
Thesis Supervisor

Accepted by ... ..

Peter Temin  
Elisha Gray II Professor of Economics  
Chairman, Departmental Committee on Graduate Studies



**ARCHIVES**



# Policies of Different Governments: Persistence and Interactions

by

Frantisek Ricka

Submitted to the Department of Economics  
on 5 August 2008, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

## Abstract

This dissertation consists of three chapters on persistence and interactions of policies of different governments in various settings.

Chapter 1 studies government policy persistence when firms face capital installation costs. In an environment, where the government is either left- or right-wing, the left-wing government almost always sets more right-wing policies when it follows a conservative administration than otherwise. Similarly, when the types of policy makers are close enough to each other, a right-wing government that succeeds a leftist one chooses more left-wing policies. However, when the preferences of the politicians are sufficiently different, a conservative government selects more right-wing fiscal policies when following a leftist administration in order to address the state of the economy it inherits.

Chapter 2 considers the political implications of tax competition between countries of different sizes. Smaller countries competing for internationally mobile capital set lower tax rates than their larger counterparts. This is because they perceive higher elasticity of capital with respect to their policy and their governments are to the right of those elected in larger countries. Then a more significant number of small countries involved in the competition with large countries not only decreases the large-country tax rates on capital, but also shifts their governments to the right. Large countries do not have a similar “right-wing” power and the presence of more of them in a competition can actually cause a left-ward shift in the governments of the competitors.

Chapter 3 compares educational achievement of 15-year olds in post-communist versus other countries. It finds that even almost 20 years after the fall of communism, the effect of the regime and its policies seems to persist in the educational system of the Eastern Bloc countries. Students in the East achieve better in mathematics and hard sciences and worse in reading than their Western counterparts. This is likely because the communist regimes supported education in the former but stifled free and interpretive thinking necessary to achieve on the latter test. While the advantage of the East may be shrinking, its disadvantage remains the same.

Thesis Supervisor: Guido Lorenzoni

Title: Pentti J. K. Kouri Career Development Assistant Professor of Economics

Thesis Supervisor: Francesco Giavazzi

Title: Visiting Professor of Economics



# Contents

<b>1</b>	<b>Capital Installation Costs and Government Policy Persistence</b>	<b>11</b>
1.1	Introduction . . . . .	11
1.2	The Basic Model . . . . .	15
1.2.1	The Setup . . . . .	15
1.2.2	Myopic Firms and Government: First-Period Left . . . . .	18
1.2.3	Myopic Firms and Government: Extensions . . . . .	26
1.3	Forward-Looking Firms and Government . . . . .	32
1.3.1	First-Period Left-Wing Government . . . . .	34
1.3.2	First-Period Right-Wing Government . . . . .	40
1.3.3	Summary of Results . . . . .	42
1.3.4	More Than Two Periods . . . . .	44
1.3.5	Government Objective: Caring About the Future . . . . .	45
1.4	Conclusion . . . . .	46
1.5	Appendix . . . . .	48
<b>2</b>	<b>The Right-Wing Power of Small Countries</b>	<b>53</b>
2.1	Introduction . . . . .	53
2.2	The Model . . . . .	56
2.2.1	The Basics . . . . .	56
2.2.2	Small Country . . . . .	59
2.2.3	Large Country . . . . .	60
2.3	Tax Competition . . . . .	61

2.3.1	Nash Competition . . . . .	61
2.3.2	Elections . . . . .	62
2.4	Composition of Competitors and Political Shifts . . . . .	65
2.4.1	Small Countries and the Shift to the Right . . . . .	65
2.4.2	More Competitors: Small Countries . . . . .	67
2.4.3	More Competitors: Large Countries . . . . .	68
2.4.4	A Policy Example: Tax Cooperation and Minimum Tax . . . . .	71
2.5	Extensions . . . . .	76
2.5.1	Small Countries with Positive Size . . . . .	76
2.5.2	Capital Ownership by Median Agents . . . . .	81
2.6	Conclusion . . . . .	83
2.7	Appendix . . . . .	84
<b>3</b>	<b>Twenty Years On: Educational Achievement in Post-Communist Countries</b>	<b>93</b>
3.1	Introduction . . . . .	93
3.2	Description of the Data . . . . .	95
3.3	Study Design and Data Processing . . . . .	97
3.3.1	Sample Weights . . . . .	97
3.3.2	Replicate Weights . . . . .	99
3.3.3	Plausible Values . . . . .	102
3.3.4	Indeces . . . . .	104
3.4	The Model . . . . .	105
3.5	Results and Analysis . . . . .	106
3.5.1	Mathematics . . . . .	106
3.5.2	Other Subjects . . . . .	109
3.5.3	Impact of the Fay's Replicates Method . . . . .	111
3.6	Conclusion . . . . .	111
3.7	Appendix: Tables . . . . .	113

# Acknowledgment

I would like to thank my adviser, Professor Guido Lorenzoni, for his help with my projects throughout the years. Without his advice I would not have made it all the way to the end.

I also want to thank Professor Francesco Giavazzi, my second adviser. His insights were instrumental to significant improvements in the various chapters of this thesis.

My great thanks go to my classmates and roommates Joao Leao and Emre Kocatulum. Our countless conversations revolved around economics and anything else imaginable. Together with my other friends here in Cambridge, it was them who created the great atmosphere, in which it is pleasure to work on one's thesis.



*To my Mom and Grandma*



## Chapter 1

# Capital Installation Costs and Government Policy Persistence

### 1.1 Introduction

When an election results in a government change, the voters have probably elected the new administration based on that party's promise to alter the policies of their country. Otherwise they would have simply re-elected the old policy makers. However, we can sometimes see that the new government changes the policies of its predecessor much less than one would expect, whether based on its pre-election rhetoric or its political profile. Thus we observe a degree of policy persistence.

The most recent change of government in Slovakia provides a good example: A socialist government was formed in 2006 after a four-year rule by a purely right-wing coalition. The socialists' electoral platform in 2002, their voting record and public statements over the next four years as well as their agenda in the 2006 elections all suggested a strong disagreement with the policies of the conservative government. They first opposed and subsequently wanted to abolish almost all of the right-wing reforms implemented between 2002-2006.

Then, contrary to what was widely expected, the new left-wing government left many of the economically conservative reforms of the previous administration virtually intact - only very slightly modifying the country's flat tax and not even decreasing the value-added tax on items such as food or books. It is nearly impossible to believe that a socialist government would

implement the same policies were it to follow a more left-wing administration<sup>1</sup>. Rather, the right-wing fiscal reforms persisted after the change of government in 2006.

This paper develops a model of persistence in fiscal policy. We describe a simple mechanism for a small country, where firms face adjustment costs in the form of capital installation costs. In this environment, a representative politician from a heterogeneous population pursues policies that maximize his welfare and that of the agents of his type. Then, when the policy maker in power changes, even without trying to affect his successor, a politician's fiscal policy choice happens to affect the subsequent government.

More specifically, our economy is populated by agents who differ according to their productivity. They elect from amongst themselves a welfare-maximizing politician. We assume an exogenous election process, where each party comes to power with a certain probability. Thus we obtain alternating policy makers in power.

In our model, firms use internationally mobile capital that requires a certain net rate of return. However, in the period when they begin using a particular amount of capital, they also have to pay a premium for installing it. Therefore, in our two-period setup, installation costs are paid on all capital present in the economy in the first period. We first show that the second-period left-wing government never induces installation of additional capital. Thus firms in the second period do not face the additional installation cost on the margin and are willing to bear a higher tax rate without removing any capital from the economy. The leftist government takes advantage of this by taxing away this "wedge" in firms' marginal cost. It sets the highest possible tax rate while ensuring that none or little capital leaves the country. The tax rate is then clearly higher in the second period than it was in the first one.

However, this tax rate must vary by how much capital is present in the economy when the government comes to power. To induce a higher amount of capital to remain in the country, the administration must choose a relatively lower tax rate. A first-period conservative government sets a lower tax rate than a first-period left-wing policy maker, and therefore attracts more capital. Then to keep this higher amount of capital employed by firms, the second-period

---

<sup>1</sup>One could think of Tony Blair's election in 1997 and subsequent maintaining of many Thatcherite policies as a more famous example of policy persistence. However, Blair was very un-Labour by 90s standards. He reformed the party first and changed its preferences prior to elections in order to capture more of the center of the electorate. Therefore, considering the altered goals of New Labour, it is conceivable that Blair would have implemented in the first place the policies he kept from the Conservatives.

left-wing government must choose a relatively lower tax rate than were it to follow a first-period leftist administration. But this is policy persistence: A smaller first-period tax rate means a second-period tax rate that is relatively lower than were the first-period tax rate higher. Importantly, the policy persistence occurs without a government implementing policies specifically designed at altering the behavior of its successor.

We further demonstrate that our model can also deliver policy persistence when a right-wing government follows a left-wing administration. However, the occurrence of the persistence depends on who is in the second-period government: When the types of policy makers who can be elected are similar enough, a left-wing government is going to pursue more right-wing policies when following a conservative government. Similarly, a right-wing government will be more leftist when it succeeds a left-wing administration. In this case we observe persistent fiscal policy irrespective of preferences of the second-period policy maker.

But when the types of politicians are sufficiently different, the situation changes. Only a leftist government is more fiscally conservative when it is elected after a period of right-wing administration. The right, when it follows the left, on the other hand, becomes even more conservative, in order to "correct" the state of the economy it inherits. Within our framework it is then the right-wing that can be expected, under some circumstances, to undertake more audacious reforms. The left, when it succeeds the right, is usually content with merely adjusting its conservative predecessor's policies.

The literature has already addressed some forms and aspects of policy persistence. For instance, Fernandez and Rodrik (1991) propose a mechanism based on individual uncertainty that causes preference for the status quo, under a direct majority rule. Individuals receive uncertain payoffs from a reform. Therefore their ex-ante expected payoffs and ex-post actual payoffs differ. If the ex-ante expected payoff is positive for a majority, the reform is adopted and actual rewards are revealed. If those are only positive for a minority, the reform can be repealed. However, when the expected payoff is negative for most voters, the reform is never adopted, even if the realized rewards were to have become positive. The above asymmetry then results in a status quo bias.

However, policies are usually implemented indirectly, by a policy maker. Thus another explanation for policy persistence could be that lobbies form to defend policies already in

place, hinder the administration's plans of change and prevent the implementation of reforms. However, Coate and Morris (1999) suggest that same, and equally powerful, lobbies could form to implement a policy in the first place. Their model therefore explores a mechanism whereby firms adapt to benefit from policies of a particular administration. Their actions, over time, increase the value they place on particular policies in place and therefore also their willingness to pay to keep those policies in place. Therefore, when a government is replaced, they can induce its successor to maintain policies that had already been implemented during the previous election period.

Coate and Morris obtain their results assuming politicians only care about being in office and about the bribes they receive from firms. While unfortunately this may frequently be the case, our mechanism for policy persistence assumes more representative politicians. Moreover, their model explores policies targeted to the various sectors of the economy rather than more general fiscal policy. Therefore we adopt the type of framework implemented by Persson and Tabellini (1994), Alesina and Tabellini (1990) or Persson and Svensson (1989), when they explore various interactions between politics and fiscal policy. They all assume heterogeneous populations of agents who, in an exogenously modelled process, elect representative welfare-maximizing governments from amongst themselves<sup>2</sup>.

Using this assumption, Alesina and Tabellini reason that changes between politicians in power are responsible for higher debt levels and show that voters will strategically delegate policy decisions to agents that represent them. However, most relevant to our case, Persson and Svensson show that a conservative government, when it knows it will be succeeded by a leftist government, will want to manipulate the level of public debt in order to affect the policies of its successor.

In their model, the conservative government prefers less spending than its left-wing opponents. Nevertheless, it will increase public spending somewhat, even if not to the level preferred by the left, when it knows it will be succeeded by a more leftist administration<sup>3</sup>. Through in-

---

<sup>2</sup>In a fully endogenous median voter model, it might be difficult to obtain switches between policy makers that we and the authors want to consider, unless there is an uncertainty introduced into the decision-making process of the median voter himself. For example, Alesina and Tabellini consider a version of their model, where it is clear, who the median voter will elect, but ex-ante it is not known who the median voter is.

<sup>3</sup>They derive their results under the assumption of perfect foresight and then, very reasonably, assume that introducing uncertainty would not change their conclusions qualitatively.

creased spending it will accumulate a higher level of debt. Since its left-wing successor will need to repay the debt, he will not be free to spend as much as he would like. Thus the first-period government deliberately alters its policy in order to affect that of its successor, who has different preferences. This creates policy persistence. It may indeed sometimes be the case that a politician pursues policies targeted at influencing his successor. However, our paper develops a framework that explains policy persistence without having to assume that governments distort their policies in order to affect whoever might follow them.

The remainder of the paper is organized as follows: Section 2 describes the environment of our model. Section 3 discusses the model under the simplifying assumption of myopic firms and government, as well as the impact of alternative assumptions on our mechanism. In section 4 we then consider the more realistic case of forward-looking firms and government. The last section concludes.

## 1.2 The Basic Model

### 1.2.1 The Setup

The setup is as follows: We model a small open economy, taking the rest of the world as exogenous. The economy is populated by agents who supply labor and competitive firms, which produce a single final good using the labor and capital from abroad. The economy has a government which can tax capital income and subsequently redistribute the tax revenue lump-sum to the agents. This form of government is employed, for instance, by Persson and Tabellini (1992).

The agents supply labor inelastically and the labor market clears. Therefore their income consists of their wage and the lump-sum transfers they receive from the government. For simplicity we will assume that the agents are risk neutral. There are two types of agents, which differ according to their observable labor productivities,  $\gamma_L$  and  $\gamma_R$ , such that  $(\gamma_L + \gamma_R)/2 = 1$ . The productivities determine their marginal product when employed by a firm and therefore wage.

The firms produce the final good according to a simple Cobb-Douglas production function,

$$F = K^\alpha X^{(1-\alpha)},$$

where  $X$  is the composite labor,

$$X = \gamma_R L(\gamma_R) + \gamma_L L(\gamma_L),$$

and  $L(\gamma)$  is the quantity of each type of labor. For simplicity we will assume that labor supply is  $L(\gamma_L) = L(\gamma_R) = 1$ .

$K$  is the capital employed by the firm. Capital does not depreciate in our model. Firms maximize their profits and therefore pay each factor of production its marginal product. All capital is internationally mobile and therefore it requires a certain net return  $r^*$  that represents its outside option. Since ours is a small open economy, we take  $r^*$  as given and determined in the world equilibrium that is outside the scope of our model. Note that because of this, ownership of capital by domestic agents does not impact the policy choice by the government: Any net capital income of agents that would enter their welfare is exogenous.<sup>4</sup>

However, to begin the production process, a firm needs to first install any new quantity of capital, which incurs an additional cost  $c$  per unit of capital. This cost is tax-deductible from profit and therefore not taxable. It is only incurred once, at the time of installation. That is, if a unit of capital is already installed at the beginning of a period, that unit requires no additional subsequent installation cost. Therefore the marginal product of capital net of tax must equal either  $r^*$  when the capital had already been installed before or  $r^* + c$  when the capital needs to be yet installed.

The results of the model are not affected qualitatively when instead of a cost of installation we assume the firms must pay a cost of removing capital that has already been installed. Also, we could assume a more general constant elasticity of substitution production function or convex costs of installation. Please see below for a discussion of results of our basic model under these

---

<sup>4</sup>We should mention that if all capital is foreign-owned, our small country will be a net exporter. Else, we could assume that some agents own capital and then the country could either be an exporter or an importer, depending on whether the amount of capital in the country would exceed that owned by those agents or not.

alternative assumptions. It is because we later use the simplifying assumptions in the more complex treatment of the model with forward-looking firms and government that we introduce the basic model with a Cobb-Douglas production function and linear installation costs.

Finally, there are two possible types of government. The policy maker is elected from among the population and therefore both types maximize the utility of a particular type of an agent in the period, in which they are in power: We will call a left-wing government an administration that maximizes the utility of an agent with lower productivity,  $\gamma_L$ , and a right-wing or conservative government one that maximizes the utility of a high-productivity agent with parameter  $\gamma_R$ . As mentioned above, the government chooses a tax rate on capital and then redistributes the tax revenue lump-sum to the agents. We shall assume that all tax rates are non-negative<sup>5</sup>. All agents in the economy receive the same transfer. Since the tax rate on capital indirectly affects the two types of agents differently (e.g. the more productive agents stand to lose more because their wages are a steeper function of capital, which in turn is a function of tax rates), we do not need to consider differential lump-sum transfers to the two types of agents. Uniform transfers do achieve redistribution in our environment.

The model has two periods. The first-period begins with a government for that period already in place, i.e. the elections for that period had already occurred. Each government type has a certain probability of coming to power in the second period. That probability is exogenous and independent of the first-period government or its policies. Therefore we do not endogenize the election process in our model: Just as Persson and Svensson (1989) and others we assume that there are factors other than economic that will determine who is elected to the government. Technically, since in this model each agent would want to vote for his own type (because that type will maximize his welfare once elected) and the number of each type of agents is the same, election outcome should be purely random and each party have a 50% chance of winning. Once we add other exogenous non-economic factors, such as candidates' declared preferences over social issues, personal charisma, etc, this probability can deviate and one type of government may be more likely to win than the other.

The game in our model proceeds as follows: In the first period, a government selects a tax

---

<sup>5</sup>In our model, subsidies to capital would be financed by negative transfers to agents. Therefore a negative tax rate on capital would imply imposing a lump-sum tax on the agents.

rate  $\tau_1$ , maximizing the utility of the agents, about whom it cares. The firms then decide how much capital to install, which in turn determines agents' wages and the size of the transfers in the economy. While in our first example the firms are short-sighted, later they are forward-looking and realize that their installation decision will have an impact on how much capital they will have readily available in the second period. Then the second-period government comes to power and selects a tax rate  $\tau_2$ , following which the firms decide to either retain the same amount of capital, install additional capital or reduce the amount of capital they employ for the final period.

### 1.2.2 Myopic Firms and Government: First-Period Left

Starting with the simplest setup, where firms and governments are short-sighted and only maximize considering the current period, we first realize that the marginal cost of capital to firms will be the sum of the installation cost, the required net return to capital and the tax on capital. The firms need to ensure that the owners of the capital receive  $r^*$  after tax. Therefore they will pay the capital owners a gross return  $r_g$  such that  $(1 - \tau_1) r_g = r^*$ . This brings the gross return to  $\frac{r^*}{1 - \tau_1}$ . Moreover, for each unit of capital they install, the firms pay installation costs  $c$ . As mentioned earlier, these are not taxed. Therefore the total cost to the firms of each unit of capital they employ in the first period will be  $\frac{r^*}{1 - \tau_1} + c$ .

Thus the firms' first-period objective function is

$$\max_{K_1, L_1(\gamma)} K_1^\alpha X_1^{(1-\alpha)} - \sum_{\gamma} w_1(\gamma) L_1(\gamma) - \left( \frac{r^*}{1 - \tau_1} + c \right) K_1,$$

which yields two first-order conditions:

$$(1 - \alpha) K_1^{\alpha} X_1^{-\alpha} \gamma L_1(\gamma) - w_1(\gamma) = 0 \text{ and}$$

$$\alpha K_1^{\alpha-1} X_1^{(1-\alpha)} - \frac{r^*}{1 - \tau_1} - c = 0.$$

Therefore

$$K_1 = \left[ \frac{\alpha X_1^{1-\alpha}}{\frac{r^*}{1 - \tau_1} + c} \right]^{\frac{1}{1-\alpha}}.$$

Since we have assumed that the labor market clears, we also have  $L_1(\gamma) = 1$ .

Government then maximizes the utility of its type of agent  $\gamma$ , which is the sum of the agent's labor income and lump-sum transfers. Its objective function is

$$\max_{\tau_1} \left[ w_1(\gamma) + \frac{1}{2} \tau_1 \frac{r^*}{1 - \tau_1} K_1 \right],$$

which becomes, after substituting for wages and capital from the firms' optimization problem,

$$\max_{\tau_1} \left[ (1 - \alpha) \left[ \frac{\alpha}{\frac{r^*}{1 - \tau_1} + c} \right]^{\frac{\alpha}{1 - \alpha}} \gamma + \tau_1 \frac{r^*}{1 - \tau_1} \left[ \frac{\alpha}{\frac{r^*}{1 - \tau_1} + c} \right]^{\frac{1}{1 - \alpha}} \right].$$

The first-order condition simplifies to

$$(1 - \gamma) \left( \frac{r^*}{1 - \tau_1} + c \right) - \frac{\tau_1 r^*}{(1 - \tau_1)(1 - \alpha)} = 0.$$

This yields the government's optimal tax rate

$$\tau^*(\gamma) = \frac{(1 - \alpha)(1 - \gamma)(r^* + c)}{r^* + (1 - \alpha)c(1 - \gamma)}.$$

The internal solution for the right-wing government's tax rate is negative,  $\tau^*(\gamma_R) < 0$ . We have earlier assumed the tax rates to be non-negative, therefore the conservative government chooses the corner solution and will set  $\tau_R = 0$ .

More generally, we see that the right-wing policy maker chooses a lower tax rate than its left-wing counterpart: This is simply because the agents, whose welfare the leftist administrator maximizes, are less productive, and thus their wages take a smaller weight in the objective function. Since wages are a decreasing function of the tax rate and both types of policy makers have the same Laffer-curve component, which peaks at a positive tax rate, in their maximizations, the government that maximizes for the less productive agents will necessarily select a higher optimal tax rate.

In this setup, let us only analyze sequences of governments where the second-period administration is left-wing and show that there will be policy persistence:

**Proposition 1** *A second-period left-wing government following a right-wing one will always*

*select a tax rate lower than a leftist government succeeding another left-wing administration.*

We begin the proof of Proposition 1 by showing the following:

**Lemma 1** *A left-wing government following a left-wing government will increase the tax rate as much as possible while keeping the capital stock unchanged.*

To prove Lemma 1 we first realize that a left-wing government following another leftist administration could induce an increase in the amount of capital employed by the firms, so that  $K_2^L > K_1^L$ . In this case, however, the firms would, on the margin, need to pay the required net return on capital  $r^*$  as well as the installation cost  $c$  - the same as in the first period. This in turn would imply that the government, with the same objective function as its predecessor, would induce the same amount of capital as the first-period administration, that is  $K_2^L = K_1^L$ . Therefore certainly no additional capital will be installed in the second period. The second-period government will thus induce an amount of capital  $K_2^L \leq K_1^L$ . At this range of amounts of capital, no new capital is being installed and thus the only expenditure the firms incur at the margin is  $r^*$ . However, for any  $K_2^L < K_1^L$ , the government will want to choose tax rate

$$\tau_2 = \frac{(1 - \alpha)(1 - \gamma/\bar{\gamma})r^*}{r^*} < \tau^*(\gamma_L).$$

We obtain this expression simply by setting  $c = 0$  in the formula for  $\tau^*(\gamma_L)$ . However, this would certainly imply an increase rather than decrease in the amount of capital installed, since then

$$K_1^L = \left[ \frac{\alpha}{\frac{r^*}{1 - \tau^*(\gamma_L)} + c} \right]^{\frac{1}{1 - \alpha}} X < \left[ \frac{(1 - \tau_2^L)\alpha}{r^*} \right]^{\frac{1}{1 - \alpha}} X = K_2^L.$$

Therefore the left-wing administration will choose to induce a corner solution, the only remaining option, where the amount of capital in the second period is the same as the first-period amount of capital. Given this constraint, however, it will set the maximum possible tax rate: The objective function of the government becomes linear in tax  $\tau_2$  when it is constrained to  $K_2^L = K_1^L$ . Specifically, the government's objective function becomes

$$\max_{\tau_2} \left[ (1 - \alpha)K_1^\alpha X^{1 - \alpha} \gamma + \frac{1}{2} \tau_2 r_2 K_1 \right],$$

where  $r_2$  is the second-period gross return on capital and  $r_2 = \alpha K_1^{\alpha - 1} X^{1 - \alpha}$ . This will yield an

always-positive first-order condition for any government:

$$\frac{1}{2}\alpha K_1 X^{1-\alpha} > 0.$$

Therefore the government will set the limiting tax rate  $\tau_2$  that we obtain from ensuring that  $K_2^L = K_1^L$ , where  $K_2^L$  results from a non-installation margin for the firms. Specifically,

$$K_1^L = \left[ \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right]^{\frac{1}{1-\alpha}} X = \left[ \frac{(1-\tau_2^L)\alpha}{r^*} \right]^{\frac{1}{1-\alpha}} X = K_2^L, \quad (1.1)$$

implying

$$\tau_2 = 1 - \frac{r^*}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} > \tau^*(\gamma_L).$$

A left-wing government following another leftist administration thus actually increases its tax rate. This then proves Lemma 1.

Now we consider a left-wing government coming into power after a conservative predecessor. In this case, two types of situation may arise. It may be that the amount of capital installed under the first-period right-wing administration is larger than the amount of capital employed by firms not installing any new capital and facing  $\tau^*(\gamma_L)$ . In other words, remembering that the first-period conservative government sets zero tax rate,

$$\left[ \frac{\alpha}{r^* + c} \right]^{\frac{1}{1-\alpha}} X > \left[ \frac{(1-\tau^*(\gamma_L))\alpha}{r^*} \right]^{\frac{1}{1-\alpha}} X.$$

In such a case, the government sets  $\tau_2 = \tau^*(\gamma_L)$  because it finds it optimal to induce a decrease in the amount of capital used by the firms.

Alternatively, if the converse holds, the government will again choose the maximum tax rate that still induces  $K_2^L = K_1^R$ , which in this case, analogous to the derivation in (1.1), will be

$$\tau_2 = 1 - \frac{r^*}{\frac{r^*}{1-\tau^*(\gamma_R)} + c}.$$

Since

$$\tau^*(\gamma_L) < 1 - \frac{r^*}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \text{ and } 1 - \frac{r^*}{\frac{r^*}{1-\tau^*(\gamma_R)} + c} < 1 - \frac{r^*}{\frac{r^*}{1-\tau^*(\gamma_L)} + c},$$

a second-period left-wing administration following a right-wing government will always select a tax rate lower than a leftist government following another left-wing administration. Thus we have shown policy persistence in the particular case of a second-period left-wing policy maker under the simple assumption of myopic firms and governments. This also concludes the proof of Proposition 1.

We shall now employ the diagram in Figure 1 to aid us in considering our mechanism more intuitively. The installation costs essentially create two welfare functions for the government as a function of the capital tax rate it chooses. One quasi-concave function relates the welfare to the tax rate when firms have to pay installation costs on the margin. This function applies when firms are installing additional capital. Therefore it is always the relevant function in the first period, which we begin with no capital installed by firms. It also applies in the second period whenever firms are, in equilibrium, bringing more capital into the economy than the stock already present from the previous period. This is the lower function graphed in Figure 1,  $W_l$ .

When no new capital is being installed, i.e. when firms in the second period are either remaining at the same level of capital stock or reducing the amount of capital they employ, firms no longer need to pay installation costs at the margin. Therefore for any level of tax rate, they will desire a higher capital stock. This in turn means that the welfare level will be higher for all tax rates when firms do not face installation costs. The welfare function when effectively  $c = 0$  is the upper function in Figure 1,  $W_h$ . The two welfare functions only meet when under either  $c > 0$  or  $c = 0$  there is no capital present in the economy, i.e. when the tax rate is 100%. Of course, the government would always prefer to be maximizing over  $W_h$ , rather than  $W_l$ . However, welfare values on  $W_h$  are only achievable if the corresponding amount of capital had already been installed in the previous period. Otherwise, only values on  $W_l$  are attainable. Therefore the shape of the full second-period welfare function will depend on how much capital is present in the economy at the start of the period.

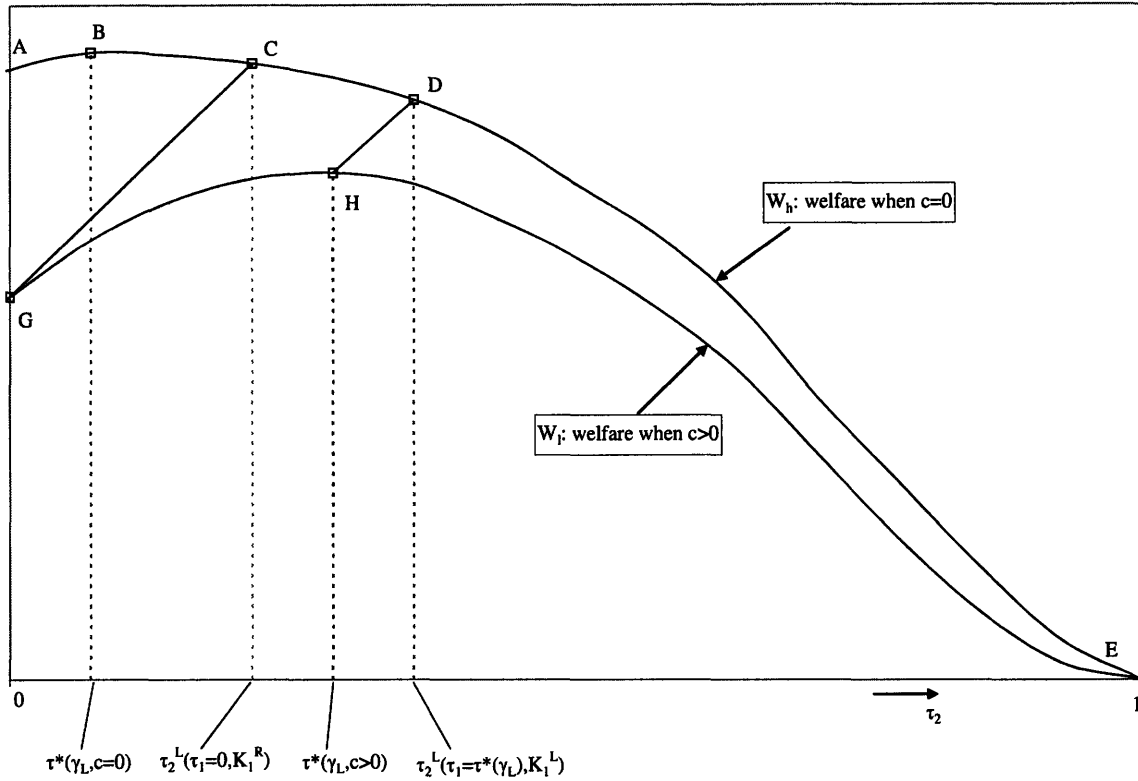


Figure 1-1: Welfare for the Second-Period Left-Wing Government

First-period left-wing government is maximizing solely over  $W_i$  graphed in Figure 1. Therefore it chooses tax rate  $\tau^*(\gamma_L, c > 0)$ , which brings it to point H on the graph. The second-period government following a first-period leftist administration then faces a composite welfare function: On the interval  $\tau_2 \in [0, \tau^*(\gamma_L, c > 0)]$ , it is on  $W_i$ , between G and H. In this interval it would set a tax rate lower than its predecessor, therefore more capital would be installed, firms would face installation costs and therefore capital levels and welfare would be on  $W_i$  for any tax rate.

Then for  $\tau_2 \in [\tau^*(\gamma_L, c > 0), \tau_2^L(\tau_1 = \tau^*(\gamma_L, c > 0), K_1^L)]$ , it is free to increase the tax rate without losing any capital in the economy. Welfare is represented by the linear segment HD. The tax rate in this interval is not low enough to attract additional capital, but not so high that any capital would leave the economy. This is the wedge caused by installation costs: When the

firms no longer face the cost of installation, they are willing to bear a higher tax rate and yet remain at the same level of capital stock. Because of this, welfare is linear and increasing in tax rate in this interval. Wages and tax base are constant, but tax revenue rises as taxes increase. If the government does in the end choose a tax rate on the segment HD (and it will), it will necessarily choose point D, because it will want to take advantage of the lack of installation costs faced by the firms.

In the last interval,  $\tau_2 \in [\tau_2^L(\tau_1 = \tau^*(\gamma_L, c > 0), K_1^L), 1]$ , when the welfare is between D and E, tax is sufficiently high that capital leaves the economy relative to the first period. Since the firms do not pay installation costs here, we are on  $W_h$ .

From the graph we can see (and we showed above) that the second-period government will choose the tax rate associated with point D. This kink point is the maximum of the combined welfare function. We know that  $\tau^*(\gamma_L, c > 0) > \tau^*(\gamma_L, c = 0)$ , that is, in terms of Figure 1, B is to the left of H. H is necessarily to the left of D, therefore B is also to the left of D. The maximum of  $W_h$  is attained at a lower tax rate than the rate, at which capital will start fleeing the country in the second period. Therefore the amount of capital at the maximum of  $W_h$  is higher than that already installed in the economy in the first period. But that means that  $W_h$  is not attainable at point B. The kink point D really affords the highest level of welfare and the government will set as high a tax as possible while inducing all the capital installed in the previous period to remain in the economy, so that  $K_2^L = K_1^L$ .

Now we realize that a first-period conservative government selects a tax rate lower than  $\tau^*(\gamma_L, c > 0)$ , because we showed earlier that  $\tau_R = 0 < \tau^*(\gamma_L, c > 0)$ . Therefore it attracts more capital than its first-period leftist counterpart. The welfare function of the second-period left-wing government now looks somewhat different. On the interval  $\tau_2 \in [0, \tau_2^L(\tau_1 = 0, K_1^R)]$ , it linearly increases, and thus is represented by the segment GC in Figure 1. Then it is on  $W_h$  between C and F, i.e. for tax rates  $\tau_2 \in [\tau_2^L(\tau_1 = 0, K_1^R), 1]$ .

As the graph is drawn, the second-period left-wing government will select point C and the associated tax rate. In other words, it will again set the highest possible tax rate that still keeps all the capital that was previously installed in the economy. This is conditional upon the maximum of  $W_h$ , B, being to the left of C. In the opposite case, i.e. if B were to the right of C, the government would select point B and tax rate  $\tau^*(\gamma_L, c = 0)$ . In this case the government

would be inducing a flight of capital in the second period, because it would consider the tax rate that preserves the previous period's installed capital to be too low.

We can see very clearly the workings of our mechanism: Segment GC is to the left of the parallel segment HD because the former is associated with the higher level of capital accumulated under a first-period conservative government whereas the latter with less capital attracted under a first-period leftist administration. Then the kink point C in the second-period government's welfare function is to the left of the kink point D. That is, the welfare maximum is achieved at a lower tax rate when more capital is present in the economy at the beginning of the second period due to a lower tax rate in the previous period.

As we noted, the mechanism also works when B is to the right of C, but still between C and D. Then the second-period government selects the kink point D with a higher tax rate when it follows a leftist predecessor. It chooses the maximum B and a lower tax rate when it succeeds a conservative administration.

When does the mechanism fail? Only if B were to the right of both C and D. That is, only if the maximum of  $W_h$  were attained at a tax rate, where capital installed in the first period had already started leaving the economy. Then no second-period government would select the kink point. Rather, whoever it would follow, the administration would choose point B and the associated tax rate  $\tau^*(\gamma_L, c = 0)$ . Since the fiscal policy would be chosen independently of the identity of the previous government and its tax rate, there would be no policy persistence. We already showed that this is not the case in our basic setup.

We can see that the mechanism is quite general and only dependent on three conditions: The two welfare functions, under  $c = 0$  and  $c > 0$ , are quasi concave. There is a positive difference between them for all tax rates  $\tau_2 \in [0, 1]$ . Finally, the tax maximizing of  $W_h$  is either lower than or not much higher than that maximizing  $W_l$ . Since, however, the mathematics of proving this involves rather complicated sufficient conditions without clear interpretations, we resort, throughout this paper, to employing specific functional forms.

### 1.2.3 Myopic Firms and Government: Extensions

#### Removal Costs

We could consider the alternative assumption of a cost of removal of capital instead of a cost of installation. Specifically, installing any amount of capital does not incur any additional cost in either the first or second period, but removing some of already installed capital at the beginning of the second period will cost the firms an amount  $c$  per unit of capital. While the mechanics of the more complete model, which we shall discuss below, are simpler under the installation-cost setup, let us first ensure that the qualitative results of the model would remain unchanged in the above simple environment and that Proposition 1 still holds.

A cost of removal assumption preserves a wedge between the firms' marginal costs depending on their desired capital stock in the second period relative to the first period. Above, the marginal cost difference was between installing additional capital on one hand and remaining at the same level or decreasing the level of capital on the other. Here, with a cost of removal, firms' marginal cost is higher when they remove capital versus a lower cost when they either install new capital or remain at the same capital level.

This wedge in marginal costs ensures that a not-too-left-wing leftist second-period government will still want to tax away the already-installed capital, without causing capital flight. Given that it inherits an economy with more capital installed by firms when it follows a right-wing administration, it will set a lower tax rate to achieve this than if its predecessor is another left-wing policy maker.

Mathematically, in this case, first-period firms maximize

$$\max_{K_1, L_1(\gamma)} K_1^\alpha X_1^{(1-\alpha)} - \sum_{\gamma} w_1(\gamma) L_1(\gamma) - \frac{r^*}{1 - \tau_1} K_1,$$

yielding

$$K_1 = \left[ \frac{(1 - \tau_1)\alpha}{r^*} \right]^{\frac{1}{1-\alpha}} X,$$

whereas second-period firms that are removing some of their capital will be maximizing

$$\max_{K_2, L_2(\gamma)} K_2^\alpha X_2^{(1-\alpha)} - \sum_{\gamma} w_2(\gamma) L_2(\gamma) - \frac{r^*}{1 - \tau_2} K_2 - c(K_1 - K_2),$$

which results in

$$K_2 = \left[ \frac{\alpha}{\frac{r^*}{1-\tau_2} - c} \right]^{\frac{1}{1-\alpha}} X.$$

First-period governments, after maximizing for their agents as before, will choose tax rates

$$\tau_1^L = (1 - \gamma_L)(1 - \alpha) \text{ and } \tau_1^R = 0.$$

Following the arguments presented above, a second-period left-wing government following another left-wing administration will set a tax rate

$$\tau_2^L = 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^L} + c},$$

whereas after a period of right-wing government it will select either

$$\tau_2^L = 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^R} + c},$$

or

$$\tau_2^L = (1 - \gamma_L)(1 - \alpha) = \tau_1^L.$$

We again know that

$$\tau_1^L < 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^L} + c} \text{ and } 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^R} + c} < 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^L} + c},$$

and therefore the same result holds as before under the assumption of installation costs: a second-period left-wing government elected after a right-wing administration will always select a tax rate lower than a leftist policy maker following another left-wing administration. Proposition 1 therefore holds.

### Convex Installation Costs

We may want to examine whether some of our simplifying assumptions affect the working of the mechanism we described above. While we could write the model in general terms or explore the simplifications simultaneously, this would yield rather complicated sufficient conditions,

which may not be easily interpretable. Therefore, in the setting of our simple model, i.e. with myopic firms and a second-period left-wing government, we shall first consider non-linear installation costs and then a constant elasticity of substitution production function with elasticity of substitution other than 1. In both cases we will want to again prove Proposition 1.

Here, we shall assume that the installation costs take the form  $ck^\rho$ , where  $c$  is as before,  $k \equiv K/X$  and  $\rho$  is a parameter such that  $\rho \geq 1$ . Of course, when  $\rho = 1$ , we have linear installation costs.

Firms' optimization problem now yields

$$\alpha k^{\alpha-1} - c\rho k^{\rho-1} = \frac{r^*}{1-\tau} \quad (1.2)$$

and

$$w = \gamma(1-\alpha)k^\alpha.$$

The first-period government's maximization problem is, as before,

$$\max_{\tau_1} \left[ \gamma(1-\alpha)k_1^\alpha + \frac{\tau_1 r^*}{1-\tau_1} k_1 \right].$$

This yields the first-order condition

$$\frac{\partial W}{\partial \tau_1} = \gamma(1-\alpha)\alpha k_1^{\alpha-1} \frac{\partial k_1}{\partial \tau_1} + \frac{r^*}{(1-\tau_1)^2} k_1 + \frac{\tau_1 r^*}{1-\tau_1} \frac{\partial k_1}{\partial \tau_1} = 0.$$

Here, we can obtain  $\frac{\partial k_1}{\partial \tau_1}$  by totally differentiating (1.2), which yields

$$\frac{\partial k_1}{\partial \tau_1} = \frac{r^*}{(1-\tau_1)^2 \left[ \alpha(\alpha-1)k_1^{\alpha-2} - c\rho(\rho-1)k_1^{\rho-2} \right]}. \quad (1.3)$$

Then the FOC simplifies to

$$(\gamma-1)(1-\alpha)\alpha k_1^{\alpha-1} + c\rho(1-\rho)k_1^{\rho-1} + \frac{\tau_1 r^*}{1-\tau_1} = 0. \quad (1.4)$$

We then obtain that the second derivative, evaluated at the point, where (1.4) holds, is negative

whenever

$$(1 - \alpha)\alpha k_1^{\alpha-2}[(1 - \alpha)(1 - \gamma) - 1] < c\rho^2(\rho - 1)k_1^{\rho-2}.$$

Since  $(1 - \alpha)(1 - \gamma) - 1 < 0$ , this always holds, and therefore any local extremum of the welfare function must be a local maximum. Since the function is continuously differentiable on  $\tau_1 \in (0, 1)$ , it follows that it is quasi-concave on that interval.

Now we want to know when our mechanism functions in this altered environment. As we mentioned earlier, it only needs to be verified that the capital stock remains unchanged under a left-wing government following a leftist administration.

Remember, before we had the second-period left-wing government succeeding a right-wing administration taxing the capital at as high a rate as possible while not inducing any of the capital to leave the economy. Or, if it inherited "too much" capital in the economy, it chose its optimal tax rate from the maximization problem with firms not facing any installation costs, because at that tax rate capital was flowing out of the country rather than additional capital being installed.

However, when a left-wing policy maker followed another leftist government, he always set the tax rate at the higher possible level that induced the firms to remain at their capital levels from the first period. This ensured that this policy maker set a higher tax rate than the one following a right-wing government.

Therefore a sufficient condition for our mechanism to work is that a left-wing government after another leftist administration prefers to maintain the same capital level to choosing an even higher tax rate and inducing the capital to leave. Because if that were the case, he would also certainly prefer less capital and a higher tax rate after a right-wing administration, thus choosing the same fiscal policy after both. Remember, in terms of Figure 1, we need that the second-period leftist government chooses point D.

Mathematically, we want to ensure that  $\partial k_1 / \partial c < 0$ , so that the amount of capital in the economy at the optimal tax rate of the government maximizing under the assumption of  $c = 0$  is higher than when the government optimizes for a positive installation cost. This is to say that more capital would be installed at point B than at point H. But since H and D have the same amount of capital, we are also saying that at B there is more capital in place than at kink point D. But B is not attainable in that case: On  $W_h$ , only points with less capital than

at D are achievable. Therefore when  $\partial k_1/\partial c < 0$ , we necessarily have that B is to the left of D, which we earlier showed to be sufficient for the functioning of our mechanism.

This will ensure that a left-wing policy maker following another leftist government will always choose the corner solution (kink point D) and the economy will retain the same level of capital it attracted in the first period. We already know from (1.3) that  $\partial k_1/\partial \tau_1 < 0$ , therefore we want to show that  $\partial \tau_1/\partial c > 0$ . Totally differentiating (1.4), we obtain that

$$\partial \tau_1/\partial c > 0 \Leftrightarrow \frac{(1-\alpha)(1-\gamma) - (1-\rho)}{(1-\alpha)\alpha k_1^{\alpha-2}[(1-\alpha)(1-\gamma) - 1] - c\rho^2(\rho-1)k_1^{\rho-2}} < 0,$$

which, again, holds for all  $\rho \geq 1$ .

From (1.4) we see that holding the tax rate constant,

$$\frac{\partial W}{\partial \tau_1 \partial \gamma} = (1-\alpha)\alpha k_1^{\alpha-1} > 0.$$

Since the welfare function is quasi-concave, we still have that the right-wing government will choose a lower tax rate in the first period. Then a second-period left-wing government will either choose to keep all the capital from the first period in the economy, or induce some of it to leave. In both cases though it will choose a lower tax rate than were it to follow another left-wing administration and Proposition 1 holds.

### Constant Elasticity of Substitution Production Function

We can perform a very similar analysis to the one above for a constant elasticity of substitution function and constant installation costs to verify that our basic mechanism does not depend on our original assumption of an elasticity of substitution of one between labor and capital.

Our production function now changes to

$$f(k) = (\alpha k^s + 1 - \alpha)^{1/s},$$

where  $\sigma \equiv \frac{1}{1-s}$  is the elasticity of substitution between the factors of production. Firms'

optimization yields an explicit expression for capital stock in the economy,

$$k_1 = \left[ \frac{\left( \frac{1}{\alpha} \left( \frac{r^*}{1-\tau_1} + c \right) \right)^{\frac{s}{1-s}} - \alpha}{1-\alpha} \right]^{-\frac{1}{s}}$$

The government's FOC simplifies to

$$(1-\gamma) \left[ \left( \frac{1}{\alpha} \left( \frac{r^*}{1-\tau_1} + c \right) \right)^{\frac{s}{1-s}} - \alpha \right] - \frac{\tau_1 r^*}{(1-\tau_1)(1-s)\alpha^{\frac{s}{1-s}}} \left( \frac{r^*}{1-\tau_1} + c \right)^{\frac{s}{1-s}-1} = 0. \quad (1.5)$$

Once again, it can be shown that the second derivative of the welfare function is negative at any tax rate that satisfies the above condition, therefore also that the welfare function is quasi-concave.

As before, we want to determine whether  $\partial k_1 / \partial c < 0$ . We can easily see that

$$\partial k_1 / \partial c < 0 \Leftrightarrow \frac{\partial \left( \frac{r^*}{1-\tau_1} + c \right)}{\partial c} < 0 \Leftrightarrow \partial \tau_1 / \partial c < -\frac{(1-\tau_1)^2}{r^*} \quad (1.6)$$

Total differentiation of (1.5) yields

$$\partial \tau_1 / \partial c = -\frac{\frac{(1-\tau_1)^2}{r^*} \left[ \left( \frac{r^*}{1-\tau_1} + c \right) [(1-\gamma)s-1] - \left( \frac{s}{1-s} - 1 \right) \frac{\tau_1 r^*}{1-\tau_1} \right]}{\left( \frac{r^*}{1-\tau_1} + c \right) [(1-\gamma)s-1] - \left( \frac{s}{1-s} - 1 \right) \frac{\tau_1 r^*}{1-\tau_1} - \left( \frac{r^*}{1-\tau_1} + c \right) \frac{1}{1-s}}$$

Then (1.6) holds whenever

$$-\left( \frac{r^*}{1-\tau_1} + c \right) \frac{1}{1-s} < 0.$$

Since this is always satisfied, we again have that the second-period left-wing government chooses to retain all capital installed in the economy under a left-wing predecessor. This then implies that our mechanism in the basic model continues to work under a more general production function specification. Left-wing government selects a higher tax rate following another leftist administration than were it to follow a conservative policy maker and Proposition 1 still holds.

### 1.3 Forward-Looking Firms and Government

Now we need to consider the more realistic case where the firms and the government are forward-looking and they take into account the future period when making their first-period decisions. Specifically, we will assume that the firms maximize their profits over both periods, since by installing capital in the first period at a cost they clearly make a two-period investment decision. They discount the future period at the rate  $\beta$ .

The governments still only maximize the utility of their agents in the period when they are in power. This means first- and second- period governments maximizing for the same  $\gamma$  have identical objective functions. In other words, the second-period utility of the agents with productivity  $\gamma$  does not enter the objective function of a first-period government. But the first-period administrations do realize the impact of their policies on the decisions taken in the future period and thus on the expectations regarding those decisions formed in the first period. In this sense they are forward-looking: They take the future period's variables to be functions of its first-period decision regarding fiscal policy.

Let us define  $R \equiv \gamma_L/\gamma_R$  to be a measure of how distant the two types of government are from each other. In the environment of forward-looking firms and government we will then want to show the following:

**Proposition 2** *There is a cutoff value  $\tilde{R}_2 \in (0, 1)$  such that for any  $R > \tilde{R}_2$  there will be policy persistence for both types of government: When the right- and left-wing government types are sufficiently close to each other, the right will be more left-wing when it follows a left-wing government than otherwise and the left will be more right-wing when it succeeds a right-wing administration than otherwise. Moreover, there is a cutoff value  $R_R \in (0, 1)$  such that for any  $R < \min\{\tilde{R}_2, R_R\}$  there will be policy persistence for the left-wing and policy reversal for the right-wing government: When the right- and left-wing government types are sufficiently far from each other, the left will be more right-wing when it follows a right-wing administration than otherwise, but the right will be even more right-wing when it succeeds a left-wing government than otherwise.*

We begin proving Proposition 2 by realizing that our equilibrium has changed somewhat: Firms take the tax rates as given in each period and the governments take the response func-

tions of the firms as given. More specifically, first-period firms observe  $\tau_1$  set by the first-period government and form correct expectations regarding  $\tau_2$  under either type of subsequent government that will be induced by the capital accumulated by firms in the first period.

We can solve for the equilibrium backwards. Second-period firms take their already-installed capital and second-period tax rate as given, choosing an optimal second-period capital level,  $K_2 = \kappa_2(K_1, \tau_2)$ . Second-period government maximizes its objective function, taking the firms' reaction function as given, arriving at  $\tau_2 = \tau^*(\gamma, K_1)$ . First-period firms take  $\tau_1$  as given, yielding an optimal amount of capital installed in the first period,  $K_1 = \kappa_1(\tau_1)$ . In equilibrium this will also depend on the firms' forecasts of the second period tax rates  $\tau_2(\gamma_L)$  and  $\tau_2(\gamma_R)$ , as well as  $p_L$  and  $p_R$ , the probabilities of a left-wing and a right-wing government coming to power in the second period, respectively. Combining the functions  $\tau^*(\cdot)$  and  $\kappa_1(\cdot)$  yields  $K_1$  as a function of the first-period tax rate only, which the first-period government can use in its maximization problem, selecting a value for  $\tau_1$ .

We begin our solution by assuming that the first-period government takes as given the fact whether additional capital will be installed, capital will remain at the same level, or some capital will be uninstalled in the second period. Later, with a few additional assumptions, we will relax this condition to arrive at the complete solution, where the first-period policy maker realizes he can affect which one of the three above possibilities actually occurs.

Firms in the second period thus select

$$K_2 = \left[ \frac{(1 - \tau_2)\alpha X^{1-\alpha}}{r^*} \right]^{\frac{1}{1-\alpha}} \quad \text{if } K_2 \leq K_1,$$

$$K_2 = \left[ \frac{\alpha X^{1-\alpha}}{\frac{r^*}{1-\tau_2} + c} \right]^{\frac{1}{1-\alpha}} \quad \text{if } K_2 > K_1$$

and

$$K_2 = K_1 \quad \text{if} \quad \left[ \frac{(1 - \tau_2)\alpha X^{1-\alpha}}{r^*} \right]^{\frac{1}{1-\alpha}} > K_1 \quad \text{and} \quad \left[ \frac{\alpha X^{1-\alpha}}{\frac{r^*}{1-\tau_2} + c} \right]^{\frac{1}{1-\alpha}} \leq K_1.$$

Therefore let us solve the problem by looking at cases differentiated by whether the second-period government selects a tax rate that induces the firms to increase, decrease or retain the same level of capital.

### 1.3.1 First-Period Left-Wing Government

When the first-period government is left-wing, we can show the following:

**Lemma 2** *Let us assume that the first-period government takes as given the fact whether additional capital will be installed, capital will remain at the same level, or some capital will be uninstalled in the second period. Then the stock of capital remains unchanged when a left-wing government follows another leftist administration. There is a cutoff value  $\tilde{R} \in (0, 1)$  such that when a right-wing government succeeds a left-wing one, for all  $R \geq \tilde{R}$  the stock of capital also does not change whereas for all  $R < \tilde{R}$  it increases.*

To prove Lemma 2, let us first consider an equilibrium with a first-period left-wing government that is followed either by another left-wing government under which firms have the same level of capital, or by a right-wing government that induces the firms to increase their capital stock. Symbolically,  $K_2^R > K_1^L$  and  $K_2^L = K_1^L$ . Then the first-period firms maximize

$$\begin{aligned} & \max_{K_1, L_1(\gamma)} ((K_1^L)^\alpha (X_1^L)^{1-\alpha} - \sum w_1^L(\gamma) L_1^L(\gamma) - \left( \frac{r^*}{1 - \tau_1^L} + c \right) K_1^L) \\ & + \beta [p_L ((K_1^L)^\alpha (X_2^L)^{1-\alpha} - \frac{r^*}{1 - \tau_2^L} K_1^L - \sum w_2^L(\gamma) L_2^L(\gamma)) \\ & + p_R ((K_2^R)^\alpha (X_2^R)^{1-\alpha} - \left( \frac{r^*}{1 - \tau_2^R} + c \right) (K_2^R - K_1^L) - \frac{r^*}{1 - \tau_2^R} K_1^L - \sum w_2^R(\gamma) L_2^R(\gamma))]. \end{aligned}$$

This yields a first-period capital stock

$$K_1^L = \left[ \frac{\alpha(1 + \beta p_L)}{\frac{r^*}{1 - \tau_1^L} + \frac{\beta p_L r^*}{1 - \tau_2^L} + c(1 - \beta p_R)} \right]^{\frac{1}{1-\alpha}} X.$$

If  $K_2^R > K_1^L$ , additional capital is installed under a second-period right-wing government and therefore necessarily  $\tau_2^R = 0$ . Also, since  $K_2^L = K_1^L$ , we have

$$1 - \tau_2^L = \frac{r^*}{\frac{r^* + c}{1 - \tau_1^L} - \beta p_R c},$$

and therefore

$$K_1^L = \left[ \frac{\alpha}{\frac{r^*}{1-\tau_1^L} + c(1-\beta p_R)} \right]^{\frac{1}{1-\alpha}} X.$$

First-period left-wing government then maximizes

$$\max_{\tau_1} \left[ w_1(\gamma_L) + \frac{1}{2} \tau_1 \tau_1 K_1 \right],$$

which yields the optimal tax rate

$$\tau_1^L = \frac{(1-\gamma_L)[r^* + c(1-\beta p_R)]}{r^*/(1-\alpha) + (1-\gamma_L)c(1-\beta p_R)}.$$

Now we need to determine under which conditions it is in fact true that  $K_2^R > K_1^L$  and  $K_2^L = K_1^L$ . A set of sufficient conditions for  $K_2^L = K_1^L$  is that  $\tau_2^L \geq \tau^*(\gamma_L, c = 0)$  and  $\tau^*(\gamma_L, c > 0) \geq \tau_2^{ent}$ , where  $\tau_2^{ent}$  is the tax rate at which additional capital would start entering the economy. Under these conditions the amount of capital installed by the firms neither increases nor decreases. Importantly, when the two conditions hold, the government finds it optimal to in fact set the second-period tax rate equal to  $\tau_2^L$ .

The government's welfare function is once again composed of three sections: For taxes lower than  $\tau_2^{ent}$  it is on a quasi-concave schedule peaking at  $\tau^*(\gamma_L, c > 0)$ , because firms are installing capital at these tax rates and therefore their marginal cost of capital includes  $c$ . For taxes above  $\tau_2^L$  it is on a quasi-concave curve peaking at  $\tau^*(\gamma_L, c = 0)$ , because for these tax rates, firms are reducing the amount of installed capital. We can easily see that  $\tau^*(\gamma_L, c > 0) > \tau^*(\gamma_L, c = 0)$ . For a given tax rate, any point on the latter curve is above the corresponding point on the former schedule, because at the marginal cost for firms without  $c$ , the firms will have more capital installed at any tax rate. In the interval  $[\tau_2^{ent}, \tau_2^L]$  the welfare function is linear and increasing: here, the government is free to increase the tax rate without any reduction in capital stock by the firms.

Therefore when  $\tau_2^L \geq \tau^*(\gamma_L, c = 0)$  holds, the government prefers to induce no reduction in capital stock and  $\tau^*(\gamma_L, c > 0) \geq \tau_2^{ent}$  ensures that it will certainly not cause an increase in the amount of capital installed.

We have that

$$\begin{aligned}\tau_2^L &\geq \tau^*(\gamma_L, c = 0) \Rightarrow \frac{r^*(1-\alpha)(1-\gamma_L) + c(1-\beta p_R)}{r^* + c(1-\beta p_R)} \geq (1-\alpha)(1-\gamma_L) \\ &\Rightarrow 1 \geq (1-\alpha)(1-\gamma_L),\end{aligned}$$

which always holds. Similarly,

$$\begin{aligned}\tau_2^{ent} &\leq \tau^*(\gamma_L, c > 0) \Rightarrow \frac{(r^* + c)(1-\alpha)(1-\gamma_L) - \beta p^R c}{r^* - \beta p^R c + (1-\alpha)(1-\gamma_L)c} \leq \frac{(1-\alpha)(1-\gamma_L)(r^* + c)}{r^* + (1-\alpha)(1-\gamma_L)c} \\ &\Rightarrow 1 \geq (1-\alpha)(1-\gamma_L).\end{aligned}$$

We now need to consider the conditions necessary for  $K_2^R > K_1^L$ . For this to occur, the second-period right-wing government must see setting a tax that will induce additional installation of capital as more beneficial than a tax rate that results in the capital level remaining at the level inherited from the previous period. Namely

$$welfare^R[K_2^R > K_1^L] > welfare^R[K_2^R = K_1^L]. \quad (1.7)$$

Note that a right-wing government will never want to induce a decrease in the level of capital if the left-wing government, facing the same conditions, desires the capital level to remain at the same level.

We defer any further calculation regarding (1.7) until we have first investigated the case when  $K_2^R = K_1^L$ . Here, both types of government will set the same tax rate in the second period. To determine this tax rate, we first need to solve for the optimal level of first-period capital installed by the firms that expect to retain the exact same level of capital in the second period regardless of the type of government. Therefore the firms maximize

$$\begin{aligned}\max_{K_1, L_1(\gamma)} &\left[ (K_1^L)^\alpha (X_1^L)^{1-\alpha} - \sum w_1^L(\gamma) L_1^L(\gamma) - \left( \frac{r^*}{1-\tau_1^L} + c \right) K_1^L + \right. \\ &\left. + \beta \left( (K_1^L)^\alpha (X_2)^{1-\alpha} - \frac{r^*}{1-\tau_2} K_1^L - \sum w_2(\gamma) L_2(\gamma) \right) \right] \quad (1.8)\end{aligned}$$

This results in a first-period capital stock

$$K_1^L = \left[ \frac{\alpha(1+\beta)}{\frac{r^*}{1-\tau_1^L} + \frac{\beta r^*}{1-\tau_2} + c} \right]^{\frac{1}{1-\alpha}} X.$$

Since, under both possible second-period administrations,  $K_2 = K_1^L$ , we have

$$\tau_2^L = \tau_2^R = 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^L} + c}.$$

Subsequently

$$K_1^L = \left[ \frac{\alpha}{\frac{r^*}{1-\tau_1^L} + c} \right]^{\frac{1}{1-\alpha}} X.$$

From the government maximization problem, we obtain

$$\tau_1^L = \tau^*(\gamma_L) = \frac{(1-\alpha)(1-\gamma_L)(r^*+c)}{r^*+(1-\alpha)c(1-\gamma_L)}.$$

As before, a set of sufficient conditions for  $K_2^L = K_1^L$  is that  $\tau_2^L \geq \tau^*(\gamma_L, c=0)$  and  $\tau^*(\gamma_L, c>0) \geq \tau_2^{ent}$ . Here we have

$$\begin{aligned} \tau_2^L &\geq \tau^*(\gamma_L, c=0) \Rightarrow \frac{(1-\alpha)(1-\gamma_L)r^*+c}{r^*+c} \geq \frac{(1-\alpha)(1-\gamma_L)(r^*+c)}{r^*+(1-\alpha)c(1-\gamma_L)} \Rightarrow \\ &\Rightarrow 1 \geq (1-\alpha)(1-\gamma_L), \end{aligned}$$

which always holds. Since in this case,  $\tau_2^{ent} = \tau^*(\gamma_L, c>0)$ , it is always true that  $\tau^*(\gamma_L, c>0) \geq \tau_2^{ent}$ .

Now we return to considering condition (1.7), since it is precisely when it does not hold that we have  $K_2^R = K_1^L$ . The inequality translates to

$$w_2(\gamma_R, \tau_2=0) > w_2(\gamma_R, \tau_2=\tau_2^L) + \frac{1}{2}\tau_2^L \frac{r^*}{1-\tau_2^L} K_1^L.$$

This further yields, after some manipulation,

$$1 > \left( \frac{r^* + c}{\frac{r^*}{1-\tau_1^L} + c} \right)^{\frac{\alpha}{1-\alpha}} \cdot \left[ 1 + \left( 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^L} + c} \right) \frac{\alpha}{2(1-\alpha)} (1+R) \right]. \quad (1.9)$$

Now we may plug in for  $\tau_1^L$ , obtaining

$$1 > \left( 1 - (1-\alpha) \frac{1-R}{1+R} \right)^{\frac{\alpha}{1-\alpha}} \cdot \left[ 1 + \frac{r^*(1-\alpha) \frac{1-R}{1+R} + c}{r^* + c} \cdot \frac{\alpha}{2(1-\alpha)} (1+R) \right].$$

We cannot solve for a simple closed-form expression for  $R$ , at which (1.9) holds. However, let us show that (1.9) is satisfied for all  $R \in [0, \tilde{R}]$ , whereas it does not hold for any  $R \in (\tilde{R}, 1]$ , where  $\tilde{R} \in [0, 1)$  is implicitly defined by

$$1 = \left( 1 - (1-\alpha) \frac{1-\tilde{R}}{1+\tilde{R}} \right)^{\frac{\alpha}{1-\alpha}} \cdot \left[ 1 + \frac{r^*(1-\alpha) \frac{1-\tilde{R}}{1+\tilde{R}} + c}{r^* + c} \cdot \frac{\alpha}{2(1-\alpha)} (1+\tilde{R}) \right],$$

or, if no solution exists, it is set to  $\tilde{R} = 0$ .

At  $R = 1$ , (1.9) becomes  $1 > 1 + \frac{c}{r^*+c} \cdot \frac{\alpha}{(1-\alpha)}$ , therefore (1.9) is never satisfied when  $R = 1$ . On the other hand, at  $R = 0$ , (1.9) simplifies to  $1 > \alpha^{\frac{\alpha}{1-\alpha}} \left[ 1 + \frac{r^*(1-\alpha)+c}{r^*+c} \cdot \frac{\alpha}{2(1-\alpha)} \right]$ , which may or may not hold; in fact, (1.9) will hold at  $R = 0$  for some reasonable values of parameters.

Let us now determine whether the right-hand side (RHS) of (1.9) is increasing. This will, of course, be the case when its first derivative is positive, i.e. when, after some algebra,

$$R^2(2-\alpha)(c-r^*(1-\alpha)) + 2R(1+\alpha)(c-r^*(1-\alpha)) + 4(r^*+c) - 3\alpha r^* - \alpha c - \alpha^2 r^* > 0.$$

When  $c - r^*(1-\alpha) \geq 0$ , i.e.  $r^* \leq \frac{c}{1-\alpha}$ , this always holds, which means that the RHS of (1.9) is increasing. When, on the other hand,  $r^* > \frac{c}{1-\alpha}$ , we have two zeros,

$$R_{1,2} = \frac{-(1+\alpha) \pm \sqrt{(1+\alpha)^2 - \frac{(2-\alpha)(4(r^*+c)-3\alpha r^*-\alpha c-\alpha^2 r^*)}{c-r^*(1-\alpha)}}}{2-\alpha}.$$

These are real when  $r^* \geq \frac{8\alpha-7}{9(1-\alpha)}c$ . Since  $\frac{c}{1-\alpha} > \frac{8\alpha-7}{9(1-\alpha)}c$ , they are always real in the second

range for parameter  $r^*$ . Since one of  $R_1, R_2$  is certainly negative, we know that the RHS of (1.9) is either increasing for  $r^* > \frac{c}{1-\alpha}$ , else it is increasing and subsequently decreasing on the range  $R \in [0, 1]$ .

Given the above, we have indeed determined that (1.9) is either never satisfied - when it does not hold even at  $R = 0$ , or it holds up to a certain value of  $R < 1$ . In other words, (1.9) is satisfied for all  $R \in [0, \tilde{R}]$ , whereas it does not hold for any  $R \in (\tilde{R}, 1]$ . This concludes the proof of Lemma 2.

### Manipulating the Successor

Now we turn our attention to the fact that the first-period left-wing government's policy choice can in fact affect whether the amount of capital installed in the second period increases or remains at the same level - and that the policy maker realizes this when setting  $\tau_1^L$ . In other words, contrary to what we assumed so far, a forward-looking first-period administration does not take as given whether its successor increases, decreases or maintains the level of capital in the economy. The first-period government knows its decision will affect which of those three actions will be induced by the second-period government. Expectations of a particular action - increasing, decreasing, or maintaining the level of capital - by the second-period firms can have a discrete effect on the first-period welfare of agents. Then the first-period government may set a policy specifically designed to cause the second-period government to induce that action, the expectations of which will be the most beneficial to agents in the first period.

Considering the above solution, we see immediately that for any tax level set in the first period, more capital is installed in the economy if the firms expect the second-period right-wing government to induce an increase in the amount of capital. Therefore, if the relative position of the two types of government is precisely  $\tilde{R}$ , the first-period government will certainly prefer the case when  $K_2^R > K_1^L$ . But by the same argument, if we are at  $\tilde{R} + \varepsilon$ , where  $\varepsilon > 0$  is a small number, the first-period left-wing administrator will want to "deviate" and set a tax rate slightly higher than the one suggested by the above solution so as to induce the second-period right-wing government that might follow him to choose  $\tau_2^R = 0$  and therefore cause an increase in the amount of capital installed by the firms.

In fact, the same could theoretically occur with a first-period left inducing a second-period

left or a first-period right inducing the next period's right to cause an additional capital installation. However, we shall rule out such options by assumption. It is unreasonable to expect an administration to set a higher tax rate in a given period simply in order to create the expectation of an administration with the same preferences setting a lower tax rate in the next period. More importantly, the mere possibility of this occurring is simply a function of our model being a finite, two-period one. Imagine an infinite-horizon model where every period with the same amount of capital installed at the start of that period will lead to the same policy result. In such a setting, it is impossible for, say, a left-wing administration that is causing capital to enter in a given period to induce additional capital installation in the next period, unless some underlying variables or parameters changed. Therefore agents and, crucially, firms are not going to expect such an additional installation. Thus either administration cannot make firms believe it will induce more capital being installed in the next period under an administration with the same preferences. While for ease of computation and exposition, we only consider a two-period model, it would be unrealistic not to rule out such an option by assumption.

We return to the case where a left-wing government sufficiently close in its preferences to its potential right-wing successor will deliberately choose a higher tax rate in order to cause that successor to induce a capital increase in the economy. In fact, they will choose the smallest possible tax rate that is high enough such that that happens. If we call that tax rate  $\tau_1^{man}$ , where  $\tau_1^{man} > \tau^*(\gamma_L)$  (the manipulation occurs), we can show that such a "manipulation" of the successor will occur for  $R \in (\tilde{R}, \tilde{R}_2]$ , where  $\tilde{R}_2 \in (\tilde{R}, 1)$ . Please see the appendix for the proof of this result.

### 1.3.2 First-Period Right-Wing Government

Let us now proceed to the case when the first-period government is right-wing. In the second period, a right-wing government will always set a tax rate that induces the firms to remain at the same level of capital as in the first period, because as we will see below, a first-period right-wing government always sets a zero tax rate. Therefore we will focus on the conditions for the two options regarding a second-period left-wing government: a second-period left-wing government can also set a tax rate that preserves the level of capital inherited from the first period, or it can induce a reduction of firms' capital stock.

When  $K_1^R = K_2^L$ , we have firms maximizing (1.8), yielding

$$\tau_1^R = \tau^*(\gamma_R) = \frac{(1-\alpha)(1-\gamma_R)(r^*+c)}{r^*+(1-\alpha)c(1-\gamma_R)}.$$

Since this expression is negative, as before, the first-period tax rate is  $\tau_1^R = 0$ . Then

$$K_1 = \left[ \frac{\alpha}{r^*+c} \right]^{\frac{1}{1-\alpha}} X$$

and subsequently we have

$$1 - \tau_2^L = 1 - \tau_2^R = (1 - \tau_1^R) \frac{r^*}{r^*+c}, \text{ thus } \tau_2^L = \tau_2^R = \frac{c}{r^*+c}.$$

To have in fact  $K_1^R = K_2^L$ , it is sufficient that  $\tau_2^L > \tau^*(\gamma_L)$  and  $\tau^*(\gamma_L) > \tau_2^{ent}$ . Since  $\tau_2^{ent} < 0$ , the latter condition holds always. The former holds whenever

$$\frac{c}{r^*+c} > \frac{(1-\alpha)(1-\gamma_L)(r^*+c)}{r^*+(1-\alpha)c(1-\gamma_L)},$$

namely, whenever the governments' types are sufficiently close to each other. This translates into

$$R > \frac{(1-\alpha)(r^*+2c) - c}{(1-\alpha)(r^*+2c) + c} \equiv R_R.$$

On the other hand, when  $K_1^R > K_2^L$ , firms are maximizing

$$\begin{aligned} & \max_{K_1, L_1(\gamma)} ((K_1^R)^\alpha (X_1^R)^{1-\alpha} - \sum w_1^R(\gamma) L_1^R(\gamma) - \left( \frac{r^*}{1-\tau_1^R} + c \right) K_1^R) \\ & + \beta [p_L ((K_2^L)^\alpha (X_2^L)^{1-\alpha} - \frac{r^*}{(1-\tau_2^L)} K_2^L - \sum w_2^L(\gamma) L_2^L(\gamma)) \\ & + p_R ((K_1^R)^\alpha (X_2^R)^{1-\alpha} - \frac{r^*}{(1-\tau_2^R)} K_1^R - \sum w_2^R(\gamma) L_2^R(\gamma))]. \end{aligned}$$

This yields a first-period capital stock

$$K_1^L = \left[ \frac{\alpha(1 + \beta p_R)}{\frac{r^*}{1 - \tau_1^R} + \frac{\beta p_R r^*}{1 - \tau_2^R} + c} \right]^{\frac{1}{1 - \alpha}} X.$$

We have

$$1 - \tau_2^R = \frac{r^*}{\frac{r^*}{1 - \tau_1^R} + c},$$

and therefore

$$K_1^R = \left[ \frac{\alpha}{\frac{r^*}{1 - \tau_1^R} + c} \right]^{\frac{1}{1 - \alpha}} X,$$

which in turn implies  $\tau_1^R = 0$ . For firms to decrease their capital stock under a second-period left-wing government, it is sufficient that  $\tau_2^L < \tau^*(\gamma_L)$ , which means

$$R < \frac{(1 - \alpha)(r^* + 2c) - c}{(1 - \alpha)(r^* + 2c) + c} = R_R.$$

We have thus shown that if the first-period government is right-wing, a second-period right-wing government selects a tax rate that induces firms to remain at the same level of capital stock. If a second-period left-wing government is close enough in its objective function to the right-wing, it will do the same, else it will induce a decrease in the firms' capital.

### 1.3.3 Summary of Results

We are now ready to summarize our results in the following tables.

Table1: Second-Period Tax Rates After a First-Period Left-Wing Government

Tax after L1	$R \in [0, \tilde{R}]$	$R \in [\tilde{R}, \tilde{R}_2]$	$R \in [\tilde{R}_2, 1]$
$\tau_2^L$	$\frac{r^*(1 - \alpha)(1 - \gamma_L) + c(1 - \beta p_R)}{r^* + c(1 - \beta p_R)}$	$> \frac{r^*(1 - \alpha)(1 - \gamma_L) + c(1 - \beta p_R)}{r^* + c(1 - \beta p_R)}$	$\frac{(1 - \alpha)(1 - \gamma_L)r^* + c}{r^* + c}$
$\tau_2^R$	0	0	$\frac{(1 - \alpha)(1 - \gamma_L)r^* + c}{r^* + c}$

Table 2: Second-Period Tax Rates After a First-Period Right-Wing Government

Tax after R1	$R \in [0, R_R]$	$R \in [R_R, 1]$
$\tau_2^L$	$(1 - \alpha)(1 - \gamma_L)$	$\frac{c}{r^* + c}$
$\tau_2^R$	$\frac{c}{r^* + c}$	$\frac{c}{r^* + c}$

Now we only need to compare the various tax rates to determine when we in fact observe policy persistence.

It is clear that for a second-period right-wing government, there will be policy persistence when  $R > \tilde{R}_2$ . In that case, it will set a higher tax rate following a left-wing administration than otherwise. On the other hand, when  $R < \tilde{R}_2$  we will have policy reversal: If a right-wing government follows a left-wing one, it will set a more "extreme" right-wing policy. In other words, if the preferences of the two policy makers are similar enough, a left-wing administrator will induce its right-wing successor to choose more left-wing policies. However, if the two types of government are sufficiently different from each other, a second-period right-wing policy maker will select more right-wing policies to "address" the state of the economy he inherits from his predecessor.

The results are somewhat more complicated for a second-period left-wing government. Since

$$\frac{r^*(1 - \alpha)(1 - \gamma_L) + c(1 - \beta p_R)}{r^* + c(1 - \beta p_R)} > (1 - \alpha)(1 - \gamma_L), \text{ and}$$

$$\frac{(1 - \alpha)(1 - \gamma_L)r^* + c}{r^* + c} > (1 - \alpha)(1 - \gamma_L),$$

there will be policy persistence when  $R < R_R$ . Similarly, since

$$\frac{(1 - \alpha)(1 - \gamma_L)r^* + c}{r^* + c} > \frac{c}{r^* + c},$$

we will also have policy persistence when  $R > R_R$  and  $R > \tilde{R}_2$ .

Thus as long as  $R_R \geq \tilde{R}_2$ , there will always be policy persistence impacting a second-period left-wing government. However, if  $R_R < \tilde{R}_2$ , the results are ambiguous when  $R \in [R_R, \tilde{R}_2]$ . This also concludes the proof of Proposition 2.

Therefore we have two important results. First, when a country is governed by policy

makers, who have sufficiently similar preferences, we will observe policy persistence: A left-wing administrator following a right-wing one will shift his policies to the right, and a right-wing policy maker who succeeds a leftist government will choose more left-wing policies.

Second, when the policy makers differ significantly in their preferences, there will only be policy persistence by the left-wing governments: They will be more conservative following a right-wing predecessor. However, a conservative government succeeding a left-wing administration will choose more right-wing policies to attempt to "correct" the state of the economy it inherits from its predecessor.

### 1.3.4 More Than Two Periods

Our model could be easily transformed into a more-than-two-period model, under certain conditions.

Let us assume that there are two industries in the economy, with overlapping generations of investments: In each period, one of the industries needs to invest from scratch. The other industry enters the period with a certain amount of capital installed. It is realistic to assume that the capital income in each industry is taxed at a different tax rate. Real-world governments will want to attract new investment projects with tax rebates, etc.

If we make the additional assumption of labor force being split and tied to a particular industry, our two-period model fully applies. We will observe persistence in the tax rate applied to the incumbent industry. There will be no persistence within our model as far as the tax rate levied on the entering industry is concerned.

When we allow labor to move between sectors, the model with forward-looking firms and government becomes rather complicated and certainly very difficult to solve. However, we can always consider, in the spirit of our basic model, a simple case of myopic firms and government and a left-wing government following either a right- or left-wing predecessor.

Since our firms are perfectly competitive and exhibit constant returns to scale, a labor market equilibrium ensures not only that wages are equal in the entrant and incumbent firm sector, but also that the capital-labor ratio is the same in the two sectors. Then, as long as the government does not want to shut down a sector entirely, which would certainly be quite undesirable, setting a tax rate in one of the sectors fully determines the tax rate in the other

sector. The tax rates need not be the same unless both sectors operate on the same margin - that is, unless both sectors are installing capital in a given period (one of the sectors is definitely installing capital in every period). However, a decrease in one of the taxes will lead to a decrease in the other one and vice versa.

Then a government that inherits an incumbent sector with a certain amount of installed capital will want to retain as much as possible of that capital by choosing the maximum possible tax rate in the that sector, which achieves this. Comparatively, if it succeeds a right-wing administration that induced firms to install more capital in the then-entrant sector, it will need to set a lower tax rate in the now-incumbent sector. This will, however, necessitate a lower tax in the entrant sector as well.

Therefore with free labor movement between the two sectors in an overlapping generations model of firms, we will have policy persistence in a longer-than-two period horizon setting.

### 1.3.5 Government Objective: Caring About the Future

In our model thus far, the first-period government certainly is forward-looking and realizes the impact of its current-period actions on the next period's variables. The current expectations of those variables then in turn feed into the first-period welfare and thus its objective function. However, the next period's welfare of its agents does not enter its objective function directly. Therefore we could amend its objective function to

$$\max_{\tau_1} \left[ w_1(\gamma) + \frac{1}{2} \tau_1 r_1 K_1 + \delta \left[ p_R \left( w_2^R(\gamma) + \frac{1}{2} \tau_2^R r_2^R K_2^R \right) + p_L \left( w_2^L(\gamma) + \frac{1}{2} \tau_2^L r_2^L K_2^L \right) \right] \right].$$

This would lead to similar, albeit algebraically significantly more complicated, interior solution to the one we obtained when we in effect assumed that  $\delta$ , the first-period government's discount factor, was zero. More importantly, however, it would be difficult to fully characterize the solution due to the additional complexity caused by "manipulation" of the successor.

Namely, a first-period left-wing government, when  $\delta = 0$ , might want, in our above solution, to induce its right-wing successor to set a tax rate such that  $K_2^R > K_1^L$ , even though an interior solution of the problem would lead to  $K_2^R = K_1^L$ . It does so in order to take advantage of the additional capital installation by firms that expect with positive probability to install even more

capital in the second period. However, when  $\delta > 0$ , the opposite may occur: If the government cares sufficiently about the welfare of its agents in the next period and the interior solution leads to  $K_2^R > K_1^L$  and therefore  $\tau_2^R = 0$ , it may choose a lower tax rate in the first period so as to induce  $K_2^R = K_1^L$ .

Therefore for some values of  $R$  and depending on the other parameters of the model, we may see the first-period administrator "manipulating" its successor in either direction. However, our basic result of policy persistence for high enough values of  $R$  remains. For  $R > \tilde{R}(\delta)$ , where  $\tilde{R}(\delta)$  is  $\tilde{R}$  modified for  $\delta > 0$ , we have an interior solution such that  $K_2^R = K_1^L$ , therefore a first-period left can only try to affect its right-wing successor in the same way that we considered when  $\delta = 0$ . Then we will have a threshold value  $\tilde{R}_2(\delta)$ , such that for all  $R > \tilde{R}_2(\delta)$  there will be policy persistence.

## 1.4 Conclusion

In this paper we have explored a simple mechanism for fiscal policy persistence. Our environment includes a welfare-maximizing representative policy maker, as opposed to politicians concerned with holding power or bribes in some previous contributions. The model relies on the presence of installation costs of capital. These costs create a wedge in the marginal cost of capital for the firms between the first and second period. In the first period, firms must install capital and therefore pay the required net return to rent the capital from the world economy as well as pay for its installation. In the second period, provided they are not installing additional capital, firms only need to pay the required net return and thus face lower marginal cost of capital.

A government can take advantage of this: It can induce firms to maintain the same level of capital they installed in the first period and yet tax them more. However, say, a left-wing government inherits more capital from a conservative predecessor, who prefers lower taxes that attract higher levels of investment, than from a preceding left-wing government. Then to prevent the capital already installed from leaving the country, it will need to set lower tax rates when the amount of capital is higher - i.e. if the preceding government was conservative and had set lower tax rates in the first place. Thus we have policy persistence.

This persistence arises in an environment when the government does not want to directly alter the expected behavior of the representative of the population in the next period. The policies of the preceding government affect the current administration's actions even when the former did not strive for such an impact. This is in contrast to the work of, for instance, Persson and Svensson, where the concern about future periods distorts the decisions of the current policy maker.

The mechanism in our model functions under several alternative assumptions and also under the more complete specification where firms and government are forward-looking. We also show that not only will there be a two-way policy persistence when government types that alternate in power are similar enough: a left-wing administrator will render the subsequent conservative government more left-wing and vice versa. Interestingly, in our setup, when the types of government are sufficiently different, the influence a government happens to exert on its successor depends on the type of government: There will be policy persistence when a leftist government follows a conservative one. However, once a right-wing administration different enough from a left-wing policy maker succeeds it, it will implement even more right-wing policies than it otherwise would and we observe policy reversal.

Our results can also aid in explaining why otherwise very similar countries, run by governments from the same part of the political spectrum, may exhibit rather different policies. In our model, for instance, there will always be the same amount of capital present in the economy under a second-period government that is sufficiently right-wing. The administration either inherits it from a conservative predecessor or it induces firms to reach the same level of installed capital if it comes to power after a leftist government. However, the fiscal policy in either case is going to be rather different: The tax rate is going to be lower in the latter case.

There are several directions, in which this work could be extended. A fuller analysis of a model with more than two periods and forward-looking firms and government, while likely very complicated, could bring richer results. With a longer horizon, policies of a current government would be influenced by the entire path of policy makers in previous periods. Beyond simple policy persistence, one could study how different government actions can depend on the composition and order of the set of preceding administrations.

Also, the election process could be partly endogenized, so that the probability of being

elected would depend on the state of the economy. For instance, with a large amount of capital already present in the economy, voters might be more likely to elect a left-wing government that would redistribute more of the income generated by that capital.

Finally, our model has been built for a small country that takes the world required net return on capital as given. A model for a large country or a world composed of such countries would at least partly endogenize that return. The government of a large country would realize that attracting capital by lowering its own tax rate would also increase the world equilibrium net return and thus bring less capital into the country than otherwise. Therefore this alternative assumption could further reduce the policy differences between the various types of policy makers.

## 1.5 Appendix

### Manipulating a Right-Wing Successor by a Left-Wing Government

Here, we show that there is a value  $\tilde{R}_2, \tilde{R}_2 \in (\tilde{R}, 1)$ , such that a left-wing government when  $R \in (\tilde{R}, \tilde{R}_2]$  will deliberately choose a higher tax rate in order to cause its potential right-wing successor to induce a capital increase in the economy. The government will choose the smallest possible tax rate  $\tau_1^{man}$ , where  $\tau_1^{man} > \tau^*(\gamma_L)$ , that is high enough such that that happens.

The first-period left-wing policy maker's welfare, when additional capital is expected to be installed under a second-period right-wing government, peaks at a lower tax rate than when it is not. Simply compare

$$\begin{aligned} \tau_1^L(K_2^R > K_1^L) &= \frac{(1 - \gamma_L)[r^* + c(1 - \beta p_R)]}{r^*/(1 - \alpha) + (1 - \gamma_L)c(1 - \beta p_R)} < \\ &< \frac{(1 - \alpha)(1 - \gamma_L)(r^* + c)}{r^* + (1 - \alpha)c(1 - \gamma_L)} = \tau_1^L(K_2^R = K_1^L), \end{aligned}$$

which leads to

$$1 > (1 - \alpha)(1 - \gamma_L).$$

Therefore at  $R > \tilde{R}$ , welfare when  $K_2^R > K_1^L$  for the first-period administrator, who is considering increasing his tax above  $\tau_1^L(K_2^R = K_1^L)$  precisely in order to achieve  $K_2^R > K_1^L$ , is strictly

decreasing. He will thus consider setting the lowest possible tax rate, which we will call  $\tau_1^{man}$ , that indeed achieves  $K_2^R > K_1^L$ , and subsequently compare his welfare when choosing such a tax rate with welfare achieved when he selects  $\tau_1^L (K_2^R = K_1^L)$ .

This boundary tax rate that is just high enough to cause a second-period right wing government, when  $R > \tilde{R}$ , to induce an increase in the amount of capital installed, is implicitly defined by

$$welfare^R[K_2^R > K_1^L] = welfare^R[K_2^R = K_1^L], \text{ i.e.}$$

$$(1 - \alpha) \left( \frac{\alpha}{r^* + c} \right)^{\frac{\alpha}{1-\alpha}} \gamma_R = (1 - \alpha) \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c} \right)^{\frac{\alpha}{1-\alpha}} \gamma_R + \left( 1 - \frac{r^*}{\frac{r^*}{1-\tau_1^{man}} + c} \right) \left( \frac{r^*}{1-\tau_1^{man}} + c \right) \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c} \right)^{\frac{1}{1-\alpha}} \quad (1.10)$$

Then the first-period left-wing policy maker indeed chooses  $\tau_1^{man}$  when

$$welfare_1^L[\tau_1^L = \tau_1^{man}, K_2^R > K_1^L] > welfare_1^L[\tau_1^L = \tau^*(\gamma_L), K_2^R = K_1^L], \quad (1.11)$$

which can be re-written as

$$(1 - \alpha) \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{\alpha}{1-\alpha}} \gamma_L + \frac{\tau_1^{man} r^*}{1 - \tau_1^{man}} \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{1}{1-\alpha}} > > (1 - \alpha) \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{\alpha}{1-\alpha}} \gamma_L + \frac{\tau^*(\gamma_L) r^*}{1 - \tau^*(\gamma_L)} \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}}.$$

Note that if the above holds, then a second-period left-wing administration will certainly choose a tax rate than guarantees  $K_2^L = K_1^L$ : The second-period government will choose  $K_2^L = K_1^L$  as long as

$$welfare_2^L[\tau_2^L > \tau_1^{man}, K_2^L = K_1^L] > welfare_2^L[\tau_2^L = \tau^*(\gamma_L), K_2^L > K_1^L].$$

Since

$$welfare_2^L[\tau_2^L > \tau_1^{man}, K_2^L = K_1^L] > welfare_1^L[\tau_1^L = \tau_1^{man}, K_2^R > K_1^L] \text{ and}$$

$$welfare_2^L[\tau_2^L = \tau^*(\gamma_L), K_2^L > K_1^L] = welfare_1^L[\tau_1^L = \tau^*(\gamma_L), K_2^R = K_1^L],$$

then (1.11) implies that  $K_2^L = K_1^L$ . This is important so that our assumption that a second-period left-wing government does not induce more capital installation than its left-wing predecessor holds.

We proceed by realizing that at  $\tilde{R}_2$ , (1.11) holds with equality. We want to show that there is a unique such  $\tilde{R}_2 > \tilde{R}$ . For that it is sufficient to show that the difference between the two welfare values in (1.11) is decreasing at  $\tilde{R}_2$ : We know that the difference is positive at  $\tilde{R}$ , therefore it will need to decrease all the way to zero at  $\tilde{R}_2$ . The only way there could be more than a single value of  $R$  such that the difference is zero would be if for the higher values of such  $R$  the difference increased up to a point of equality between the welfare values in (1.11) or if it stayed at zero level for an interval rather than at a single value of  $R$ . This, of course, cannot occur if we show that at any point where the welfares in (1.11) equal, the derivative of the welfare difference is negative.

Factoring out  $\gamma_R$ , we can write the welfare difference as

$$\begin{aligned} \Delta W &= (1/\gamma_R) (welfare^L[\tau_1^L = \tau_1^{man}, K_2^R > K_1^L] - welfare^L[\tau_1^L = \tau^*(\gamma_L), K_2^R = K_1^L]) = \\ &= (1 - \alpha) \left( \frac{\alpha}{\frac{r^*}{1 - \tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{\alpha}{1 - \alpha}} R + \frac{\tau_1^{man} r^*}{2(1 - \tau_1^{man})} \left( \frac{\alpha}{\frac{r^*}{1 - \tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{1}{1 - \alpha}} (1 + R) - \\ &\quad - (1 - \alpha) \left( \frac{\alpha}{\frac{r^*}{1 - \tau^*(\gamma_L)} + c} \right)^{\frac{\alpha}{1 - \alpha}} R - \frac{\tau^*(\gamma_L) r^*}{2(1 - \tau^*(\gamma_L))} \left( \frac{\alpha}{\frac{r^*}{1 - \tau^*(\gamma_L)} + c} \right)^{\frac{1}{1 - \alpha}} (1 + R). \end{aligned}$$

Then

$$\begin{aligned} \frac{\partial \Delta W}{\partial R} &= \left[ (1 - \alpha) + \frac{\tau_1^{man} r^*}{2(1 - \tau_1^{man})} \right] \left( \frac{\alpha}{\frac{r^*}{1 - \tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{\alpha}{1 - \alpha}} - \left[ (1 - \alpha) + \frac{\tau^*(\gamma_L) r^*}{2(1 - \tau^*(\gamma_L))} \right] \cdot \\ &\quad \cdot \left( \frac{\alpha}{\frac{r^*}{1 - \tau^*(\gamma_L)} + c} \right)^{\frac{\alpha}{1 - \alpha}} - \left[ \frac{r^*}{1 - \tau_1^{man}} + c(1 - \beta p_R) \right]^{\frac{1}{\alpha - 1} - 1} \frac{\alpha^{\frac{1}{1 - \alpha}} r^*}{(1 - \tau_1^{man})^2} \cdot \\ &\quad \cdot \left[ \frac{c(1 - \beta p_R)}{2} + \frac{r^* [(1 - \alpha) + \tau_1^{man}]}{2(1 - \tau_1^{man})(1 - \alpha)} \right] \frac{\partial \tau_1^{man}}{\partial R} + \left[ \frac{r^*}{1 - \tau^*(\gamma_L)} + c \right]^{\frac{1}{\alpha - 1} - 1} \cdot \\ &\quad \cdot \frac{\alpha^{\frac{1}{1 - \alpha}} r^*}{(1 - \tau^*(\gamma_L))^2} \left[ \frac{c}{2} + \frac{r^* [(1 - \alpha) + \tau^*(\gamma_L)]}{2(1 - \tau^*(\gamma_L))(1 - \alpha)} \right] \frac{\partial \tau^*(\gamma_L)}{\partial R} \end{aligned}$$

We also know that

$$\frac{\partial \tau^*(\gamma_L)}{\partial R} = \frac{\partial \left[ \frac{(1-\alpha)(r^*+c)(1-R)}{r^*(1+R)+(1-\alpha)c(1-R)} \right]}{\partial R} = -\frac{2r^*(1-\alpha)(r^*+c)}{[r^*(1+R)+(1-\alpha)c(1-R)]^2} < 0.$$

Then totally differentiating (1.10), we obtain, after some algebra,

$$\frac{\partial \tau_1^{man}}{\partial R} = \frac{(1-\alpha) [\tau_1^{man} r^* (1 - \tau_1^{man}) + c(1 - \tau_1^{man})^2]}{r^* [2 + \alpha(R+1) \frac{\tau_1^{man} r^* (1 - \tau_1^{man}) + c}{r^* (1 - \tau_1^{man}) + c}]} > 0.$$

This in turn means that to show  $\frac{\partial \Delta W}{\partial R} < 0$ , it is sufficient to show that

$$0 \geq (1-\alpha) \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1-\beta p_R)} \right)^{\frac{1}{1-\alpha}} + \frac{\tau_1^{man} r^*}{2(1-\tau_1^{man})} \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1-\beta p_R)} \right)^{\frac{1}{1-\alpha}} - \\ -(1-\alpha) \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}} - \frac{\tau^*(\gamma_L) r^*}{2(1-\tau^*(\gamma_L))} \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}}.$$

Let us proceed by contradiction and therefore assume that the above expression is in fact positive. Using the fact that we are evaluating  $\frac{\partial \Delta W}{\partial R}$  at  $\tilde{R}_2$ , where (1.11) holds with equality, we obtain

$$\frac{\tau_1^{man} r^*}{2(1-\tau_1^{man})} \left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1-\beta p_R)} \right)^{\frac{1}{1-\alpha}} - \frac{\tau^*(\gamma_L) r^*}{2(1-\tau^*(\gamma_L))} \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}} \leq 0. \quad (1.12)$$

This states that the former tax revenue is lower than the latter. But since the welfares for the  $\gamma_L$  agent are equal at both points, this necessarily implies that at  $\tau_1^{man}$  and  $K_2^R > K_1^L$ , the agents' wages and therefore installed capital is higher than at  $\tau^*(\gamma_L)$  and  $K_2^R = K_1^L$ , i.e.

$$K_1^L[\tau_1^{man}, K_2^R > K_1^L] > K_1^L[\tau^*(\gamma_L), K_2^R = K_1^L]$$

and therefore

$$\left( \frac{\alpha}{\frac{r^*}{1-\tau_1^{man}} + c(1-\beta p_R)} \right)^{\frac{1}{1-\alpha}} > \left( \frac{\alpha}{\frac{r^*}{1-\tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}}.$$

We also know that  $\tau_1^{man} > \tau^*(\gamma_L)$  and therefore

$$\frac{\tau_1^{man} r^*}{2(1 - \tau_1^{man})} > \frac{\tau^*(\gamma_L) r^*}{2(1 - \tau^*(\gamma_L))}.$$

This leads to the conclusion that

$$\frac{\tau_1^{man} r^*}{2(1 - \tau_1^{man})} \left( \frac{\alpha}{\frac{r^*}{1 - \tau_1^{man}} + c(1 - \beta p_R)} \right)^{\frac{1}{1-\alpha}} - \frac{\tau^*(\gamma_L) r^*}{2(1 - \tau^*(\gamma_L))} \left( \frac{\alpha}{\frac{r^*}{1 - \tau^*(\gamma_L)} + c} \right)^{\frac{1}{1-\alpha}} > 0,$$

which is a contradiction with (1.12).

We have thus shown that  $\frac{\partial \Delta W}{\partial R}$  at  $\tilde{R}_2$  is negative and therefore, by the above argument, first-period's left-wing governments choose  $\tau_1^L = \tau_1^{man}$  for  $R \in (\tilde{R}, \tilde{R}_2]$ .

## Chapter 2

# The Right-Wing Power of Small Countries

### 2.1 Introduction

The enlargement of the European Union in May 2004 was probably the single most significant change in the membership of the European bloc in its history. Sometimes dubbed big bang, 10 new countries joined the elite group. They represent formidable competition for the original EU-15. Apart from their low labor costs, most of the new countries have implemented rather low corporate tax rates, in order to attract even more investment in their still transforming economies. The old EU members are slowly responding to the new challenge in the tax competition within the EU. The citizens of the old members might themselves be recognizing the changed situation and are electing governments more capable of responding to it - for example, the recent slight shift to the right in Germany and France.

In this paper, we find a framework that helps explain the right-ward shift in some EU governments in response to increased competition from the new members of the Union. The economies of the EU-10 are on average, and as opposed to the EU-15, poorer and smaller. Our mechanism is based on the second difference.

When we consider the current corporate tax rates in all of the Union's 27 countries, it is clear that they are positively related to their size (see Figures 1 and 2). Old and new small members have rates as low as 10%, whereas the 5 largest EU members all tax capital at over 30%.

Huizinga and Nicodeme (2005) provide rigorous estimates that support our casual observation when they find a positive impact of a country's GDP on its corporate tax rate.

We build a model of tax competition between countries of different sizes that incorporates the economic and political equilibria. Countries compete for internationally mobile capital. Since they are equally productive, the capital responds and moves according to the countries' various capital tax rates. The fiscal policies are decided by representative governments from heterogeneous populations, which are elected by the median agents in each country.

First, we show in our environment populated by many large and small countries that, when the countries are run by the same type of government, the small countries will set a lower tax rate than large ones. This is because each of the small countries cannot change the world equilibrium much (or at all) and therefore perceives a higher elasticity of capital with respect to its own tax rate.

The second reason in our model why large countries set a higher tax rate than small ones is that their governments are to the left of those elected by their smaller competitors. The voters in large countries, just as their governments, understand that their country has an impact on the world equilibrium. Therefore the mechanism proposed by Persson and Tabellini (1992) applies here: Before elections and the subsequent non-cooperative tax setting take place, the electorate in the large countries understands that the other countries will respond positively to its own government's actions. However, after the elections, during the Nash game where countries set the tax rates simultaneously, governments take the other countries' fiscal policies as given. Given this, the agents will always prefer a higher tax rate ex-ante, before the game begins, than ex-post, when the game is played. To have this tax rate implemented, they will vote for a government with preference for higher taxes than they themselves. Since it is the median voter who ultimately chooses the representation, the government will be to the left of the median of the population.

Once again, though, small countries affect the world equilibrium little or not at all. Therefore in our model of heterogeneous countries a change in the policy of a small country does not cause a shift in other countries' tax rates. But then it follows that the median voter in a small country chooses himself or an agent of the same type to become the policy maker and set the tax rate. The government in small countries is thus to the right of that in large ones and we have another

reason for small countries setting a lower tax on capital in equilibrium.

Our model then produces its main result: When faced with relatively more small countries in a tax competition, the large country governments shift to the right. This occurs for two reasons. First, as we have already mentioned, smaller countries are tougher competitors for scarce capital. They set lower tax rates because they perceive higher elasticity of capital with respect to their fiscal policy and because their governments are to the right of those in large countries. That, in turn, decreases the equilibrium tax rate set by large countries as they are forced to respond to the challenge posed by a larger fraction of small competitors. Smaller absolute tax rates then diminish the difference between desired ex-ante and ex-post tax rates. The need to elect a left-wing government decreases. Second, smaller countries respond less (or, in our basic setup, not at all) to changes in the policy of large countries. This is again due to the higher elasticity of capital they face when they alter their own tax rate. A more significant proportion of the competitors a large country faces reacts less to its own actions when that large country is competing with relatively more small countries. That, once again, reduces the need for the median agent to elect a representative that is poorer than himself.

Small countries thus have what we call a "right-wing power". Their mere presence in a tax competition environment causes a right-ward move in the type of policy makers elected in large countries. We also show that the same result holds when new small countries are simply added to a group of countries competing for capital. However, the same is not always true for additional large countries joining the tax-setting game. Here, two effects combine. More large countries represent an increased level of competition. Even if they are not as challenging opponents as small countries, a higher number of large competitors in a setup without small countries would still induce lower equilibrium tax rates and a right-ward move in the type of administration. But these are the countries that respond more significantly to changes in others' fiscal policies. Therefore bringing up their numbers when small countries are present can result in a left-ward shift in large countries' governments and an increase in tax rates, because then proportionately more countries will be responsive in the Nash competition. We demonstrate that the second effect dominates when a minimum mass of small countries is also competing for capital. Thus larger countries do not have the same right-wing power their smaller counterparts possess.

This paper employs the framework of tax competition and coordination in its attempt to explain the political impact of country size heterogeneity in an environment with freely mobile capital. The vast body of work on tax competition is well summarized in Krogstrup (2003) and specifically as it pertains to the European Union in Krogstrup (2002) and Nicodeme (2006). Relative to the size of the literature, few authors have considered the interaction between countries of different size. Under different assumptions - such as quadratic production function - and mostly only for two-country interactions, Kanbur and Keen (1993), Wilson (1991), Bucovetsky (1991) and Peralta and van Ypersele (2005) have all proved that smaller countries set lower tax rates than their larger counterparts. In their models, small countries perceive a higher elasticity of capital with respect to their own tax rate.

However, the literature does not consider the interaction between country size heterogeneity and the political equilibrium. It is this interplay that generates the second reason, for which in our environment large countries set a higher tax rate than small ones - their governments are to the left of those elected by their smaller competitors. Equally importantly, it results in the right-wing power of small countries.

The paper proceeds as follows: The next section introduces the model. Section 3 considers ex-post and ex-ante tax competition, where we show that smaller countries choose lower tax rates and larger countries elect more left-wing governments. In section 4 we show our main result that a change in the composition of countries, with which a large country is competing, will determine how left-wing its government will be. We also evaluate a minimum tax proposal by some of EU's policy makers in the framework of our model. Section 5 provides computational solutions for an expanded version of our model. The last section concludes.

## 2.2 The Model

### 2.2.1 The Basics

Our model has one period. The world consists of large and small countries. We have  $n$  large countries (L), each of which has a mass 1 of agents. We also have a mass  $s$  of small countries (S), each of which has an infinitesimally small mass of agents.

There are perfectly competitive firms with access to a CRS technology in all the countries.

Firms produce a single good employing labor and capital, for which they compete in the factor markets and therefore pay them their marginal products.

Throughout our model, we shall assume the Cobb-Douglas production function, so that production is  $F(K, L) = K^\alpha L^{1-\alpha}$ , where  $K$  is the capital and  $L$  is the total effective labor employed. We can re-write the production function as  $f(k) = k^\alpha$ , where  $k$  is capital per unit of effective labor.

Agents are risk-neutral. They are immobile between countries. They each have one unit of labor, which they supply inelastically to the firms. Labor market clears, therefore all agents are employed. Some are also endowed with capital, which they rent to firms. Agents' endowment is the only source of capital in the world. Since our model only has one period, no decisions regarding capital accumulation are made.

Agents differ in their labor productivity  $\gamma$ , and capital endowment  $\theta$ . The mean of each distribution is one. Agents' productivity  $\gamma$  turns the one unit of labor they supply to the market into  $\gamma$  units of effective labor supply. We shall assume that for the median agent  $\gamma < 1$ . We will make the simplifying assumption that while agents' productivity  $\gamma$  is positive and continuously distributed across all agents, only agents with above-average productivity are also endowed with an amount of capital  $\theta > 0$ . Moreover,  $\theta$  is strictly increasing in  $\gamma$  for agents with values of  $\gamma$  above the mean. Therefore the above will imply that the median agent or anyone less productive will not be endowed with any capital<sup>1</sup>. However, since the mean capital endowment of an agent in every country is one, per capita supply of capital will also be one.

Thus the total effective labor employed by firms in any country is equal to the mass of agents in that country. Also, the total amount of capital in the world is simply  $K^T = n + s$ , in other words, equal to the size of the world population. Throughout the model we will assume that  $n \geq 2$  and therefore  $K^T \geq 2$ . This is because, as we will see later, interesting dynamics in our setup only occur when at least two large countries compete for capital.

Capital is perfectly mobile between the countries and therefore responds to the various tax incentives it faces. Ultimately, thanks to arbitrage, it earns the same endogenously determined net rate of return in all countries.

---

<sup>1</sup>Only the more productive agents with a higher labor income own capital. In a richer environment, these would be the agents able to save more and invest their savings, thus resulting in a capital endowment.

The government in each country only taxes capital. It collects tax revenue and then re-distributes it lump-sum to all the agents in the country, such that they all receive the same payment from the government. Since agents with different productivity earn different labor income, and therefore are affected differently by variations in capital stock caused by the tax rate on capital, the administration can achieve redistribution even under the assumption of uniform lump-sum transfers. We assume that the tax rate on capital is non-negative<sup>2</sup>.

Each agent's welfare is then the sum of his labor and net capital income and the lump-sum transfers from the government. Agent's wage will be  $w = \gamma(1 - \alpha)k^\alpha$  and his capital income  $\theta(1 - \tau)\alpha k^{\alpha-1}$ , where  $\tau$  is the tax rate imposed on capital returns. Then an agent's welfare will be

$$W(\tau, \gamma, \theta) = \gamma(1 - \alpha)k^\alpha + \theta(1 - \tau)\alpha k^{\alpha-1} + \tau\alpha k^\alpha.$$

In each country, the agents elect from amongst themselves a government. Whichever of the candidates receives the majority of the votes wins the election. The government then makes policy decisions regarding the capital tax rate  $\tau$ . It chooses the tax rate, which is optimal for the agent elected, that is, it maximizes the welfare of that particular agent with certain parameters  $\gamma$  and  $\theta$ .

Finally, the timing in our model is as follows: First, voters in each country simultaneously elect their respective governments. Then, once elected, the various governments simultaneously announce and commit to a capital tax policy. Finally, capital moves between countries until net return on capital is equalized and then production takes place.

We thus look for a two-part equilibrium: An economic equilibrium where governments play a non-cooperative Nash game when setting their respective tax rates. In the game, the policy makers take the other countries' policy decisions as given. The second part of the solution is a political equilibrium that requires that each government be elected and therefore preferred by a majority of voters in their country. When deciding, voters take foreign countries' election outcomes as given: Elections in all countries occur simultaneously and thus voters cannot affect another country's choice of policy maker. However, they realize that the foreign policy makers will set their policies in response to those of their own government. Therefore they certainly

---

<sup>2</sup>In our model, subsidies to capital would be financed by negative transfers to agents. Therefore a negative tax rate on capital would imply imposing a lump-sum tax on the agents.

do not take the foreign policies as given. Then the voters and policy makers maximize the same welfare function, but subject to different constraints. This, as we will see, will lead to a difference between the optimal policies implemented by the government at the time of the Nash game and those desired by the voters before the elections. Voters then strategically delegate at the time of the elections to achieve their optimal policy in the Nash game.

## Politics

Let us consider what parameters  $\gamma$  and  $\theta$  enter the government's maximization problem, that is, which agent is elected to the government.

We use the single-crossing property of Gans and Smart (1996) to show that the majority of voters always elects the agent preferred by the median voter. In particular, we show that if an agent prefers a higher tax rate to a lower one, than a "poorer" agent will certainly also prefer the higher tax rate, and vice versa (see Appendix 1). Then the agents' preferences satisfy the single crossing property in each country and we can use the result of Gans and Smart that a Condorcet winner exists and he represents the optimum for the median voter.

To ensure that the Condorcet winner is elected, we will make one additional assumption regarding the political process: Let us assume that the election is never a single pairwise competition of candidates equally preferred by the median voter. In such a case he would be indifferent between the two candidates and they would both stand the same chance of winning. We would have an equilibrium with a random election outcome.

Rather, we have pairwise competitions between candidates, one of whom is always preferred by the median voter. Then the Condorcet winner beats any opponent. The candidate preferred by the median agent wins the election with certainty.

### 2.2.2 Small Country

An infinitesimally small country on its own has no effect on the world economy. While a mass of such small countries may together affect the world equilibrium, any individual small country will necessarily take as exogenous the world net return on capital, which is equalized between the countries and actually determined endogenously within the model.

Due to international arbitrage, the net return on capital present in the economy must

equal the world net return on capital,  $r^*$ , taken as given by the small country. Therefore  $(1 - \tau_S)\alpha k_S^{\alpha-1} = r^*$ , and thus  $k_S = \left(\frac{(1-\tau_S)\alpha}{r^*}\right)^{\frac{1}{1-\alpha}}$ , where  $\tau_S$  stands for the tax rate chosen in a small country. We will only consider symmetric equilibria, where identical countries also behave identically and therefore all small countries select the same tax rate.

A small country government thus maximizes the total labor and capital income of and transfers received by the agent it represents; i.e.

$$\max_{\tau_S} \gamma(1 - \alpha) \left(\frac{(1 - \tau_S)\alpha}{r^*}\right)^{\frac{\alpha}{1-\alpha}} + \tau_S \alpha \left(\frac{(1 - \tau_S)\alpha}{r^*}\right)^{\frac{\alpha}{1-\alpha}}.$$

Note that capital income in the case of the small country does not depend on the government's policy, because net capital return is perceived to be constant.

The maximization problem yields  $\tau_S = (1 - \alpha)(1 - \gamma)$ . Since we have assumed taxes to be non-negative, the government will choose the corner solution of zero taxation whenever it cares about an agent with an above-average labor productivity.

### 2.2.3 Large Country

A large country realizes that its choice of capital tax rate will have an impact on the world net capital return. For example, if it decreases its capital tax rate, capital will flow into the country, but as that also implies that, given the fixed total amount of capital in the world, there will be capital outflow elsewhere, the net return on capital will also rise. Therefore it will perceive a lower elasticity of capital with respect to its tax rate than a small country, which believes the only effect of its actions to be the flow of capital (and no change in the net return on capital).

Let us for now assume that the median voter in a large country will prefer either himself or someone poorer than himself, i.e. an agent with a lower value of  $\gamma$ , to be the policy maker. This is to simplify our subsequent optimization problems by setting  $\theta = 0$ . We will later see that the median voter will indeed prefer to elect an agent with the same or smaller labor productivity.

Because the net return on capital must equal between countries, if our large country's capital stock is  $k'_L$  and tax rate  $\tau'_L$  (we will denote with primes the variables for that particular large country, for which we are performing a given calculation at the time), every other large country's capital stock can be obtained from  $\alpha(1 - \tau'_L)k_L'^{\alpha-1} = \alpha(1 - \tau_L)k_L^{\alpha-1}$ . Thus we have

$k_L = k'_L \left( \frac{1-\tau_L}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}}$ . Similarly, for small countries we already know that  $k_S = \left( \frac{(1-\tau_S)\alpha}{r^*} \right)^{\frac{1}{1-\alpha}}$  and we can again express the net world return as  $r^* = \alpha(1 - \tau'_L)k'_L{}^{\alpha-1}$ , therefore obtaining  $k_S = k'_L \left( \frac{1-\tau_S}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}}$ .

Then, since the total world capital is  $K^T = n + s$ , we will have

$$k = \frac{K^T}{A'}, \text{ where } A' \equiv 1 + (n-1) \left( \frac{1-\tau_L}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}} + s \left( \frac{1-\tau_S}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}}. \quad (2.1)$$

Then the government maximization problem becomes

$$\max_{\tau'_L} [\gamma(1-\alpha) + \tau'_L\alpha] \left( \frac{K^T}{A'} \right)^\alpha,$$

which yields the first order condition

$$-\frac{A'-1}{(1-\tau'_L)(1-\alpha)} [\gamma(1-\alpha) + \tau'_L\alpha] + A' = 0. \quad (2.2)$$

The optimal tax rate is the solution to (2.2).

Note also that there is a one-to-one correspondence between the type  $\gamma$  and the tax rate  $\tau'_L$  chosen by that type, since, as we show in the Appendix, preferences of agents in large countries are quasi-concave and have a unique maximum on the interval  $(0, 1)$ . When we hold  $\tau'_L$  constant, differentiating the LHS of (2.2) we have

$$\frac{\partial W}{\partial \tau'_L \partial \gamma} = -\frac{A'-1}{(1-\tau'_L)} < 0.$$

The desired policy  $\tau'_L$  is then a decreasing function of the type  $\gamma$ .

## 2.3 Tax Competition

### 2.3.1 Nash Competition

We can now show that for two governments representing their median agent, a government in a large country will set a higher tax rate than a government in a small country. This is simply due to the fact that the small country government perceives a higher elasticity of capital with

respect to the tax rate it selects than the same type of government in a large country.

Let us assume, by contradiction, that the large countries select the same tax rate as the small countries and therefore in equilibrium all states choose the tax rate  $\tau = (1 - \gamma)(1 - \alpha)$ . Then let us show that the expression in (2.2), i.e. the derivative of welfare with respect to the tax rate, is positive, and thus optimal tax rate in large countries is higher than that in small ones. The left-hand side of (2.2) becomes

$$-\frac{K^T - 1}{[1 - (1 - \gamma)(1 - \alpha)](1 - \alpha)} [\gamma(1 - \alpha) + (1 - \gamma)(1 - \alpha)\alpha] + K^T > 0,$$

which simplifies to  $1 > 0$ , which is always true.

Since the preferences of the median agent are quasi-concave (see Appendix 2), the tax rate in the large countries is higher than in small countries if they are all run by median agents.

### 2.3.2 Elections

We are now ready to show that a large country elects a government to the left of the median agent, because the voters at the time of the elections and the government, when it sets its fiscal policy, perceive the international tax competition differently. Whereas ex-post the government takes other countries' tax rates as given when choosing its tax policy, ex-ante the voters realize that other countries' tax rates will respond to their own country's policy changes. Since the tax policies of the various countries are strategic complements, each agent will perceive a lower elasticity of capital with respect to the tax rate before the election than if he is in the government, deciding about the optimal fiscal policy.

We have already shown that it is the median agent who chooses the policy maker. However, in this case, he will not find it optimal to elect himself or an agent with the same preferences, since he realizes that he himself would not actually implement his desired ex-ante tax policy. Since, as we have seen above, the tax rate that is implemented in the Nash equilibrium is decreasing in the type of policy maker, the median agent will elect someone poorer than himself if he desires a higher tax rate ex-ante than he himself would choose in the Nash competition, and vice versa.

Therefore we can think of the median agent wanting a certain ideal tax rate at the time

of the election as him choosing the optimal policy maker to run the government. He therefore maximizes his welfare with respect to the parameter  $\gamma'_p$  of the policy maker, where  $\gamma$  is his own productivity. For now, we shall assume that he will elect a poorer agent, so that we can disregard election candidates with positive holdings of capital. We then have

$$\frac{\partial W}{\partial \gamma'_p} = \frac{\partial W}{\partial \tau'_L} \cdot \frac{d\tau'_L}{d\gamma'_p} + \frac{\partial W}{\partial \tau_L} \cdot \frac{d\tau_L}{d\gamma'_p} = 0, \quad (2.3)$$

since the tax rate chosen by other large countries,  $\tau_L$ , will also depend on the type of the policy maker  $\gamma'_p$ . That policy maker is the competitor for the other large countries and therefore influences their own policy. Remember that small countries have a constant fiscal policy.

From (2.3) we can then obtain

$$\frac{\partial W}{\partial \tau'_L} + \frac{\partial W}{\partial \tau_L} \cdot \frac{d\tau_L}{d\tau'_L} = 0,$$

which is our ex-ante first-order condition. We can also interpret it as simply the derivative of the welfare function with respect to  $\tau'_L$ , where the median agent realizes ex-ante (as opposed to ex-post) that other countries' tax rates, i.e.  $\tau_L$ , will respond to changes in his own tax policy. The agent, once again, does not actually choose and implement the tax rate  $\tau'_L$ . He views the problem before the elections and therefore takes only the types of the foreign governments (rather than their policies) as given. Then he knows what his ideal tax rate is: The tax rate  $\tau'_L$  that maximizes his welfare when the other countries' fiscal policies are functions of  $\tau'_L$  that arise from the Nash game optimal responses. He then elects the agent, who will implement his desired  $\tau'_L$ .

The ex-ante first order condition of a large country's median agent becomes

$$\left( \frac{1 - \tau_L}{1 - \tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n-1)\partial\tau_L/\partial\tau'_L}{(1 - \tau_L)(1 - \alpha)} [\gamma(1 - \alpha) + \tau'_L\alpha] - \frac{A' - 1}{(1 - \tau'_L)(1 - \alpha)} [\gamma(1 - \alpha) + \tau'_L\alpha] + A' = 0. \quad (2.4)$$

Here,  $\partial\tau_L/\partial\tau'_L$  is, once again, the response of any other large country's policy to a change in the tax rate of that large country, for which (2.4) holds. Since the response occurs when the Nash policy game is played, we can obtain  $\partial\tau_L/\partial\tau'_L$  from (2.2), since that condition holds for any large country.

We can write (2.2) as

$$F \equiv - \left[ 1 + (n-2) \left( \frac{1-\tau_L}{1-\tau_L''} \right)^{\frac{1}{1-\alpha}} + \left( \frac{1-\tau_L'}{1-\tau_L''} \right)^{\frac{1}{1-\alpha}} + s \left( \frac{1-\tau_S}{1-\tau_L''} \right)^{\frac{1}{1-\alpha}} \right] \cdot [(\gamma - \Delta\gamma)(1-\alpha) + \tau_L''\alpha - (1-\tau_L'')(1-\alpha)] + (\gamma - \Delta\gamma)(1-\alpha) + \tau_L''\alpha, \quad (2.5)$$

where  $\gamma_p'' = \gamma - \Delta\gamma$ .

We have rewritten (2.2) for a particular large country that selects its fiscal policy  $\tau_L''$ , while separating the country with tax rate  $\tau_L'$  from the remaining large countries, which all choose  $\tau_L$ . This is because we want to know how much  $\tau_L''$  will respond to an exogenous change in tax rate  $\tau_L'$  of the large country, for which (2.4) is written. It is important to realize that all the remaining large countries in the world will also respond to a change in  $\tau_L'$ . Also, the country selecting  $\tau_L'$  realizes that  $\tau_L$  and  $\tau_L''$  are chosen by identical countries and therefore  $\tau_L = \tau_L''$  as well as response in  $\tau_L$  is identical to the response in  $\tau_L''$ , i.e.  $d\tau_L/d\tau_L' = d\tau_L''/d\tau_L'$ .

Then we have

$$0 = dF/d\tau_L' = \partial F/\partial\tau_L' + \partial F/\partial\tau_L \cdot d\tau_L/d\tau_L' + \partial F/\partial\tau_L'' \cdot d\tau_L''/d\tau_L',$$

which implies

$$d\tau_L/d\tau_L' = d\tau_L''/d\tau_L' = - \frac{\partial F/\partial\tau_L'}{\partial F/\partial\tau_L + \partial F/\partial\tau_L''}. \quad (2.6)$$

This leads to, when we incorporate the fact that  $\tau_L = \tau_L''$ ,

$$\frac{d\tau_L}{d\tau_L'} = \left( \frac{1-\tau_L'}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} \left[ (A+1-n) \left( \frac{1-\tau_L'}{1-\tau_L} \right) + \frac{(1-\tau_L')(A-\alpha)}{\left[ \gamma_p + \frac{\tau_L\alpha}{1-\alpha} - (1-\tau_L) \right]} \right]^{-1},$$

where

$$A = n - 1 + \left( \frac{1-\tau_L'}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} + s \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}}.$$

$A$  here is evaluated from the point of view of the country choosing  $\tau_L'$ . From (2.2) we also have

$$\gamma_p + \frac{\tau_L\alpha}{1-\alpha} = \frac{A(1-\tau_L)}{(A-1)},$$

which we can substitute into the former to obtain

$$d\tau_L/d\tau'_L = \left( \frac{1 - \tau'_L}{1 - \tau_L} \right)^{\frac{1}{1-\alpha}-1} [(A + 1 - \alpha)(A - 1) - (n - 2)]^{-1}. \quad (2.7)$$

We can see that, since  $A \geq n - 1$ , we will have  $d\tau_L/d\tau'_L > 0$ . The tax rates are strategic complements.

Then for any tax rate that satisfies (2.2), the expression in (2.4) will be positive, which, since preferences are quasi-concave, implies that the ex-ante optimal tax rate will be higher than the ex-post optimal tax rate. The median voter will therefore indeed want to elect an agent to the left of himself in order to achieve his ex-ante desired tax policy.

Thus large countries will have higher tax rates than their smaller competitors for two reasons: The economic reason implies that even if they have governments with identical preferences, the one in the large country will select a higher tax rate. Moreover, since individual small countries do not affect the world equilibrium and therefore do not induce any fiscal response by changing their own policies (i.e.  $d\tau_L/d\tau_S = 0$ ), the above mechanism does not apply to them and they elect median types to run the government. But then the large countries will have governments to the left of those in the small countries, and those governments will set even higher tax rates than the same type would.

## 2.4 Composition of Competitors and Political Shifts

### 2.4.1 Small Countries and the Shift to the Right

We now show that if a large country faces competition from a relatively larger number of small countries, its government will be more right-wing. We proceed by performing a comparative static exercise where we increase the total size of small countries (their total mass)  $s$  and decrease the number of large countries  $n$  so as to keep  $K^T = n + s$  constant. The size of the world is held constant so that we can determine the political impact of competing with relatively more small countries, as separate from the effect of simply competing with a higher number of countries. A competition between more countries alone, small or large, can cause a right-ward shift in the governments of large countries, therefore we want to demonstrate our

result net of this impact. We will be comparing symmetric equilibria, therefore, once again, all large countries' tax rates will be equal to each other.

Remember that  $\Delta\gamma$  is the difference in productivity between the median agent and the agent elected to the government in a large country, so that large country governments maximize for an agent with parameter  $\gamma_p = \gamma - \Delta\gamma$ . In equilibrium,  $\gamma_p$  will be the same for all large countries. We want to show that the equilibrium value of  $\Delta\gamma$  decreases as  $s$  increases.

We begin by showing that the equilibrium tax rate in large countries decreases when the mass of small countries increases (once again, the tax rate in small countries remains unchanged). First, we can rewrite (2.4) in a symmetric equilibrium as  $0 = G(\tau_L, s)$ , where

$$G(\tau_L, s) \equiv \frac{(n-1)[\gamma(1-\alpha) + \tau_L\alpha]}{(1-\tau_L)(1-\alpha)[(A+1-\alpha)(A-1) - (n-2)]} + A \left( 1 - \frac{\gamma(1-\alpha) + \tau_L\alpha}{(1-\tau_L)(1-\alpha)} \right) + \frac{\gamma(1-\alpha) + \tau_L\alpha}{(1-\tau_L)(1-\alpha)}. \quad (2.8)$$

$G(\tau_L, s)$  is simply the ex-ante derivative of the welfare function with respect to the country's own tax rate  $\tau'_L$ , evaluated at the point  $\tau'_L = \tau_L$ .

By implicit function theorem, we have

$$\frac{\partial \tau_L}{\partial s} = - \frac{\partial G(\tau_L, s)/\partial s}{\partial G(\tau_L, s)/\partial \tau_L}.$$

We show in Appendix 3 that  $\partial G(\tau_L, s)/\partial s < 0$  and  $\partial G(\tau_L, s)/\partial \tau_L < 0$ , therefore  $\partial \tau_L/\partial s < 0$ : The equilibrium large-country tax decreases when  $s$  increases.

When we write (2.2) for  $\gamma_p$ , we obtain

$$\Delta\gamma = \gamma + \frac{\tau_L\alpha}{1-\alpha} - \frac{A(1-\tau_L)}{(A-1)}. \quad (2.9)$$

Differentiating with respect to  $s$ , realizing that both  $\tau_L$  and  $A$  are functions of  $s$ , gives

$$\partial \Delta\gamma/\partial s = \frac{\alpha(\partial \tau_L/\partial s)}{1-\alpha} - \frac{[(\partial A/\partial s)(1-\tau_L) - A(\partial \tau_L/\partial s)](A-1) - (\partial A/\partial s)A(1-\tau_L)}{(A-1)^2}$$

Let us proceed by contradiction: We want to show that  $\partial \Delta\gamma/\partial s < 0$ , therefore we will

assume that  $\partial\Delta\gamma/\partial s > 0$ , i.e.

$$(\partial\tau_L/\partial s) \left( \frac{\alpha}{1-\alpha} + \frac{A}{A-1} \right) + (\partial A/\partial s) \frac{1-\tau_L}{A-1} > 0. \quad (2.10)$$

Since  $\partial\tau_L/\partial s < 0$ , (2.10) can only possibly hold if  $\partial A/\partial s > 0$  holds as well.

However, combining (2.2) and (2.4), we have

$$\Delta\gamma = \frac{(n-1)[\gamma(1-\alpha) + \tau_L\alpha]}{(A-1)(1-\alpha)[(A+1-\alpha)(A-1) - n + 2]}$$

Here,

$$\frac{\partial\Delta\gamma}{\partial s} = \frac{[\gamma(1-\alpha) + \tau_L\alpha] - (\partial\tau_L/\partial s)(n-1)\alpha}{(1-\alpha)(A-1)[(A+1-\alpha)(A-1) - n + 2]} - \frac{(n-1)[\gamma(1-\alpha) + \tau_L\alpha][A-1 + (\partial A/\partial s)[(3A+1-2\alpha)(A-1) - n + 2]]}{[(A-1)[(A+1-\alpha)(A-1) - n + 2]]^2} \quad (2.11)$$

If, as we concluded above,  $\partial A/\partial s > 0$ , then we necessarily arrive at  $\partial\Delta\gamma/\partial s < 0$ , which is a contradiction. Therefore we have shown that in fact  $\partial\Delta\gamma/\partial s < 0$ .

Therefore as  $s$  increases, we will not need as low a value of  $\gamma$  to satisfy (2.2) at the optimal tax rate from (2.4). Large countries will elect more right-wing governments (even though still to the left of the median and of the governments of the small countries) when they face proportionally more competition from small countries than otherwise.

## 2.4.2 More Competitors: Small Countries

While the above certainly is an interesting comparative static exercise, in practice we can rarely compare two instances of tax competition where in one of the cases we replace large countries with small countries. Rather, we can observe additional small countries joining the competition. We have already noted that increased competition alone can drive governments to the right. Here, that effect combines with the one we described in the previous section. Therefore our preceding result suggests that additional small countries in the tax competition will also shift the large country policy maker type to the right. We want to nevertheless verify formally that this change occurs when  $s$  increases and we keep  $n$  constant.

As in the previous section, we show in Appendix 3 that  $\partial G(\tau_L, s)/\partial s < 0$ , therefore

$\partial\tau_L/\partial s < 0$ . Therefore simply increasing the mass of small countries also decreases the equilibrium tax rate in the large countries.

We still need  $\partial A/\partial s > 0$  for (2.10) to hold. Since (2.11) becomes

$$\frac{\partial\Delta\gamma}{\partial s} = \frac{(\partial\tau_L/\partial s)(n-1)\alpha}{(1-\alpha)(A-1)[(A+1-\alpha)(A-1)-n+2]} - \frac{(n-1)[\gamma(1-\alpha)+\tau_L\alpha](\partial A/\partial s)[(3A+1-2\alpha)(A-1)-n+2]}{[(A-1)[(A+1-\alpha)(A-1)-n+2]]^2}$$

there is once again a contradiction; because the above implies that  $\partial\Delta\gamma/\partial s < 0$  when  $\partial A/\partial s > 0$ .

Therefore our result from the previous section holds: Governments in large countries shift to the right when additional small countries join the tax competition.

### 2.4.3 More Competitors: Large Countries

A similar condition does not hold in general when we increase the number of large countries competing for capital. We already know that when we simply increase the proportion of large countries in the tax competition, while keeping the size of the world constant, their governments will actually move to the left of the political spectrum. This is a corollary of our result that decreasing the proportion of small countries moves the governments to the left. Clearly, when the size of the world is constant, a smaller proportion of small countries necessarily means a higher proportion of large ones and vice versa. Now we might want to consider the implication of simply adding new large countries into the game, keeping  $s$  constant.

While the algebra of the model does not permit us to find an explicit solution to the problem, we can show that when no small countries are present, increasing the number of large countries will shift the large-country governments to the right. Only the economic effect of increased tax competition is present and the subsequent lower equilibrium taxes are going to cause a right-ward shift in governments.

However, in general the effect of increasing  $n$  is ambiguous and depends on  $s$ . When small countries are present, large countries joining the competition have a political effect as well: They will respond to the policy changes of other large countries in the Nash game. In our model, this is the reason why large countries elect governments to the left of the median. Additional large

countries when small countries are present then imply a higher proportion of a large country's competitors being "responsive". This effect then causes the median voters to elect more left-wing representatives. There is a minimum mass of small countries in the competition (a mass comparable in size to the total size of large countries involved), for which the political effect dominates and additional large countries will cause a move to the left. Therefore the effect of more large countries involved in a tax competition depends on the composition of existing competitors and can actually be opposite to that of an increase in the mass of small countries competing for capital.

Let us begin by considering the case when  $s = 0$  to demonstrate that when no small countries are competing for capital with large ones, additional large countries will always cause a shift to the right in the policy maker elected. In equilibrium (2.4) becomes

$$H \equiv \frac{(n-1)[\gamma(1-\alpha) + \tau_L \alpha]}{(1-\tau_L)(1-\alpha)[(n+1-\alpha)(n-1) - (n-2)]} + n - (n-1) \frac{\gamma(1-\alpha) + \tau_L \alpha}{(1-\tau_L)(1-\alpha)} = 0. \quad (2.12)$$

Here, we can express  $\tau_L$  explicitly, namely it is

$$\tau_L = \frac{n(1-\alpha)[(n+1-\alpha)(n-1) - (n-2)] - (n-1)^2(n-\alpha)(1-\alpha)\gamma}{n(1-\alpha)[(n+1-\alpha)(n-1) - (n-2)] + \alpha^2(n-1)^2(n-\alpha)}.$$

We can substitute the above into (2.9) and then differentiate it with respect to  $n$  to obtain a rather complicated explicit expression for  $\frac{\partial \Delta \gamma}{\partial n}$ . We also find that  $\frac{\partial \Delta \gamma}{\partial n} < 0$  for all  $n > \frac{1+\sqrt{1+8\alpha}}{4}$ , which, even for a value  $\alpha = 1$  implies that  $\frac{\partial \Delta \gamma}{\partial n} < 0$  for all  $n > 1$ .

Therefore when no small countries are participating in the tax competition, additional large countries will cause a shift in the governments of the original competitors to the right - as well as decrease the tax rate in each of those countries. There is no political effect present here: For a given large country, when it is competing solely with other large countries, more competitors do not represent an increase in the proportion of countries that respond to its policy changes in the Nash game.

However, when  $s > 0$ , the situation changes.  $G$  defined in (2.8) is now a function of  $\tau_L$  and  $n$ . We still have that  $\partial G(\tau_L, n)/\partial \tau_L < 0$ , therefore  $\partial \tau_L/\partial n > 0$  whenever  $\partial G(\tau_L, n)/\partial n > 0$ .

We can obtain that, holding the equilibrium tax rate constant,

$$\partial G(\tau_L, n)/\partial n > 0 \Leftrightarrow A^2 - \alpha A - 2An - 2A + 2\alpha + \alpha n + 2 + \left[ \frac{\gamma(1-\alpha) + \tau_L \alpha}{(1-\tau_L)(1-\alpha)} - 1 \right] > 0. \quad (2.13)$$

Since we know from (2.4) that  $-1 < \frac{\gamma(1-\alpha) + \tau_L \alpha}{(1-\tau_L)(1-\alpha)} - 1 < 0$ , (2.13) will hold whenever

$$-n^2 + \bar{s}^2 - 2n - (2 + \alpha)\bar{s} + 2\alpha + 1 > 0,$$

where  $\bar{s} \equiv s \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}}$ . The only non-negative solution to the corresponding quadratic equation for  $\bar{s}$  is

$$\bar{s} = \frac{2 + \alpha + \sqrt{4n^2 + 8n - 4\alpha + \alpha^2}}{2}.$$

This value is of the same order of magnitude as  $n$  (given that  $n \geq 2$ ). Moreover, as  $n$  increases,  $\bar{s}$  approaches  $n + 2 + \alpha/2$ .

Therefore  $\partial G(\tau_L, n)/\partial n > 0$ , and thus equilibrium large-country tax rate increases when  $n$  rises, for all  $\bar{s} \in \left( \frac{2+\alpha+\sqrt{4n^2+8n-4\alpha+\alpha^2}}{2}, \infty \right)$ . The actual mass of small countries  $s$  is in fact always smaller than  $\bar{s}$ , since  $\left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} > 1$ , because, as we have already shown,  $\tau_S < \tau_L$ .

Therefore for any number of large countries  $n$ , when there is a sufficiently large mass of small countries competing with them, additional large countries in the competition will increase the equilibrium tax rate adopted by large countries. That mass of small countries is comparable to the total mass of large countries.

Then differentiating (2.9) with respect to  $n$ , realizing that  $\tau_L$  is a function of  $n$ , gives us

$$\frac{\partial \Delta \gamma}{\partial n} = (\partial \tau_L / \partial n) \left( \frac{\alpha}{1-\alpha} + \frac{A}{A-1} \right) + \left( 1 + s \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} \frac{(\partial \tau_L / \partial n)}{(1-\alpha)(1-\tau_L)} \right) \frac{(1-\tau_L)}{(A-1)^2}.$$

Then we have that  $\frac{\partial \Delta \gamma}{\partial n} > 0$  certainly holds whenever  $\partial \tau_L / \partial n > 0$ . Thus by showing that when a minimum number of small countries is present in the competition, additional large countries will increase the equilibrium large-country tax rate means we have also shown that the large-country governments will shift to the left. It is important to note that the tax rate increasing is a sufficient, not a necessary condition for the left-ward shift of elected policy maker. Therefore there is yet another reason (apart from the fact that  $\bar{s} > s$ ), for which we may observe that

additional large countries do not cause a right-ward shift in governments of large countries even for values of  $s < \frac{2+\alpha+\sqrt{4n^2+8n-4\alpha+\alpha^2}}{2}$ .

A higher number of large countries competing for capital will therefore not necessarily have the same effect as a larger mass of small countries: The large countries only have the "right-wing power" as long as no or relatively few small countries are present. Otherwise, the political effect of additional large countries takes over: A significantly higher fraction of a large country's competitors becomes responsive to its policy changes, so that the country's median voter will elect a more left-wing government. More competition in the form of more large countries will then cause a leftward rather than a rightward shift in the governments of large countries.

#### 2.4.4 A Policy Example: Tax Cooperation and Minimum Tax

As a policy exercise in the context of our model, it is interesting to consider the relatively recently renewed proposals for tax harmonization and, in particular, a minimum tax rate rule in the European Union. Immediately upon the enlargement of the Union in 2004, the German Chancellor Schroeder and subsequently the French finance minister Sarkozy criticized the low corporate tax rates in the new countries as tax dumping. They in fact accused the new entrants of intentions to finance parts of their budgets through EU transfers from the old and richer members, rather than through tax revenue. The politicians suggested that while some level of tax competition was desirable, the new members of the bloc were simply lowering their corporate tax rates too far and therefore a minimum tax rate rule should be imposed in the Union. In this section, we demonstrate that in our model such a regime would always improve the welfare for the median agents in the large countries. More interestingly, we also show that provided that the total mass of small countries is not too small, median agents in those countries would also benefit from a minimum tax rule - even though they themselves would oppose it.

The literature on tax cooperation and tax harmonization is almost as rich as the literature on tax competition itself. The research thus far has not arrived at any conclusive results. Rather, depending on the assumptions made, cooperation may be harmful or beneficial. In a setup such as ours, but with homogeneous countries, tax harmonization would clearly increase welfare<sup>3</sup>.

---

<sup>3</sup>If the competing countries were identical, they would all set the same tax rate in equilibrium and they would

However, if we redefined government as maximizing public spending rather than welfare of a particular agent, tax competition can limit wasteful spending through limiting the amount of revenue the government can collect. Edwards and Keen (1996) provide a treatment of the issue and consider intermediate types of government. Also, Kehoe (1989) demonstrates that coordination is undesirable when tax competition helps solve governments' credibility problems. More recently, Brueckner (2001) shows that for a particular type of tax cooperation taxes may actually decrease and therefore welfare can fall relative to free tax competition.

Moreover, while some level of cooperation may indeed be desirable, it can be difficult to implement. In particular, we would need either institutions powerful enough to prevent individual governments from deviating in their policy from the cooperative equilibrium levels, or incentives that would render such deviations undesirable. Also, full tax cooperation may be problematic, especially when tax competition occurs between a larger number of countries, due to the complexity of the decision process that is the real-world equivalent of maximizing the weighted sum of welfares of the respective countries.

For both of the above reasons, here we shall simply evaluate solely in the context of our model a particular and easily implementable form of tax cooperation<sup>4</sup> that has been suggested in direct relation to the subject discussed above - namely the one proposed by the German and French policy makers. The large (and rich) countries would select a minimum corporate tax rate for all the members of the Union. It would be the tax rate that the large countries would set in a tax competition, however, it would be binding on the small countries, which, in a purely competitive environment, would choose lower rates. Therefore no implementation mechanism would be necessary for the large countries. The small countries, on the other hand, would require an incentive not to deviate and decrease their corporate tax rates. Given that for the foreseeable future these countries are mostly net recipients of transfers from the EU budget, i.e. indirectly mostly from the large, wealthy countries, making the transfers conditional upon sufficiently high a tax rate could provide such an incentive. While requiring independent

---

all have the same amount of capital per capita in the economy. Therefore with full tax harmonization, they would retain the same amount of capital, but would be able to set a much higher tax rate. In fact, since supply of capital would be inelastic, all countries would tax away all capital income. This would increase lump-sum transfers to agents and thus their welfare.

<sup>4</sup>Here we suggest that the mechanism for tax cooperation is easily implementable once approved. Of course, under the current EU rules, dissent by a single member of the Union would preclude such an approval.

countries to render any payments whenever they amend their tax policy in a certain manner may be unenforceable, simply reducing the size of payments they are due in the first place certainly is very implementable.

We now want to investigate the welfare impact of such a minimum tax in the large and small countries. Of course, the impact on the welfare of entire countries would depend on the weights we would place on the various agents and on the distribution of productivities and endowments among these agents. Therefore we shall only consider the welfare of the median agents

Note that such an approach might actually pose a problem if the total supply of capital in our world were elastic. Then consider a complete tax cooperation, where a planner maximizes the welfare of all median agents across the countries at the same time. As noted before, it increases median agents' welfare in our model. However, with elastic total supply of capital, the higher taxes in the cooperation case would decrease the supply of capital in the world, and therefore also production and total welfare. Then even if indeed preferred by the median voters, tax harmonization would certainly not be as desirable.

Returning to our setup, median agents in the large countries would be unequivocally better off. First, since we start from a point, where  $\tau_L > \tau_S$ , we also have  $k_L < k_S$ . There is less capital per capita in the large countries. Once all countries tax capital at the same rate,  $k_L = k_S = 1$ , i.e. capital moves from small countries to the large ones. From (2.4) we know that when  $\tau_S$  increases and thus  $A$  decreases for any given tax rate in the large countries, the derivative of the welfare function becomes positive. Since preferences of the median agent are quasi-concave, this implies that  $\tau_L$  will be higher in all large countries in the new equilibrium. Therefore the minimum tax rate the large countries will want to establish will be in fact higher than their own prevailing tax rate under an unconstrained tax competition that includes a positive mass of small countries.

More importantly, though, it means that the minimum tax rule will bring both a higher tax rate and more capital per capita in the large countries, thus increasing wages and transfers, thereby resulting in a higher level of welfare for the median agents. The median voters in those countries will therefore always support a minimum tax rate rule of the form described above.

The impact on the median agents in the small countries is slightly more complicated. Clearly, the only reason for small countries in our model to select a lower tax rate is that they, each on

their own, do not affect the world equilibrium net return on capital. This effect works directly and indirectly - through the small countries electing governments to the right of those in large countries. However, a mass  $s$  of small countries does impact the world equilibrium and therefore a social planner maximizing median-agent welfare for all small countries can improve on that welfare by taking this into account. Let us therefore compare the tax rate this planner would choose to the tax rate large countries would like to impose on the small countries. Given the quasi-concavity of preferences, once we show that the planner would select a tax rate higher than the minimum tax rate required by the large countries, we will know that the median agents in the small countries would be better off under a minimum tax rate rule than when left to compete for capital freely.

When  $s = 0$ , this social planner would perform the same maximization as individual small-country governments. Therefore at zero total mass of small countries, those countries are necessarily worse off at a higher tax rate imposed by the large countries.

When  $s > 0$ , we can write the planner's first-order condition similarly to (2.4), i.e.

$$\left[ \left( \frac{1 - \tau_L}{1 - \tau_S} \right)^{\frac{1}{1-\alpha}} \frac{n(\partial\tau_L/\partial\tau_S)}{(1 - \tau_L)(1 - \alpha)} - \frac{A_S - s}{(1 - \tau_S)(1 - \alpha)} \right] [\gamma(1 - \alpha) + \tau_S\alpha] + A_S = 0, \quad (2.14)$$

where

$$A_S \equiv n \left( \frac{1 - \tau_L}{1 - \tau_S} \right)^{\frac{1}{1-\alpha}} + s.$$

The response of large countries to a cooperative mass of small countries in the above is a modified version of (2.7), i.e.

$$\partial\tau_L/\partial\tau_S = s [(A_L + 1 - \alpha)(A_L - 1) - (n - 1)]^{-1},$$

where  $A_L$  is analogous to the one defined in (2.1).

The FOC of large countries depends on how they perceive the minimum tax rule. Since we want to show that a planner for the small countries prefers a higher tax than that imposed on the small countries by their large counterparts, let us consider the option that results in a higher tax rate selected by the large countries: We shall assume that any particular large country believes that a response by other large countries to its own policy will induce the same

change in the minimum tax rule and therefore an equal increase or decrease in the small country tax rate. Mathematically, this implies that the modified FOC for the large countries is

$$0 = \left[ \left( \frac{1 - \tau_S}{1 - \tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{s(\partial\tau_S/\partial\tau'_L)}{(1 - \tau_S)(1 - \alpha)} + \left( \frac{1 - \tau_L}{1 - \tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n - 1)(\partial\tau_L/\partial\tau'_L)}{(1 - \tau_L)(1 - \alpha)} - \frac{A_L - 1}{(1 - \tau'_L)(1 - \alpha)} \right] \cdot [\gamma(1 - \alpha) + \tau'_L\alpha] + A_L, \quad (2.15)$$

where we have not only  $\tau_S = \tau_L$ , but also

$$\partial\tau_S/\partial\tau'_L = \partial\tau_L/\partial\tau'_L = [(A_L + 1 - \alpha)(A_L - 1) - (n - 2) - s]^{-1}.$$

We can see that  $\partial\tau_L/\partial\tau'_L < \partial\tau_L/\partial\tau_S$  whenever

$$(s - 1)[(A_L + 1 - \alpha)(A_L - 1) - (n - 1) - s] > 0.$$

This always holds as long as  $s > 1$ .

Let us see what the sign is of the derivative of the welfare function for the planner for small countries when countries are in an equilibrium where  $\tau_S = \tau_L$  is imposed upon the small countries. (2.15) holds and therefore (2.14) simplifies to

$$\partial W/\partial\tau_S = (s - 1)(1 - \partial\tau_L/\partial\tau'_L) + n(\partial\tau_L/\partial\tau_S - \partial\tau_L/\partial\tau'_L).$$

Since  $\partial\tau_L/\partial\tau'_L < \partial\tau_L/\partial\tau_S$  and  $\partial\tau_L/\partial\tau'_L < 1$  for  $s > 1$ , for those sizes of the total mass of small countries we will have  $\partial W/\partial\tau_S > 0$ .

Therefore, since his preferences are quasi-concave, the planner for the mass of small countries will ideally desire a tax rate higher than that imposed upon the small countries by the large ones under our minimum-tax rule. Then he will prefer the minimum tax to the tax chosen by the individual small countries in a free tax competition, since the latter is lower than the former.

Thus small countries are also better off under the minimum-tax regime we have described, provided their total mass is equal or greater than that of a single large country. Since no planner exists that would force them to adopt a higher tax rate that would maximize their total welfare

(i.e. there is no mechanism to induce a cooperation between the small countries in the face of competition from their large counterparts), they will certainly not want to voluntarily join such a regime and will voice their opposition to it - just as the real-world new entrants to the EU have in 2004 and since then. However, the regime would improve the welfare of all parties concerned (or, at the very least, of the median agent in each country). Our model then suggests that since the potential enforcement mechanism - the transfers to poorer members, who also happen to be small - already exists, the Union might want to try setting a minimum corporate tax rate for its members despite disagreement from some of them, as difficult as that might be under the current voting rules of the EU.

## **2.5 Extensions**

### **2.5.1 Small Countries with Positive Size**

Our assumption of infinitesimally small countries competing with large countries may not seem particularly realistic, however, it yields a very plausible premise in our model - that small countries do not consider their impact on the world (or, specifically, on the group of countries, with which they most directly compete for capital). The new small entrants to the EU almost certainly do not take into account any effect on the EU equilibrium at all when making their individual tax-rate decisions. This results in the small countries in our model always electing the median agent to represent them, rather than a government to the left of the median as is the case in large countries.

However, in conjunction with the standard (and algebraically manageable) Cobb-Douglas function, the assumption of infinitesimal size of small countries results in a constant tax rate implemented by them. That is, they do not react to changes in the competition-area equilibrium by modifying their tax policies. This in turn implies that the difference between the ex-ante and ex-post tax rate for the large countries will be smaller than in a more general case. Only other large countries respond to a change in the tax policy of a particular large country with an increase or decrease of their own corporate tax rates. Small countries do not respond at all.

To verify that our results hold in a more general setting whose assumptions do not imply the above simplifications, we want to modify our model so that each individual small country

will have a positive size. First, these countries will understand that modifying their own tax rate will alter the world net return on capital. Therefore they will perceive a lower elasticity of capital with respect to their tax rate than infinitesimally small countries. Still, though, since they are smaller than their large competitors, their impact on the world equilibrium is not as significant, the elasticity of capital is therefore larger and thus also their tax rates are smaller than those implemented by the large countries. Similarly, because their tax policy modifications induce a smaller response in other countries, small and large, than a policy change by large countries, they will have a smaller disparity between the median agent's ex-ante and ex-post optimal tax rates. This in turn will spell a government to the right of those elected in large countries.

Second, when large countries compete with proportionately more small countries (or additional small countries enter the tax competition), their governments will still shift to the right. The small countries' lower tax rates will induce a decrease in the rates of the large countries, which will in turn reduce the difference between the desired ex-ante and ex-post tax rates by the median agents in those countries.

All of the above is difficult to show algebraically, therefore we will rely on a computational solution for a selected calibration in order to verify that our results hold. First, we have to modify the two first-order conditions that define our symmetric equilibrium. The median agents in the small countries will ex-ante desire the tax rate that satisfies

$$0 = G_S \equiv \left[ \left( \frac{1 - \tau_S}{1 - \tau'_S} \right)^{\frac{1}{1-\alpha}} \frac{(s - \sigma)(\partial \tau_S / \partial \tau'_S)}{(1 - \tau_S)(1 - \alpha)} + \left( \frac{1 - \tau_L}{1 - \tau'_S} \right)^{\frac{1}{1-\alpha}} \frac{n(\partial \tau_L / \partial \tau'_S)}{(1 - \tau_L)(1 - \alpha)} - \frac{A_S - 1}{(1 - \tau'_S)(1 - \alpha)} \right] \cdot [\gamma(1 - \alpha) + \tau'_S \alpha] + A_S, \quad (2.16)$$

where

$$A_S \equiv \sigma + n \left( \frac{1 - \tau_L}{1 - \tau'_S} \right)^{\frac{1}{1-\alpha}} + (s - \sigma) \left( \frac{1 - \tau_S}{1 - \tau'_S} \right)^{\frac{1}{1-\alpha}},$$

$s$  is still the total mass of small countries and  $\sigma < 1$  is the size of each individual small country.

In the large countries it will be

$$0 = G_L \equiv \left[ \left( \frac{1-\tau_S}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{s(\partial\tau_S/\partial\tau'_L)}{(1-\tau_S)(1-\alpha)} + \left( \frac{1-\tau_L}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n-1)(\partial\tau_L/\partial\tau'_{;L})}{(1-\tau_L)(1-\alpha)} - \frac{A_L-1}{(1-\tau'_L)(1-\alpha)} \right] \cdot [\gamma(1-\alpha) + \tau'_L\alpha] + A_L, \quad (2.17)$$

where  $A_L$  is analogous to the one defined in (2.1). The couples  $\partial\tau_S/\partial\tau'_S$  and  $\partial\tau_L/\partial\tau'_S$ , as well as  $\partial\tau_S/\partial\tau'_L$  and  $\partial\tau_L/\partial\tau'_{;L}$  will now each be solutions to systems of two linear equations.

We can rewrite the large and small country ex-post FOCs for a particular large country with tax  $\tau''_L$  and a particular small country with tax  $\tau''_S$ , separating in both the large country with tax  $\tau'_L$  from the remaining large countries. Then we obtain, similarly to (2.5),

$$0 = F_L \equiv - \left[ 1 + (n-2) \left( \frac{1-\tau_L}{1-\tau''_L} \right)^{\frac{1}{1-\alpha}} + \left( \frac{1-\tau'_L}{1-\tau''_L} \right)^{\frac{1}{1-\alpha}} + s \left( \frac{1-\tau_S}{1-\tau''_L} \right)^{\frac{1}{1-\alpha}} \right] \cdot [(\gamma - \Delta\gamma_L)(1-\alpha) + \tau''_L\alpha - (1-\tau''_L)(1-\alpha)] + (\gamma - \Delta\gamma_L)(1-\alpha) + \tau''_L\alpha,$$

and

$$0 = F_S \equiv - \left[ \sigma + (n-1) \left( \frac{1-\tau_L}{1-\tau''_S} \right)^{\frac{1}{1-\alpha}} + \left( \frac{1-\tau'_L}{1-\tau''_S} \right)^{\frac{1}{1-\alpha}} + (s-\sigma) \left( \frac{1-\tau_S}{1-\tau''_S} \right)^{\frac{1}{1-\alpha}} \right] \cdot [(\gamma - \Delta\gamma_S)(1-\alpha) + \tau''_S\alpha - (1-\tau''_S)(1-\alpha)] + (\gamma - \Delta\gamma_S)(1-\alpha) + \tau''_S\alpha.$$

Here we realize that both large and small country governments will be to the left of the median agent, however, not by the same amount. Therefore we call the difference in productivity between the median agent and the government he elects  $\Delta\gamma_L$  in a large country and  $\Delta\gamma_S$  in a small country. Once again, this is assuming a symmetric equilibrium where all countries of the same size have the same government.

The left-ward shift in a small country is the solution to the equation

$$\Delta\gamma_S = \left[ \left( \frac{1-\tau_S}{1-\tau'_S} \right)^{\frac{1}{1-\alpha}} \frac{(s-\sigma)(\partial\tau_S/\partial\tau'_S)}{(1-\tau_S)(1-\alpha)} + \left( \frac{1-\tau_L}{1-\tau'_S} \right)^{\frac{1}{1-\alpha}} \frac{n(\partial\tau_L/\partial\tau'_S)}{(1-\tau_L)(1-\alpha)} \right] \cdot [\gamma(1-\alpha) + \tau'_S\alpha] (1-\tau'_S)(A_S-1)^{-1}, \quad (2.18)$$

whereas in a large country it is

$$\Delta\gamma_L = \left[ \left( \frac{1-\tau_S}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{s(\partial\tau_S/\partial\tau'_L)}{(1-\tau_S)(1-\alpha)} + \left( \frac{1-\tau_L}{1-\tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n-1)(\partial\tau_L/\partial\tau'_{L})}{(1-\tau_L)(1-\alpha)} \right] \cdot [\gamma(1-\alpha) + \tau'_L\alpha] (1-\tau'_L)(A_L-1)^{-1}, \quad (2.19)$$

Then, as in (2.6), we have

$$d\tau_L/d\tau'_L = -\frac{\partial F_L/\partial\tau'_L + \partial F_L/\partial\tau_S \cdot d\tau_S/d\tau'_L}{\partial F_L/\partial\tau_L + \partial F_L/\partial\tau''_L}$$

and

$$d\tau_S/d\tau'_L = -\frac{\partial F_S/\partial\tau'_L + \partial F_S/\partial\tau_L \cdot d\tau_L/d\tau'_L}{\partial F_S/\partial\tau_S + \partial F_S/\partial\tau''_S}.$$

Solving this system of equations, we obtain

$$d\tau_L/d\tau'_L = \frac{(\partial F_L/\partial\tau'_L)(\partial F_S/\partial\tau_S + \partial F_S/\partial\tau''_S) - (\partial F_S/\partial\tau'_L)(\partial F_L/\partial\tau_S)}{(\partial F_S/\partial\tau_L)(\partial F_L/\partial\tau_S) - (\partial F_L/\partial\tau''_L + \partial F_L/\partial\tau_L)(\partial F_S/\partial\tau_S + \partial F_S/\partial\tau''_S)}$$

and

$$d\tau_S/d\tau'_L = \frac{(\partial F_S/\partial\tau'_L)(\partial F_L/\partial\tau_L + \partial F_L/\partial\tau''_L) - (\partial F_S/\partial\tau_L)(\partial F_L/\partial\tau'_L)}{(\partial F_S/\partial\tau_L)(\partial F_L/\partial\tau_S) - (\partial F_L/\partial\tau''_L + \partial F_L/\partial\tau_L)(\partial F_S/\partial\tau_S + \partial F_S/\partial\tau''_S)}.$$

Rewriting  $F_L$  and  $F_S$  so that we separate the small country with tax  $\tau'_S$  from the remaining small countries and include the large country with tax  $\tau'_L$  with the other large countries (with tax rate  $\tau_L$ ), we can similarly obtain the expressions for the partial derivatives in (2.16).

## Calibration

Now we can solve the system of two equations with two unknowns, (2.17) and (2.16), having substituted in for the leftward shift parameters  $\Delta\gamma_S$  and  $\Delta\gamma_L$  from (2.18) and (2.19), respectively. To do so, we calibrate the model: We set  $\gamma = 3/4$ , implying that the wages of the median agents will be at 75% of the average level in each economy. Further, we choose the standard share of capital in the production function,  $\alpha = 1/3$ . We compute the equilibrium values of  $\tau_L, \tau_S, \Delta\gamma_L$  and  $\Delta\gamma_S$  for the baseline model of small countries with infinitesimal populations as well as for small countries with size  $\sigma = 1/10$  and  $\sigma = 1/4$  (remember, the large countries

have size one).

We perform our computations in three series. First, we keep the size of the world constant, so that as we increase  $n$ , there is a compensating decrease in  $s$ . Second, we set  $n = 3$  and vary the value of  $s$ , and finally, we anchor the total mass of small countries at  $s = 3$  and vary the number of small countries. When small countries have infinitesimal populations, interesting interactions only occur for at least two large countries, therefore we vary  $n$  so that  $n \geq 2$ . We consider values of  $s \geq 0$ .

### **Computational Results**

The results of the computations are presented in Figures 3-8. We can see that introducing positive size for our small countries does not significantly alter the results of the model. As expected, the large country tax rates are higher for any combination of  $n$  and  $s$  because they are facing less of a tough competitor: The small countries themselves have higher taxes because they realize they affect, albeit not too significantly, the world equilibrium net return on capital. Also, their governments are slightly to the left of the median, even if much less so than in large countries.

As expected, the tax rate in the large countries decreases and their governments shift to the right as a function of higher proportion or number of small countries, not only in the baseline case of infinitesimal small countries, but also when the mass of each of those countries is positive. The government only becomes more right-wing when additional large countries enter the competition when  $n$  is sufficiently large relative to  $s$ , again, even when small countries have positive mass. As we mentioned in our theoretical discussion, an increase in the tax rate is not necessary to have a left-ward shift as a consequence of more large countries joining the competition: Indeed, as  $n$  increases, keeping  $s$  constant, the equilibrium tax rate in large countries drops, and yet the government first shift slightly to the left and then to the right. This further shows that large countries do not have the "right-wing power" that their smaller counterparts possess.

### 2.5.2 Capital Ownership by Median Agents

Throughout our analysis, we assumed that median agents (and those less productive than they are) do not hold any capital. While it is certainly plausible that if those agents are less productive than the average agent in the population, they possess even less in terms of relative capital endowment, we might still want to consider a case where they do own capital stock.

We shall make two alternative simple assumptions regarding capital ownership in the economy. First, we assume that all agents with below-average productivity hold  $\gamma$  units of capital, where  $\gamma$  is the productivity of the median agent. Thus all agents are endowed with capital, but that endowment does not vary among agents with median or lower productivity. This will mean that, relative to our results thus far, the median agents in the large countries will desire a lower tax rate, both ex-post and ex-ante, because their welfare is decreasing in tax beyond the drop in wage experienced by agents with no capital endowment: Their net income from capital holdings will also decrease as a function of corporate tax implemented by their own country. Since both ex-post and ex-ante tax rates will be lower, so will the difference between them, and the median voter will elect a more right-wing government than in the case when he and those poorer than him own no capital.

On the other hand, when we assume instead that both labor productivity and capital endowment follow the same distribution in every economy, electing someone with a lower productivity will also mean choosing a representative with a lower capital endowment - both reasons for desiring a higher ex-post tax rate. Choosing a poorer representative will be more "effective" at having the median voter's ideal ex-ante tax rate implemented. A poorer agent closer in type to the median voter will, when elected, select that tax rate. The median agent, when agents to the left of him vary in their capital endowments, will not need to elect as left-wing a government to achieve his desired ex-ante tax policy. The elected policy maker will be more to the right than in the previous case.

Mathematically, the solution for small countries is exactly as before. Since small countries consider the net return on capital to their agents as exogenous, capital ownership does not enter the solution for their optimal tax rate. In the case of the large countries, we can modify the

first-order condition in (2.4) to obtain

$$0 = \left( \frac{1 - \tau_L}{1 - \tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n-1)\partial\tau_L/\partial\tau'_L}{(1 - \tau_L)(1 - \alpha)} [K^T [\gamma(1 - \alpha) + \tau'_L\alpha] - A\gamma(1 - \alpha)(1 - \tau'_L)] \\ - \frac{A - 1}{(1 - \tau'_L)(1 - \alpha)} [\gamma(1 - \alpha) + \tau'_L\alpha] + A(K^T - \gamma).$$

Let us call  $\beta$  a parameter that depends on the distribution of capital for agents below median productivity: If all those agents hold the same amount of capital as the median agent, let  $\beta = 0$ . If, on the other hand, capital endowment distribution is the same as the distribution of labor productivity, let  $\beta = 1$ . Then the difference between the productivity of the agent in the government and the median agent is defined by

$$\Delta\gamma = \left[ \frac{(A - 1)K^T}{1 - \tau'_L} + \beta A \right]^{-1} \left( \frac{1 - \tau_L}{1 - \tau'_L} \right)^{\frac{1}{1-\alpha}} \frac{(n-1)\partial\tau_L/\partial\tau'_L}{(1 - \tau_L)(1 - \alpha)} \cdot [K^T [\gamma(1 - \alpha) + \tau'_L\alpha] - A\gamma(1 - \alpha)(1 - \tau'_L)].$$

It is rather complicated to show all of our results algebraically when the median agents in the large countries hold capital. Therefore, again, we choose to computationally solve for the equilibrium in a calibrated version of our model. The calibration and values for  $n$  and  $s$  are the same as before.

### Computational Results

We present the results in Figures 9-14. As in the case of positive mass of individual small countries, capital ownership by the median agents in the various countries does not qualitatively affect our results.

However, as mentioned before, two effects are present: When median agents own capital, they desire and subsequently implement lower tax rates. Moreover, with lower taxes, the magnitude of difference between the median agent's ex-post and ex-ante optimal tax rates drops - since both of those rates decrease. Then he does not need to elect an agent as different from himself to implement his desired ex-ante tax rate. Therefore, governments are going to be more right-wing for all values of  $n$  and  $s$  when median agents in the various countries own capital.

The second effect is present when  $\beta = 1$ , i.e. when capital ownership distribution is identical to labor productivity distribution in the population. The distribution in capital ownership among agents poorer than the median, i.e. among those he considers electing, renders choosing a different policy maker than oneself more powerful: For any difference between types, the policies chosen by them ex-post will be more unlike each other. Therefore while assuming  $\beta = 1$  instead of  $\beta = 0$ , ceteris paribus, does not change the equilibrium tax rates, it will shift the large-country governments further to the right and closer to the median agent.

## 2.6 Conclusion

In this paper we have first showed that smaller countries present a challenge to their larger counterparts in a tax competition setting. Because of their size, their fiscal policy affects little or not at all the world equilibrium net return on capital. Therefore they will perceive a higher elasticity of capital with respect to their corporate tax and choose a lower tax on capital than larger countries, with which they compete. Moreover, countries with larger populations that realize the impact of their policy decisions on those of other countries will elect more left-wing governments than small countries, leading to an even larger difference between their tax rates.

Then we captured the political impact of country-size heterogeneity on the results of tax competition. Namely, we have shown that small countries have a right-wing power in a tax competition setting. Proportionately more small countries in a group of countries attempting to attract capital will lead to a right-ward shift in the governments that implement policy in the large countries, as will an addition of more small countries to the group. Additional large countries, on the other hand, do not clearly lead to a similar shift to the right.

Through this mechanism large countries are able to better respond to the tougher competition for capital resources they face when interacting with small countries. However, even when they elect a fiscally more conservative government and thus further decrease their taxes, they still lose some of their capital to their small competitors. Therefore, they have a further incentive to search for a welfare-improving partial tax cooperation or harmonization solution.

In the case of the recent EU enlargement, when mostly small countries entered the bloc, a possibly easily implementable solution presents itself: A minimum tax rate requirement backed

by the threat of reduced transfers from the Union to the budgets of countries that do not comply. We therefore explore this proposal in our framework and find that it would be beneficial to all parties concerned as long as a sufficiently large number of small countries is present in the competition - which, by now, almost certainly is the case in the European Union.

Overall, there is a multitude of factors that help determine the type of government that voters choose to elect in any particular country. Those factors are not only economic, but also social, cultural and others. Within economic determinants of government choice, taxation also does not stand alone. However, here we have attempted to show that, *ceteris paribus*, the composition of countries involved in tax competition will help determine who voters elect to represent them.

## 2.7 Appendix

### Appendix 1: Single-Crossing Property

The single-crossing property of Gans and Smart (1996) is satisfied when  $\tau > \tau'$ ,  $\gamma < \gamma'$  and  $\theta \leq \theta'$  or if  $\tau < \tau'$ ,  $\gamma > \gamma'$  and  $\theta \geq \theta'$ , then  $W(\tau, \gamma, \theta) \geq W(\tau', \gamma, \theta)$  implies  $W(\tau, \gamma', \theta') \geq W(\tau', \gamma', \theta')$ .

We note that both in a small and in a large country, less capital enters the economy when the tax rate is higher. Also, in a large country, the net return on capital will be lower when the tax rate is increased: Capital leaves the country, which increases the gross return on capital, but taxation results in a lower net rate of return: The leaving capital enters other economies and when their tax rates remain the same, their gross returns and thus also net returns on capital decrease. Therefore by arbitrage the domestic net return on capital must decrease.

Then we have  $W(\tau, \gamma, \theta) = \gamma(1-\alpha)k^\alpha + \theta(1-\tau)\alpha k^{\alpha-1} + \tau\alpha k^\alpha \geq W(\tau', \gamma, \theta) = \gamma(1-\alpha)k'^\alpha + \theta(1-\tau')\alpha k'^{\alpha-1} + \tau'\alpha k'^\alpha$  and when  $\tau < \tau'$  also  $k^\alpha < k'^\alpha$ ,  $(1-\tau)\alpha k^{\alpha-1} \leq (1-\tau')\alpha k'^{\alpha-1}$  (the latter with equality in the case of a small country). Thus this implies  $\gamma(1-\alpha)(k'^\alpha - k^\alpha) + \theta[(1-\tau')\alpha k'^{\alpha-1} - (1-\tau)\alpha k^{\alpha-1}] \leq \tau\alpha k^\alpha - \tau'\alpha k'^\alpha$  and since  $k'^\alpha - k^\alpha > 0$  and  $(1-\tau')\alpha k'^{\alpha-1} - (1-\tau)\alpha k^{\alpha-1} > 0$ , when  $\gamma < \gamma'$  and  $\theta \leq \theta'$  we necessarily also have  $\gamma'(1-\alpha)(k'^\alpha - k^\alpha) + \theta'[(1-\tau')\alpha k'^{\alpha-1} - (1-\tau)\alpha k^{\alpha-1}] \leq \tau\alpha k^\alpha - \tau'\alpha k'^\alpha$  and therefore  $W(\tau, \gamma', \theta') \geq W(\tau', \gamma', \theta')$ . Similarly, we could show the implication for  $\tau < \tau'$ ,  $\gamma > \gamma'$  and  $\theta \geq \theta'$ .

Since the agents' preferences satisfy the single crossing property in each country, a Condorcet winner exists and he represents the optimum for the median voter.

## Appendix 2: Quasi-Concave Preferences

We want to show that the ex-ante as well as ex-post preferences of the median agent in the large country and anyone poorer than him (i.e. with a lower value of the productivity parameter  $\gamma$ ) are quasi-concave. Since none of these agents own any capital, we can set  $\theta = 0$  in our analysis.

We shall proceed as follows. First, we show that at any point where the first-order condition is satisfied, the second order condition will be negative, meaning that the local extremum is in fact a local maximum. Second, we verify that at the lowest possible value of the country's tax rate,  $\tau'_L = 0$ , the slope of the welfare function is positive. Last, we show that welfare is zero at the highest possible value of the tax rate,  $\tau'_L = 1$ . Since welfare is positive at  $\tau'_L = 0$ , then we will have ascertained that the welfare function is quasi-concave for  $\tau'_L \in [0, 1]$  and attains a maximum on that interval at a point  $\tau'_L \in (0, 1)$ .

For the ex-ante welfare function, we begin by substituting for  $\partial\tau_L/\partial\tau'_L$  into the first-order condition, (2.4), obtaining

$$0 = \frac{(n-1)[\gamma(1-\alpha) + \tau'_L\alpha]}{(1-\tau'_L)(1-\alpha)[(A+1-\alpha)(A-1) - (n-2)]} - \frac{A'-1}{(1-\tau'_L)(1-\alpha)} [\gamma(1-\alpha) + \tau'_L\alpha] + A'. \quad (2.20)$$

Then the second derivative of the ex-ante welfare function is

$$\begin{aligned} \frac{\partial^2 W}{\partial \tau_L'^2} &= \frac{1}{1-\tau'_L} \left[ \frac{n-1}{1-\alpha} \left[ \frac{\gamma(1-\alpha) + \tau'_L\alpha + \alpha(1-\tau'_L)}{(1-\tau'_L)[(A+1-\alpha)(A-1) - (n-2)]} \right. \right. \\ &\quad \left. \left. - \frac{[\gamma(1-\alpha) + \tau'_L\alpha](2A-\alpha)}{(1-\tau'_L)(1-\alpha)[(A+1-\alpha)(A-1) - (n-2)]^2} \left( \frac{1-\tau'_L}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} \right] + \right. \\ &\quad \left. + A' - 1 - \frac{(A'-1)[\gamma(1-\alpha) + \tau'_L\alpha]}{1-\tau'_L} \left[ \frac{1}{1-\alpha} + \frac{1}{(1-\alpha)^2} \right] \right]. \end{aligned}$$

After substituting in from (2.20), we can simplify the above to

$$\begin{aligned} \frac{\partial^2 W}{\partial \tau_L'^2} &= \left[ \left( \frac{1 - \tau_L'}{1 - \tau_L} \right)^{\frac{1}{1-\alpha}} \frac{(2A - \alpha)}{(1 - \alpha) [(A + 1 - \alpha)(A - 1) - (n - 2)]} \left[ \frac{(A' - 1) [\gamma(1 - \alpha) + \tau_L' \alpha]}{(1 - \alpha)(1 - \tau_L')} - A' \right] \right. \\ &\quad \left. + \frac{\alpha(n - 1)}{(1 - \alpha) [(A + 1 - \alpha)(A - 1) - (n - 2)]} - 1 - \frac{(A' - 1) [\gamma(1 - \alpha) + \tau_L' \alpha]}{(1 - \alpha)^2 (1 - \tau_L')} \right] \cdot \frac{1}{1 - \tau_L'}. \end{aligned}$$

Remember, we want to demonstrate that  $\frac{\partial^2 W}{\partial \tau_L'^2} < 0$ . We can first show that

$$\left( \frac{1 - \tau_L'}{1 - \tau_L} \right)^{\frac{1}{1-\alpha}} \frac{(2A - \alpha)}{[(A + 1 - \alpha)(A - 1) - (n - 2)]} < 1$$

is satisfied whenever

$$n^2 - n(3 + \alpha) + 2(1 + \alpha) + 2(n - 1)s \left( \frac{1 - \tau_S}{1 - \tau_L} \right)^{\frac{1}{1-\alpha}} + s^2 \left( \frac{1 - \tau_S}{1 - \tau_L} \right)^{\frac{2}{1-\alpha}} > 0.$$

But this surely holds for all  $n \geq 2$ . Therefore we can simplify our problem to showing that

$$\frac{- \left( \frac{1 - \tau_L'}{1 - \tau_L} \right)^{\frac{1}{1-\alpha}} (2A - \alpha) A' + \alpha(n - 1)}{(1 - \alpha) [(A + 1 - \alpha)(A - 1) - (n - 2)]} - 1 < 0.$$

This is equivalent to

$$-(2A - \alpha)A + \alpha(n - 1) - (1 - \alpha) [(A + 1 - \alpha)(A - 1) - (n - 2)] < 0,$$

which, realizing that  $A \geq n - 1$  with certainty, holds whenever

$$- [2(n - 1) - \alpha] (n - 1) + \alpha(n - 1) < 0.$$

This once again surely holds for all  $n \geq 2$ .

We have thus shown that  $\frac{\partial^2 W}{\partial \tau_L'^2} < 0$  at any local extremum and therefore also that any local extrema of the welfare function are necessarily local maxima.

Now we can also see that at  $\tau'_L = 0$ ,

$$\frac{\partial W}{\partial \tau'_L} = \frac{(n-1)\gamma}{(A+1-\alpha)(A-1)-(n-2)} - (A'-1)\gamma + A',$$

which is clearly positive for the median agent (for whom  $\gamma < 1$ ) or for anyone with a lower-than-median productivity.

At  $\tau = 1$ , no capital is present in the country, because the net marginal return on capital is zero even when there is no capital in the economy. Then the marginal product of labor and thus wages, as well as tax revenues are zero.

Therefore we have determined that the ex-ante welfare function is quasi-concave on the interval  $\tau'_L \in [0, 1]$  for  $\gamma < 1$ , and attains its maximum on that interval for a tax rate  $\tau'_L \in (0, 1)$ . Welfare first increases as a function of the tax rate, reaches its maximum, and then subsequently drops to zero when all capital income is (or, rather, would be) taxed away.

We can easily obtain the same result for the ex-post welfare, simply realizing that we need to consider the first order condition in (2.2). The second derivative here is

$$\frac{\partial^2 W}{\partial \tau'^2_L} = \frac{1}{1-\tau'_L} \left[ A' - 1 - \frac{(A'-1)[\gamma(1-\alpha) + \tau'_L \alpha]}{1-\tau'_L} \left[ \frac{1}{1-\alpha} + \frac{1}{(1-\alpha)^2} \right] \right].$$

Once again, we substitute in from (2.2) to obtain the simplified expression

$$\frac{\partial^2 W}{\partial \tau'^2_L} = -1 - \frac{A'}{1-\alpha}.$$

This will always be negative. The derivative of the welfare function at  $\tau'_L = 0$  is

$$\frac{\partial W}{\partial \tau'_L} = -(A'-1)\gamma + A'$$

and therefore remains positive for the median or poorer agent. Our conclusion for the ex-ante welfare function thus holds for the ex-post function as well. Both are quasi-concave.

### Appendix 3: Impact of Small Countries

We first show that  $\partial G(\tau_L, s)/\partial \tau_L < 0$  when evaluated at  $G(\tau_L, s) = 0$ . We have

$$\begin{aligned} \frac{\partial G(\tau_L, s)}{\partial \tau_L} &= \frac{1}{1-\tau_L} \left[ \frac{n-1}{1-\alpha} \left[ \frac{\gamma(1-\alpha) + \alpha}{(1-\tau_L)[(A+1-\alpha)(A-1) - (n-2)]} - \right. \right. \\ &\quad \left. \left. - \frac{[\gamma(1-\alpha) + \tau_L\alpha](2A-\alpha)(A-1)}{(1-\tau_L)(1-\alpha)[(A+1-\alpha)(A-1) - (n-2)]^2} \right] + A-1 - \right. \\ &\quad \left. - \frac{(A-1)[\gamma(1-\alpha) + \tau_L\alpha]}{1-\tau_L} \left[ \frac{1}{1-\alpha} + \frac{1}{(1-\alpha)^2} \right] \right]. \end{aligned}$$

After substituting in from  $G(\tau_L, s) = 0$ , we can simplify the above to

$$\begin{aligned} \frac{\partial G(\tau_L, s)}{\partial \tau_L} &= \frac{1}{1-\tau_L} \left[ -\frac{(n-1)[\gamma(1-\alpha) + \tau_L\alpha](2A-\alpha)(A-1)}{(1-\tau_L)(1-\alpha)^2[(A+1-\alpha)(A-1) - (n-2)]^2} \right. \\ &\quad \left. - 1 - \frac{\alpha A(1-\tau_L)}{\gamma(1-\alpha) + \tau_L\alpha} - \frac{(A-1)}{1-\alpha} \left[ \frac{\gamma(1-\alpha) + \tau_L\alpha}{(1-\tau_L)(1-\alpha)} - \alpha \right] \right]. \end{aligned}$$

Since for  $G(\tau_L, s) = 0$  to hold, we must have

$$\frac{\gamma(1-\alpha) + \tau_L\alpha}{(1-\tau_L)(1-\alpha)} > 1, \quad (2.21)$$

we have shown that  $\partial G(\tau_L, s)/\partial \tau_L < 0$ .

Now, for the case when the size of the world, i.e.  $K^T = n + s$ , is constant, we show that  $\partial G(\tau_L, s)/\partial s < 0$ . We have

$$\begin{aligned} \frac{\partial G(\tau_L, s)}{\partial s} &= -\frac{[\gamma(1-\alpha) + \tau_L\alpha]}{(1-\tau_L)(1-\alpha)[(A+1-\alpha)(A-1) - (n-2)]} \\ &\quad \cdot \left[ 1 + \frac{(n-1)}{[(A+1-\alpha)(A-1) - n+2]} \left[ (2A-\alpha) \left( \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} - 1 \right) + 1 \right] \right] + \\ &\quad + \left( \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} - 1 \right) \left( 1 - \frac{\gamma(1-\alpha) + \tau_L\alpha}{(1-\tau_L)(1-\alpha)} \right) \quad (2.22) \end{aligned}$$

We see that indeed  $\partial G(\tau_L, s)/\partial s < 0$ , because (2.21) still holds.

Similarly, when we hold  $n$  constant (and let the size of the world vary as we change  $s$ ), we

obtain

$$\begin{aligned} \frac{\partial G(\tau_L, s)}{\partial s} &= -(n-1) [\gamma(1-\alpha) + \tau_L \alpha] (1-\tau_L)(1-\alpha) \left[ (2A-\alpha) \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} + 1 \right] \\ &\quad \cdot [(1-\tau_L)(1-\alpha) [(A+1-\alpha)(A-1) - (n-2)]]^{-2} + \\ &\quad + \left( \frac{1-\tau_S}{1-\tau_L} \right)^{\frac{1}{1-\alpha}} \left( 1 - \frac{\gamma(1-\alpha) + \tau_L \alpha}{(1-\tau_L)(1-\alpha)} \right), \end{aligned}$$

which is still negative.

Appendix 4: Figures

European Union Corporate Taxes as a Function of Country Size

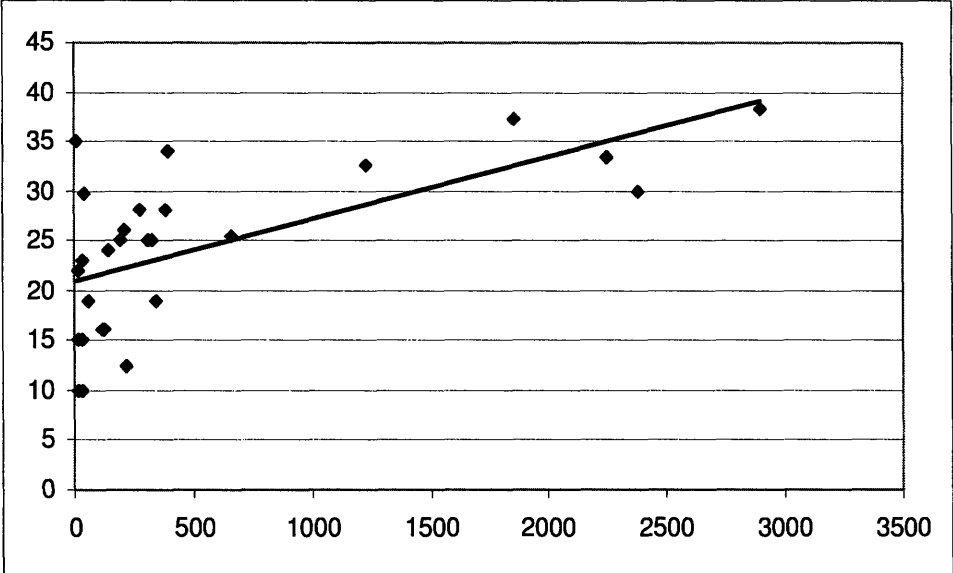


Figure 1: 2007 corporate taxes in EU 27 vs. GDP (in bil. USD)

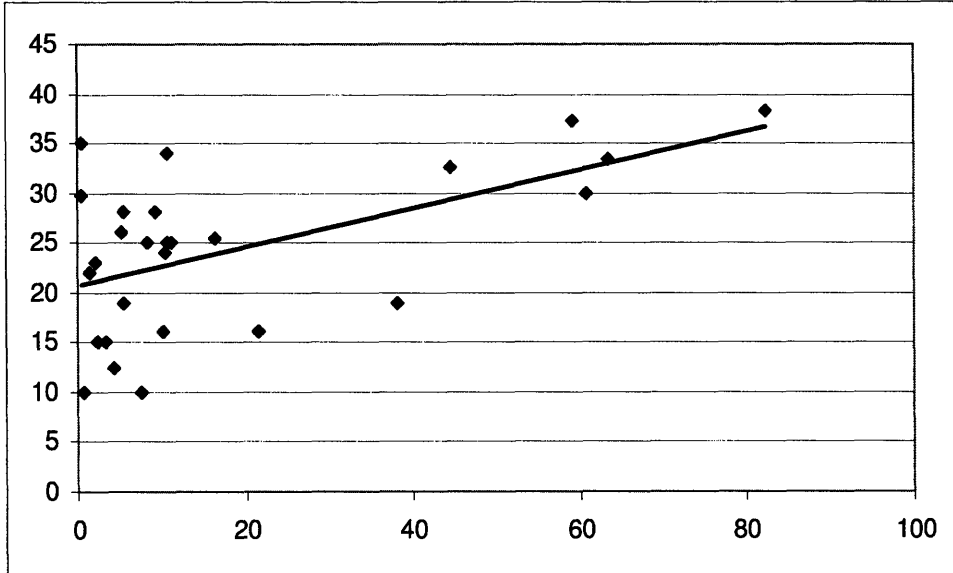


Figure 2: 2007 corporate taxes in EU 27 vs. population (mil.)

## Computations for Small Countries with Positive Size

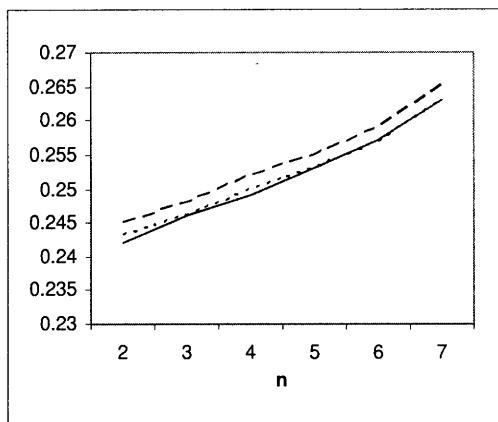


Figure 3:  $\tau_L$  when  $n + s = 7$

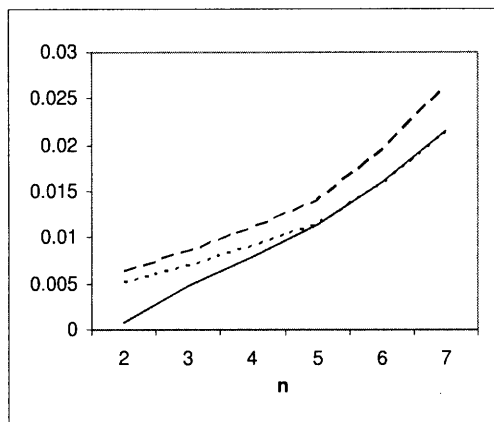


Figure 4:  $\Delta\gamma_L$  when  $n + s = 7$

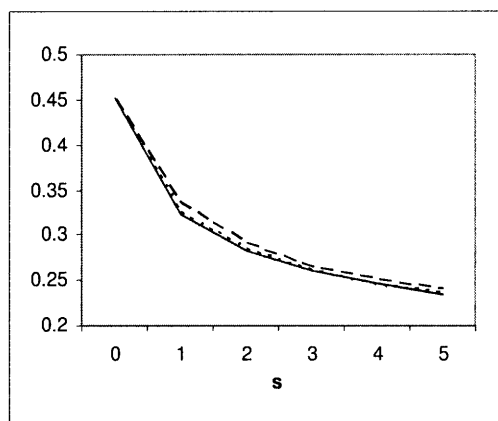


Figure 5:  $\tau_L$  when  $n = 3$

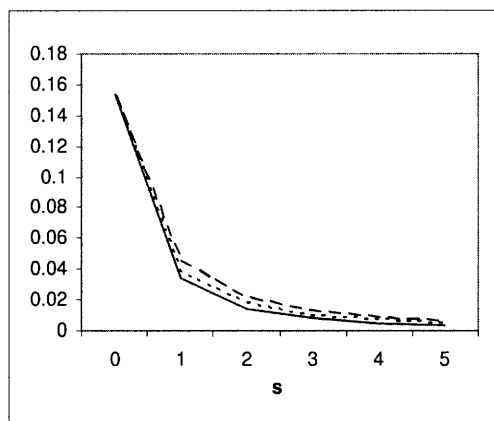


Figure 6:  $\Delta\gamma_L$  when  $n = 3$

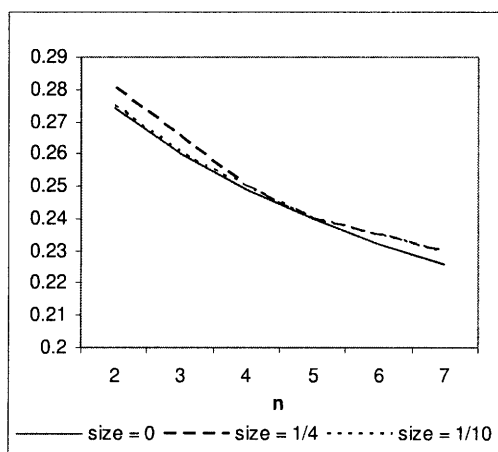


Figure 7:  $\tau_L$  when  $s = 3$

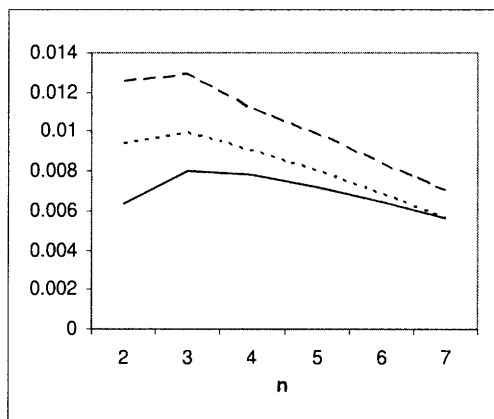


Figure 8:  $\Delta\gamma_L$  when  $s = 3$

## Computations for Median Agents with Capital Holdings

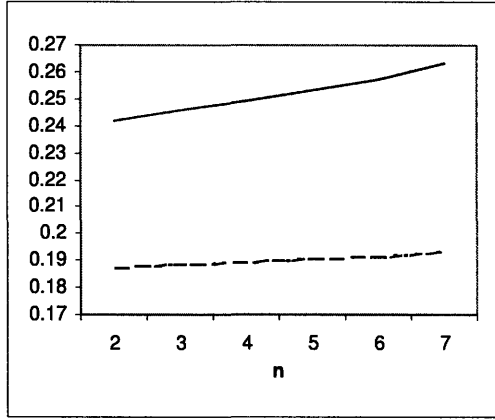


Figure 9:  $\tau_L$  when  $n + s = 7$

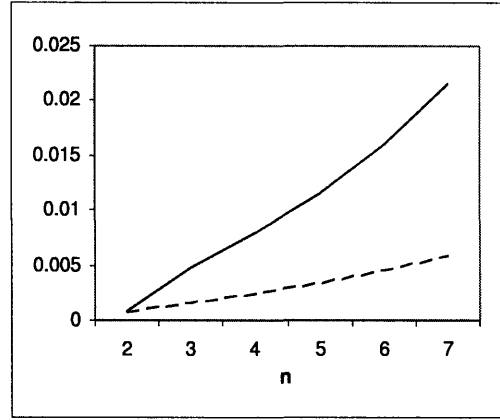


Figure 10:  $\Delta\gamma_L$  when  $n + s = 7$

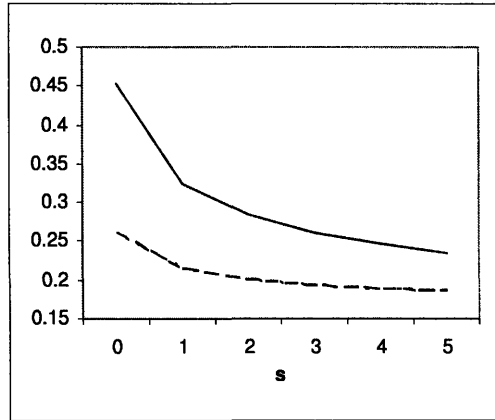


Figure 11:  $\tau_L$  when  $n = 3$

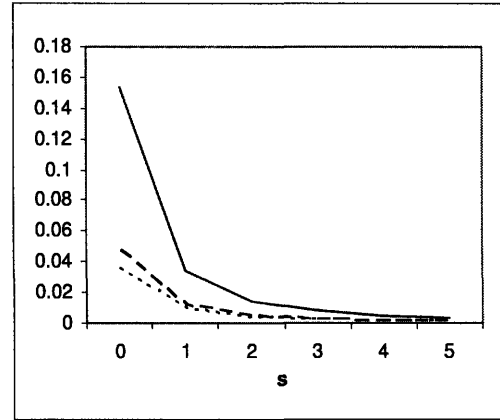


Figure 12:  $\Delta\gamma_L$  when  $n = 3$

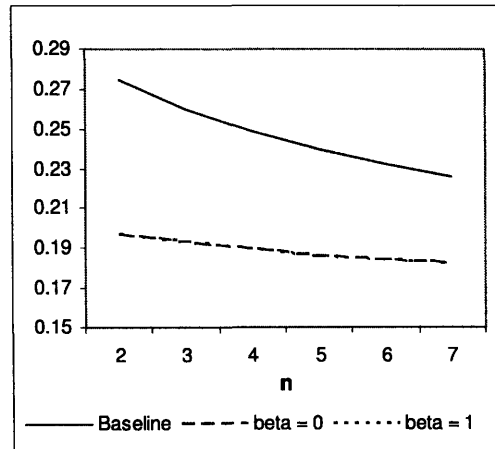


Figure 13:  $\tau_L$  when  $s = 3$

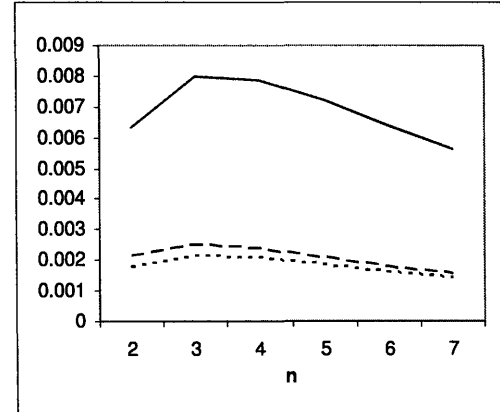


Figure 14:  $\Delta\gamma_L$  when  $s = 3$

## Chapter 3

# Twenty Years On: Educational Achievement in Post-Communist Countries

### 3.1 Introduction

Over the past almost 20 years, the Eastern European post-communist countries have taken various roads to open markets, capitalism and democracy, with varying degrees of success. Some of them are already enjoying the benefits of European Union membership, while others are still far too behind in their political and economic development.

Most of these countries have gone through turbulent two decades – economically, and quite a few politically. The public services, which were taken for granted in the at least four decades of communism, in many cases deteriorated or disappeared completely, as the countries no longer had resources for their overgenerous social programs.

However, there are areas where the original effect of communist policies may have prevailed despite the considerable amount of time that has passed since the revolutions of 1989 and thereafter. Certainly, this effect cannot persist where it would depend on the continuation of government funding - precisely the ingredient lacking in the past 20 years. Rather, it would have to be in the realm of “know-how” or “practice”. Just as in the private sphere, where

companies without foreign ownership are at a disadvantage because of the ingrained attitudes and practices developed over the long period of communism, in the public arena, one can expect to observe a rather persistent effect in education.

It may be hard to specify the precise differentiating factors of the educational system under communism and under a free and democratic government. However, in general, a free government will usually allow and support, due to its nature, a more liberal education that emphasizes independent and creative thinking. On the other hand, a communist government will stifle free thinking so necessary in the liberal arts and humanities and replace them with more repressive methods. The discipline it requires of its citizens may, however, have a positive effect on education in other areas that require hard work, such as mathematics and natural sciences.

While communism fell in the late 80s and early 90s in all Eastern European countries, it is still largely the same teachers educating the youth in the region. The teachers who themselves learnt under communism and who possibly taught pupils for years under the totalitarian regime. As this human capital is a major factor in education, there is reason to believe that the above-mentioned effects of communist policies in education still prevail in post-communist countries. That is, students there can be expected to achieve better in mathematics and natural sciences and worse in humanities than their peers from the West, controlling for other relevant factors on individual, school and national levels.

This paper intends to determine whether this is indeed the case. It employs the results of the recent 2000, 2003 and 2006 Programme for International Student Assessment (PISA) studies performed by the Organization for Economic Cooperation and Development (OECD). The PISA tested hundreds of thousands of 15-year old students in dozens of countries within OECD and elsewhere in up to four areas – mathematics, science, problem solving and reading. Therefore its results offer a perfect opportunity for testing the hypothesis of a persisting and differentiated effect of communism on the various fields of education: It is precisely in the formative years of primary and lower secondary schools, that is, before reaching the age, at which PISA conducted its tests, that the different approach and teaching methods employed in the West and in the post-communist countries will affect the youth the most.

A previous study by Olsen et al (2005) has already demonstrated, employing the PISA 2003 data, that the group of post-communist countries does indeed have common characteristics

in science achievement. Similarly, Zabulionis (2001) showed that students in post-communist countries achieved according to a similar pattern in mathematics and science, as measured by the Third International Mathematics and Science Study (TIMSS). Both papers were concerned with student responses to individual test items. Same or similar patterns between countries as to which questions were answered correctly and incorrectly would then point to related education systems and curricula in the various countries. Thus Olsen and Zabulionis have in fact shown that school systems in post-communist countries do exhibit similarities. However, they did not consider actual levels of achievement in those countries vis-a-vis achievement elsewhere, which is the goal of this paper.

This paper shows that it is indeed the case that, controlling for other factors, young students in the post-communist countries achieve better than their western counterparts in the “hard sciences” and worse in the more creative areas as represented by the reading test. It also seems that the advantage of former eastern bloc countries in sciences might now be disappearing, almost 20 years after the fall of communism. On the other hand, their lagging in the humanities persists.

The paper starts with a description of the PISA studies, the data collected and methods employed, and continues with the results of this investigation, their analysis and finally, a conclusion.

### **3.2 Description of the Data**

The OECD Technical Reports, Manuals and various other reports for the 2000, 2003 and 2006 PISA surveys provide a very good and detailed description of the studies and the methods they employed and thus have proved invaluable for this and the next section of our paper.

The 2000 PISA study was conducted in 43 countries. In 2003, the study included 41 countries and 57 countries in 2006. The studies were performed in all member states of OECD and a further 30 countries and territories, all of which are listed in Table 1.

The studies tested students in participating countries in three areas – mathematics, reading and science. The 2003 survey also included testing in problem solving. The purpose of the studies was not to test specific knowledge acquired in classes, but rather to determine a general

level of “literacy” of students in the various countries. Literacy in this case was construed as a continuous rather than binary variable, describing the ability of students to apply their general knowledge in the four areas to everyday practical problems that they may encounter in life following the completion of their compulsory education, which usually occurs around the age of testing in the PISA studies.

Let us note here that the reading tests did not simply score students on their reading ability, narrowly construed. Certainly, the reading assessment attempted to measure how well a student could retrieve information from a text. However, more importantly, it strived to examine students’ proficiency in “forming a broad general understanding”, “developing an interpretation” and “reflecting on and evaluating the content of a text” (OECD (2003)). It was specifically designed to test students’ capacity to independently and creatively combine the information contained in the text and their own outside knowledge. Therefore the reading test should be understood as more of a general assessment of abilities necessary in the fields of humanities and social science. Apart from the reading ability itself, it was concerned with the students’ interpretative skills and the capacity to independently draw conclusions based on their knowledge and on the information provided to them. It is this quality of the reading test that allows us to use the students’ reading scores as a measure of their more general capacity in the humanities and the social sciences, as well as of their independent and creative thinking.

Students tested were between 15 years 3 months and 16 years 2 months old. Thus they were tested regardless of the particular grade they were in or of a particular type of educational institution they were attending at the time. This was to find a common denominator among the significantly different educational systems of the participating countries. Only a very small percentage of schools or students were ineligible for participation in the study, for reasons such as geographical remoteness of the school or students who have had very little instruction in the language of their country (e.g. very recent immigrants). Overall, the student samples ranged from 3500 students in Iceland to over 30000 students in Mexico.

The focus of the first PISA study, administered in 2000, was reading literacy. Therefore all of the more than a quarter of a million students were tested in the subject. Also, they were examined in either mathematics or science. In 2003, the focus of the study shifted to mathematics and therefore all students were tested in that subject. Also, all students were

tested in either reading, science or problem solving. Finally, the 2006 study investigated more closely literacy in science.

Apart from the test itself, students answered a range of questions in a comprehensive survey about their social and economic background as well as their impressions from their school regarding their teachers, curriculum, atmosphere, etc. Similarly, the principals of the participating institutions responded to a number of questions about the material and educational resources available at their respective institutions, the quality of the teachers, the morale of both teachers and students and the general prevailing atmosphere.

In the end, the PISA studies compiled the responses of school principals and students from the 60 participating countries in a publicly available database, which formed the basic data source for this paper. While the basic data and test results are available for all students, some information is missing for many individuals and schools.

However, as PISA collected no financial data (e.g. concerning the funding of the individual schools), further data regarding educational spending in the participating countries was collected from the OECD, the World Bank World Development Indicators and the databases of Eurostat. The information on public expenditure in the education sector was available for all participating countries. Some post-communist countries did not have data available on private expenditure on education, which in their case is usually rather small in comparison to the public expenditure (as opposed to, say, the US, where private spending on education is significant) and thus this lack of data did not significantly affect further analysis. A larger number of countries did not have data available on spending on primary education alone, which is the figure most relevant to the analysis in this paper.

### **3.3 Study Design and Data Processing**

#### **3.3.1 Sample Weights**

Since the focus of the PISA studies was on student performance and therefore on the student sample, the studies were designed so that student weights in the survey would be similar. Ideally, the students in the population would be sampled directly and randomly. However, this would not be practically very feasible, as some schools would only have to test one or two students

and thus the survey would prove too costly, with far too many schools participating. Moreover, it would be difficult to control for school-level variables when examining the performance of the various students surveyed by PISA.

Therefore the surveys first sampled schools and subsequently selected 35 students in each school. If the schools were selected randomly, every school would have the same selection probability and therefore the same weight in the survey. However, given the different sizes of schools, the students' weights would vary widely: If a student's school was just as likely to be chosen as any other, but it would be a very small school, then that student would be much more likely to be selected for a PISA survey than a student at a larger school. Then the students' weights, which are just the inverses of the probability of selection, would also differ significantly.

To prevent this from occurring, schools were selected with probabilities proportional to their size (PPS). In order to maintain the sum of weights of all schools involved in the survey close to the number of schools in the population, schools were selected systematically. Schools were ordered according to their size and students were ordered within the schools so that every student had an ordering number. Then a sampling interval,  $s$ , was determined as the ratio of all students in the population and the number of schools to be involved in PISA. A random number was selected in the first sampling interval,  $r \in [0, s]$ , and the school that included the student with that number was chosen. The next selected school included student number  $r + s$  and so on.

Even though students and schools were selected in this manner from their respective populations in the PISA surveys, they were not equally representative of the whole student population for several reasons: some regions or sub-samples were intentionally over-represented, such as small but important regions that had to be analyzed on a national level or students with a non-official language of instruction. Similarly, students from e.g. remote areas were often under-represented. Also, the initial information regarding school size was not always correct which led in some cases to inclusion of a higher than expected proportion of students at a particular school to achieve the target number of 35 students per school. Certain schools or students did not respond to the survey and were not properly replaced in the study with their population equivalents.

For such and other reasons individual students in the study had to be assigned different

weights, taking into account all these separate factors. The final weight for each student in PISA was calculated as a product of several intermediate weights: the school base weight, which is the reciprocal of the probability of a school's inclusion in the sample; the within-school base weight, which is the reciprocal of the probability of a particular student's inclusion within a selected school; school and student non-response adjustments, which account for non-participation by schools and students in the survey that was not compensated for by replacement schools or students; a grade non-response adjustment, which is applied to schools where only students in the modal grade for 15-year olds were tested; and finally, school and student trimming factors, which were used to reduce unexpectedly large values of the school and student base weights.

### **3.3.2 Replicate Weights**

In a sample that is a simple random draw from a population, it is easy to calculate the sampling variance of a statistic of interest. Namely, since we shall be interested in the variance of linear regression estimates, the variance of those estimates can be obtained from a simple closed-form formula derived in the weighted least-squares model.

However, this standard computation of the standard error of a regression estimate would be faulty in the case of an educational study such as PISA. International surveys rarely sample students by directly randomly selecting students across the various countries. Rather, they delegate to national authorities to select schools, at which the survey will be administered, and subsequently at those schools students, who will participate in the study.

Even if the schools and students within those schools were selected at random, those students cannot be regarded as independent observations for the purposes of calculating sampling variance of the various sample statistics. This is due to the rather small variance in characteristics of students, who attend a particular school, relative to the variance between students at large in a given country. Those who study at the same school will usually be exposed to a similar curriculum taught by the same teachers. Only certain subjects or programs may be available at a school, thus educating all the students in a similar way and causing only students interested in such an education to attend the school in the first place. Perhaps even more importantly, most schools draw their student bodies based on geographical criteria - usually, the location of a student's residence. Populations tend to cluster within cities, countries, etc.

according to their socio-economic status. Therefore, at a given school, students of similar social and economic background are often to be found. Even when a particular student does not belong to the prevalent class at his school, he will be exposed to a different environment, including different resources, etc. than were he to study elsewhere.

Therefore where a random simple sample of  $m$  students is expected to cover the diversity of the population to a certain degree, a random sample of  $s$  schools and, within those, a sample of  $n$  students, such that  $m = s * n$ , will certainly represent less diversity within the population at large. This in turn implies that the variance and standard errors for any statistic based on a two-stage sample will be higher than those calculated by standard means based on a one-stage simple random sample from the population.

It turns out that simple calculation of standard errors exists for two-stage samples, provided that the sample has the following properties: Both stages have infinite populations, i.e. in our case countries have infinitely many schools and each school has infinitely many students. Then a random sample of the primary sample units (PSU) - schools - needs to be selected and within each PSU, a random sample of stage two units - students - is chosen. However, in our case the above conditions clearly do not hold: Not only the populations are finite. School sizes also vary. As mentioned earlier, schools are not selected purely randomly. Rather, they are chosen systematically and taking into account their size. Finally, special considerations are taken when selecting schools to partake in PISA, thus taking the selection process even further from a random sample.

Therefore we use computational replication methods to obtain standard errors of our statistics of interest. Specifically, a modification of the Jackknife method is employed. The Jackknife generates  $n$  replicate samples of size  $n - 1$  from an original sample of size  $n$ . For each replicate sample, it drops one of the observations from the full sample. Then, employing the squared differences between the statistic of interest calculated from the original sample and all the replicate samples, it allows for the calculation of the sampling variance of that statistic. For a simple random sample this becomes  $\sigma^2 = \frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_i - \hat{\theta})^2$  (where  $\hat{\theta}$  is the estimate of the statistic  $\theta$  based on the entire sample and  $\hat{\theta}_i$  is the estimate from the  $i^{th}$  replicate sample) and it can be shown mathematically that this is equivalent to the standard formula derived in a weighted least squares model when one deals with a simple random sample.

In the case of a stratified two-stage sample design, such as the one employed in the PISA surveys, the creation of replicate samples changes somewhat. Here, schools, as mentioned above, are selected systematically according to their size. But they are also divided into various strata, so that the final sample includes as much of the diversity of the population as possible. This increase in the diversity captured by the survey then reduces the sampling variance of statistics of interest. The stratification variables are those that affect what PISA strives to measure, i.e. academic performance. They are therefore school parameters such as public versus private, rural versus urban or vocational versus academic.

The Jackknife method then pairs together schools that are as similar as possible: First, they are in the same stratum (e.g. they are both rural schools) and then they are the schools within that stratum closest to each other in size. If the total number of PSUs is  $n$ , the Jackknife then, for each replicate sample, includes  $n/2 - 1$  pairs of schools with a weight 1 and only one school of the remaining  $(n/2)^{th}$  pair (randomly selected from between the two) with a weight 2. Technically, the other school in the  $(n/2)^{th}$  pair is assigned a weight 0. Then, if  $G$  is the number of schools in the original sample, the sampling variance for statistic  $\theta$  becomes 
$$\sigma^2 = \sum_{i=1}^G (\hat{\theta}_i - \hat{\theta})^2.$$

The problem with the basic Jackknife method is that it creates  $n/2$  replicate samples. With the number of schools included in the PISA surveys around 8000, that would mean running each linear regression, in whose estimates we will be interested, 4000 times. Here, the Balanced Repeated Replication (BRR) method provides a computationally more feasible alternative. This method removes, at random, one school out of each of the  $n/2$  pairs. Therefore the weight of the removed school becomes 0, whereas the remaining school in each pair has a weight of 2. The BRR method will provide a far greater diversity in replicate samples for a given number of samples than the Jackknife method.

However, the BRR method obviously removes half of available observations when creating its replicate samples. This is clearly a significant reduction in the number of observations relative to the complete sample. Therefore we use the Fay variant of the BRR method, whereby the dropped school from each pair is not actually weighed by 0, but rather by a factor  $k$  such that  $k \in (0, 1)$ . Then the weight placed on the remaining school in each pair is simply  $2 - k$ .

The sampling variance of a statistic of interest when employing the Fay method is  $\sigma^2 =$

$\frac{1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_i - \hat{\theta})^2$ . PISA employs Fay value of 0.5 and generates 80 replicate samples. The sample variance for any statistic then becomes  $\sigma^2 = \frac{1}{20} \sum_{i=1}^{80} (\hat{\theta}_i - \hat{\theta})^2$ .

The Fay's replicates then produce the desired unbiased and consistent estimates of standard errors for the linear estimators that are later employed in this paper, and thus avoid the risk of obtaining statistical significance due to underestimation of standard error.

### 3.3.3 Plausible Values

To scale the PISA test responses by students, a combination of a mixed coefficients multinomial logit model and a population model was used. The population model assumed that students have been sampled from a multivariate normal distribution.

The model calibration was performed first on a national, country-by-country level. The goal was to obtain for each of the subjects tested (i.e. the different dimensions of the logit model) a continuous scale that would describe the "literacy" of each student in each of the subjects tested in the study. The scores were not only to describe the number of test items answered correctly by each individual, but also to reflect their level of difficulty, which was variable between countries and determined by the proportion of students, who answered a particular item correctly.

Once the parameters of the distribution of scores based on the underlying student characteristics (i.e. their item responses) was determined through iteration, a similar procedure was followed to obtain the international parameters, which allowed the creation of a single scale, centered at 500 for OECD countries.

Then, in order to obtain consistent estimators for the population of students, plausible values rather than weighted likelihood estimators were used. The latter would be preferable if the purpose of PISA were to ascertain the knowledge of particular students as they would report the achievement of those students. However, PISA strives to survey and deduce inferences regarding an entire population of students across various countries. Therefore being able to consider individual performance and its relation to other factors is not as important.

Plausible values are thus used in international surveys such as PISA, since they are better at describing the performance of the entire population. They are random draws from posterior

distributions of student performance around the actual score achieved by each student. While a particular student scores in a certain way when tested by PISA, this does not necessarily best represent his true ability and achievement in the subject. An example may be a student whose ability is better than can be demonstrated by the test: He would achieve a full score on a test even more difficult than the one administered. Therefore PISA estimates the posterior distributions around each actual score that assign probabilities to all possible student abilities based on that score. Then it draws five vectors of plausible values from these posterior distributions.

These five values allow for better estimation of standard errors for statistics of interest than the two other options: Either using the original weighted likelihood estimates (WLE) of student scores or the expected a posteriori estimator (EAP), which is simply the mean of a posterior distribution around a student's score, produces the same or insignificantly different estimates of statistics of interest. However, the variance of those statistics will be overestimated when using WLEs and underestimated when employing EAP.

Therefore in the following analysis, parameters were estimated using all five vectors and averaged. Even though one set of plausible values alone would provide for unbiased estimates of the population, using five vectors of plausible values will insure obtaining the correct values of their standard errors. Therefore we estimate the linear regressions, in which we shall be interested, five times, once for each vector of plausible values. We then calculate each estimate value as the mean of the five various estimates. The final error variance of our estimates is going to combine the measurement error variance, or imputation variance, and the sampling variance.

The imputation variance stems from the difference between the five plausible values in each vector. If we were dealing with a perfect test that would be able to correctly capture the ability of students that it is trying to measure, there would be only one ability associated with a given score achieved in the posterior distribution for that score. All other abilities would have zero probability in the posterior distribution. But that would mean that all draws of plausible values would be identical. Therefore the difference between the various draws captures the imprecision of the PISA survey in capturing the actual ability of students.

The measurement variance is then equal to  $\sigma_M^2 = \frac{1}{p-1} \sum_{i=1}^p (\hat{\theta}_i - \hat{\theta})^2$ , where  $\hat{\theta}_i$  is the estimate of the statistic of interest based on vector  $i$  of plausible values,  $i \in \{1, \dots, p\}$ ,  $\hat{\theta}$  is the mean of

those estimates and therefore the estimate we are trying to obtain,  $\hat{\theta} = \frac{1}{p} \sum_{i=1}^p \hat{\theta}_i$ , and  $p$  is the number of vectors of plausible values. We obtain the sampling variance as the average of the variances calculated using the Fay's variant of the BRR method, i.e.  $\sigma_S^2 = \frac{1}{p} \sum_{i=1}^p \sigma_i^2$ . Then final error variance is

$$\sigma^2 = \sigma_S^2 + \left(1 + \frac{1}{p}\right) \sigma_M^2 = \frac{1}{p} \sum_{i=1}^p \sigma_i^2 + \frac{1}{p-1} \sum_{i=1}^p (\hat{\theta}_i - \hat{\theta})^2.$$

### 3.3.4 Indices

Finally, the raw data from the student and school questionnaires had to be turned into a form that could be conveniently used in the further analysis of the PISA data. For many question groups, the answers were summarized into several indices. Each index was scaled using a weighted maximum likelihood estimate, employing a one-parameter item response model. For every index, item parameters were first estimated from calibration sub-samples of 500 students for student indices and 99 schools for school indices, for each country. Then estimates were calculated for students and schools employing the previously obtained parameters. Lastly, the indices were standardized so that the mean of each index would be zero and standard deviation one for OECD countries. For all indices, more positive values mean more positive valuation of the characteristics in question.

The indices used in the following analysis include the index of economic, social and cultural status, which is composed of the highest level of education of the father or mother converted into years of schooling, the number of books at home as well as access to home educational and cultural resources (through responses to questions regarding several household items) and the highest international socio-economic index of occupational status of parents (which captures the attributes of occupations that convert parents' education into income). This index thus attempts to capture the socio-economic status of each student through their parents' occupation, education, and indirect measures of their relative wealth.

Also, several school-level indices were employed, based on principals' responses to the school questionnaire. The index of the quality of the school's educational resources captured the availability of instructional materials, computers, software, calculators, library materials, audio-visual resources and science laboratory equipment. The student behavior or school discipline

index included information regarding student absenteeism, disruption of classes by students, lack of respect for teachers by the students, student alcohol and drug abuse, etc.

### 3.4 The Model

This paper attempts to answer the question whether post-communist countries exhibit better or worse results in their educational systems than their counterparts without the experience of a communist regime in their past. This question is addressed with a simple weighted least squares model, employing the data from the PISA 2000, 2003 and 2006 studies and elsewhere, and using the plausible values and Fay's replicates techniques described above.

The variable of interest is a dummy, which takes the value of one for all post-communist countries in the sample and the value of zero for all other countries that participated in the survey. The regressions are to determine a potential effect of this dummy variable on students' test scores, controlling for other available variables that may affect their performance.

On the individual level, there are several factors that may be expected to affect a student's learning ability and thus their score on the PISA tests. First, the sex of a student may be relevant, as especially female students can be often at a significant disadvantage to their male counterparts. Also, non-native (i.e. immigrant) students will have a harder time at school due to potential language and culture barriers in their new country. Finally, and perhaps most importantly, the socio-economic status of a student as captured in this case by the index of social, economic and cultural status or by the highest parental occupational status, where the former is not available, will almost certainly play a major role in a student's achievement at school.

On the school level, it is necessary to consider the socio-economic composition of the student body, as represented by the average value of the social, economic and cultural status index for the school: Even worse-off students will benefit from well-off classmates' presence at school and in classes, as these are likely to do better due to their own status and thus have positive spillover effects on the students with a lower status.

Finally, on the national level, the potential effect of "post-communism" on educational achievement has to be controlled for educational spending in the country, which will definitely

affect the results observed in the educational system. This can be done in two ways: Spending can be expressed in absolute terms (e.g. in current USD) per student. Even though this statistic is perfectly comparable across countries and widely used, it may be quite inappropriate, as a large portion of educational expenses, such as teachers' salaries and building maintenance, is always highly correlated with the GDP of the country in question. Therefore spending can also be expressed as the total expenditure relative to the country's GDP.

An alternative way of controlling for educational spending, thanks to data available in the PISA survey, is the index of every school's educational resources, described above. It indirectly captures a significant portion of spending at a particular school, and thus actually allows for a more precise school-level rather than national-level control for educational expenditures in a country.

### **3.5 Results and Analysis**

In this section, regression results for the mathematics scores are presented for the model as described here, and accounting in turn for educational spending expressed in USD per student, relative to the country's GDP and as proxied by the school's quality of educational resources. For the other subjects, only results for model 3 are presented here, partly since the most observations are available for that model: The data on educational spending are not available for some of the countries that participated in the PISA studies, whereas the index of schools' quality of educational resources is already included in the PISA dataset.

#### **3.5.1 Mathematics**

The regression results with the mathematics achievement as the dependent variable are qualitatively very similar across the three PISA studies. Tables 2, 3 and 4 show that all the versions of the above model yield similar results for the mathematics scores on the PISA tests.

First, the sex of students is very statistically significant in all of the cases. We obtain a positive coefficient on the male dummy variable, meaning that male students achieve significantly better in mathematics than female ones. This is hardly surprising, as whatever the underlying reasons may be, boys in general indeed do better in mathematics than girls. Also quite expect-

edly, when a student is native to their country, he scores significantly higher on the test. He presumably does not have to expend as much energy trying to adapt to a foreign culture, learn a new language, etc.

When a student's socio-economic status is higher, he achieves significantly better on the mathematics test. The most important effect here seems to be that of the socio-economic composition of a student's peers at school: when a student's classmates' average socio-economic status is higher, the student on average scores significantly higher on their test. This impact is up to five times more important than a student's own status. A student from a lower-status background sees his chances for academic success much improved when he is educated in a generally higher-status environment. His classmates' parents presumably place more emphasis on education, which influences their attitude towards school and the atmosphere therein, thus creating an environment more conducive to achieving good results. Note that the 2000 PISA did not produce an index of socio-economic and cultural status of students centered at zero and with a standard deviation of one, such as the one calculated for the 2003 and 2006 PISA studies. Therefore we instead use the highest parental occupational status index, which, however, is on a 0-100 scale. Therefore the constant and the estimates on the socio-economic status produced by the 2000 regressions are significantly lower than their 2003 and 2006 equivalents.

Not surprisingly, whatever measure of educational spending is employed, whether it be in absolute terms per student in the primary education system of a given country, relative to the country's GDP per capita or expressed through the quality of a school's educational resources, its effect upon students' achievement in the mathematics test is significantly positive.

Finally, the coefficient of interest on the dummy for post-communist countries also always takes a positive value, which is statistically significant. It varies with the specific control variables employed. Its value is highest when the control variable for educational spending is the relative nation-wide primary education expenditure. Here, we can make the interesting observation that, for example in 2006, a country being post-communist had the same effect on students' mathematics performance as the government of a Western country spending an additional almost 16,000 USD<sup>1</sup> or almost 15% of its GDP per capita per student.

On the other hand, the coefficient on the post-communism dummy is at its lowest when

---

<sup>1</sup>2000 US Dollars.

the control variable is least likely to fully control for spending: The index of school's quality of educational resources does not capture, for instance, the salaries of teachers, which form a major part of educational spending and certainly affect the quality of instruction. Unfortunately, other supplementary variables such as student/teacher ratios that would aide the analysis here could not be used as the number of observations for them was too low. Once again though, the coefficient on our dummy is positive even when we are not fully controlling for country-wide spending. Since communist countries are generally poorer and thus spend less, this further assures us that post-communism does have a positive impact on achievement in mathematics.

Regressions other than those presented here were run, employing different control variables, including indices on teacher and student morale as reported by school principals, data on spending on primary and lower secondary education rather than simple primary education, and others. For all those regressions, the coefficients on variables were similar to those presented in Tables 2, 3 and 4 and the coefficient on the post-communist dummy was always significantly positive. Clearly, the way of thinking the communist regimes instilled in its citizens in schools and therefore current teachers, who grew up and were educated themselves under the regime, further instill in their students, has a very strong positive impact on students' mathematical abilities. While they may have stifled independent and creative thinking, they certainly encouraged traits necessary for high achievement in mathematics.

One potential problem that must be addressed in this regression is over-controlling. Post-communist countries have lower educational expenditure across all of its measures. This is to be expected: post-communist countries are in general in a worse economic shape than their mainly Western counterparts and thus spend less on education. Since the post-communist variable causes less spending which in turn causes lower test results, this dilutes the direct positive effect of a post-communist status on the mathematics scores of students. It is necessary to control for educational spending even despite this obvious problem. However, as the positive effect of "post-communism" can be established even in the presence of over-controlling, it is reasonable to believe that the actual effect (i.e. one that could be measured if the post-communist status of a country did not likely imply a lower level of educational spending) is indeed even higher than the one observed in the above regressions.

It is interesting to note that over-controlling is not a problem for the socio-economic index

variable. As that variable is independent of any nominal monetary value, but rather measures the education, occupational status and cultural possessions of students' parents, its values are not lower for post-communist countries. Apparently, these countries were able to achieve a similar level of "socio-economic" achievement as measured by the index as the other countries included in the study.

### 3.5.2 Other Subjects

Only one regression is presented for each of the subject areas of reading, science (and problem solving for PISA 2003) in Tables 5, 6 and 7, as once again, results of other variants of the basic model are very similar and have no bearing on the sign and significance of the variable of interest.

There are some important similarities and differences we can observe with the results we obtained for the mathematics regressions. First, male students still achieve better in science and in problem solving, albeit the difference is quite a bit smaller than in mathematics. However, in reading the difference in scores by gender reverses. It is the girls who test better in this area. This could have been expected, because girls generally perform better in the humanities and the reading achievement test on the PISA represents that academic area.

Socio-economic status of individual students and of their classmates at school has a similar positive impact on their performance in any area tested and in all years when PISA was administered. Similarly, the quality of schools' educational resources, which here captures spending on education, has a positive effect on students' achievement across fields and surveys.

Positive values on the post-communism dummy are reported for science and problem solving (in 2003) test results. These are in the same "hard science" category as mathematics and it is here that the educational systems of post-communist countries seem to over-achieve relative to their Western counterparts.

However, we notice an important caveat: The coefficient on the post-communism dummy is no longer significantly positive (it is actually not significantly different from zero) for science achievement in 2006. Whereas the advantage of post-communist countries in science appears similar to that they possess in mathematics in 2000 and 2003, it disappears in the latest PISA study. Given the significance level of all the results in our other regressions, this change is

rather meaningful: It appears that post-communist countries might be finally losing their edge in the science area, almost 20 years after the fall of the communist regimes in Eastern Europe. One might then expect a similar development to follow in the area of mathematics. It will certainly be very interesting to continue this type of investigation once data is available from the subsequent PISA studies.

Importantly, and certainly of interest to us, the coefficient on the dummy is significantly negative for the reading achievement regressions. Reading achievement of students in post-communist countries is worse than their peers in the West by a wide margin. The features of the communist regimes that cause higher achievement in mathematics and science are the same that decrease students' abilities in the humanities and social sciences: The oppressive regime discourages independent and creative thinking, necessary for succeeding in the humanities and social sciences. PISA does not test individual subjects, such as history or philosophy, in those fields, since they would require specific knowledge. Rather, the surveys are only interested in literacy in a field. For this reason, and as explained earlier, the reading tests are the ones that demonstrate students' abilities closely related to their capacity to achieve in the humanities.

We note that while the advantage that post-communist countries carry in mathematics and science might be slowly disappearing, even 20 years after the fall of communism their disadvantage in reading is not diminishing in the least. While the strengths of the educational system in the former eastern bloc are vanishing simply due to "attrition", the countries will almost certainly need to seriously reform their schools in order to close in on the Western countries in the area of humanities and social sciences.

The regressions for reading achievement are the only ones where over-controlling could pose an actual problem. However, as the absolute value of the coefficient is indeed rather large, and certainly very different from the coefficients observed for other subjects, it is very likely that the fact that a country is post-communist has still a direct negative effect on achievement of students on the reading test part of the PISA survey. Nevertheless, to assure of this result, it would be necessary to perform an analysis of how much of the lower educational expenditures in the post-communist countries can actually be explained by their political past rather than other factors (such as lower stage of development before the adoption of communism, etc). This kind of analysis is beyond the scope of this paper.

### 3.5.3 Impact of the Fay's Replicates Method

In Table 8 we present an example of the importance of employing the Fay's variant of the Balanced Repeated Replication method, rather than calculating standard errors for our estimates as if the students selected for the PISA studies represented a simple random sample. We show the estimates of the linear regression parameters and standard errors calculated using both methods for the science achievement in the 2006 PISA study. The standard method of calculation clearly underestimates the sampling variance.

We concluded above, based on Table 7, that the fact that a country is post-communist had neither a significant negative nor positive impact on the abilities of its students in science in 2006. However, if we were to have calculated the standard error for the statistic of our interest, i.e. the coefficient on the post-communism dummy, in the standard way, we would have concluded that the students from the eastern bloc countries actually fared worse in science than their Western counterparts. This is simply because the standard method, as we noted earlier, really does underestimate the sampling variance for any statistic of interest calculated from the PISA data.

## 3.6 Conclusion

This paper finds that, controlling for various variables including educational spending in a given country, 15-year olds from post-communist countries achieved better in the mathematics, science and problem solving subject areas of the PISA 2000, 2003 and 2006 surveys, whereas they did worse than students from other participating countries in reading. This implies that at the very least the primary and early secondary education in natural sciences does better in post-communist countries whereas it very likely does worse in areas requiring more creative thinking and interpretation abilities (required in the reading test). This difference apparently persists to a large extent even almost 20 years after the fall of communism in most of the surveyed countries, likely due to the slow changes in the human capital in education.

However, we have found that the advantage of post-communist countries in science achievement has disappeared by the latest PISA survey. We could conjecture that the same will happen to their edge in mathematics achievement. However, their lagging in reading performance by

their students remains at the same level throughout the years covered by the PISA surveys. Therefore it is possible that the positives of the educational systems founded under communist rule have started to evaporate 20 years after the fall of the regime - which is hardly surprising. At the same time, though, the negatives have not been reduced and would probably require a thorough school reform to alleviate.

A further detailed study would be necessary to find the exact causes of the effect that post-communism still seems to exert on students' achievement. Simple surveys of school principals or teachers as the ones conducted within the PISA studies do not suffice, as their results are not necessarily comparable across countries. A more objective study could ascertain whether the differing results are due to higher levels of discipline required by school administrators, teacher quality, or simply teacher attitudes in schools in post-communist countries versus elsewhere. Teacher technique in particular could very well explain the differences in achievement as different approaches are likely best suited for instruction in natural sciences versus in social sciences and humanities. It seems possible, however, that the same "liberal" approach being used in all instruction in non-post-communist countries may be responsible for their better results in reading and worse results in the natural sciences. On the other hand, the more "conservative" approach, which includes memorization and traditional instruction, often applied across the board in post-communist countries as part of their political legacy, may be beneficial to the instruction of natural sciences.

This paper does not attempt in any way to determine whether the educational system, on primary or any other level, is more successful in the post-communist countries or elsewhere. However, it does clearly suggest that a comprehensive study of teaching approaches employed in different subjects in the various countries could potentially lead to improvements in instruction in both groups of countries.

### 3.7 Appendix: Tables

**Table 1: Countries that participated in PISA**

(post communist countries in bold)

<b>OECD member countries</b>	<b>2000</b>	<b>2003</b>	<b>2006</b>	<b>Partner countries</b>	<b>2000</b>	<b>2003</b>	<b>2006</b>
Australia	x	x	x	<b>Albania</b>	x		
Austria	x	x	x	Argentina	x		x
Belgium	x	x	x	<b>Azerbaijan</b>			x
Canada	x	x	x	Brazil	x	x	x
<b>Czech Republic</b>	x	x	x	<b>Bulgaria</b>	x		x
Denmark	x	x	x	Chile	x		x
Finland	x	x	x	Colombia			x
France	x	x	x	<b>Croatia</b>			x
Germany	x	x	x	<b>Estonia</b>			x
Greece	x	x	x	Hong Kong	x	x	x
<b>Hungary</b>	x	x	x	Indonesia	x	x	x
Iceland	x	x	x	Israel	x		x
Ireland	x	x	x	Jordan			x
Italy	x	x	x	<b>Kyrgyz Republic</b>			x
Japan	x	x	x	<b>Latvia</b>	x	x	x
Korea	x	x	x	Liechtenstein	x	x	x
Luxembourg	x	x	x	<b>Lithuania</b>			x
Mexico	x	x	x	Macao		x	x
Netherlands	x	x	x	<b>Macedonia</b>	x		
New Zealand	x	x	x	<b>Montenegro</b>		x	x
Norway	x	x	x	Peru	x		
<b>Poland</b>	x	x	x	Qatar			x
Portugal	x	x	x	<b>Romania</b>	x		x
<b>Slovak Republic</b>		x	x	<b>Russian Federation</b>	x	x	x
Spain	x	x	x	<b>Serbia</b>		x	x
Sweden	x	x	x	<b>Slovenia</b>			x
Switzerland	x	x	x	Taiwan			x
Turkey		x	x	Thailand	x	x	x
United Kingdom	x	x	x	Tunisia		x	x
United States	x	x	x	Uruguay		x	x

**Table 2: PISA 2000: Effect of post-communism and other factors on student achievement in mathematics**

(Coefficients significant at the 5% level in bold)

	Model 1		Model 2		Model 3	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>						
Constant	<b>196.82289</b>	<b>5.26007</b>	<b>143.48898</b>	<b>5.75225</b>	<b>170.21718</b>	<b>12.30351</b>
Student is male	<b>21.92606</b>	<b>1.22497</b>	<b>21.39342</b>	<b>1.24068</b>	<b>21.67682</b>	<b>2.18628</b>
Student is native	<b>25.04103</b>	<b>2.88681</b>	<b>22.67323</b>	<b>2.88311</b>	<b>17.33115</b>	<b>3.69651</b>
Index of socio-economic status	<b>0.70423</b>	<b>0.03977</b>	<b>0.70307</b>	<b>0.03909</b>	<b>0.62674</b>	<b>0.07608</b>
<b>School factors</b>						
Average socio-economic status	<b>3.97290</b>	<b>0.11367</b>	<b>4.16494</b>	<b>0.12447</b>	<b>5.39392</b>	<b>0.25810</b>
<b>Educational spending</b>						
In USD per student, primary education	<b>0.01233</b>	<b>0.00053</b>				
As a percentage of GDP per capita, primary education			<b>5.73807</b>	<b>0.19834</b>		
Quality of schools' educational res.					<b>12.51323</b>	<b>2.50007</b>
<b>Country is post-communist</b>	<b>42.21299</b>	<b>2.09706</b>	<b>10.02432</b>	<b>1.96660</b>	<b>15.94476</b>	<b>2.51570</b>
R-squared	0.4861		0.3791		0.3095	
Number of observations	56848		56848		82887	

**Table 3: PISA 2003: Effect of post-communism and other factors on student achievement in mathematics**

(Coefficients significant at the 5% level in bold)

	Model 1		Model 2		Model 3	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>						
Constant	<b>444.29293</b>	<b>3.52162</b>	<b>416.49701</b>	<b>4.58388</b>	<b>459.65205</b>	<b>2.63797</b>
Student is male	<b>13.53783</b>	<b>1.01774</b>	<b>13.46627</b>	<b>1.00756</b>	<b>13.78901</b>	<b>1.06908</b>
Student is native	<b>24.51616</b>	<b>2.98432</b>	<b>25.82308</b>	<b>2.96728</b>	<b>21.28131</b>	<b>2.62089</b>
Index of socio-economic status	<b>14.75622</b>	<b>0.54534</b>	<b>14.72094</b>	<b>0.54718</b>	<b>15.17866</b>	<b>0.55314</b>
<b>School factors</b>						
Average socio-economic status	<b>66.10911</b>	<b>1.69213</b>	<b>63.15828</b>	<b>1.57087</b>	<b>70.34529</b>	<b>1.40374</b>
<b>Educational spending</b>						
In USD per student, primary education	<b>0.00296</b>	<b>0.00040</b>				
As a percentage of GDP per capita, primary education			<b>2.28752</b>	<b>0.19670</b>		
Quality of schools' educational res.					<b>3.75376</b>	<b>1.01936</b>
<b>Country is post-communist</b>	<b>26.38328</b>	<b>2.14600</b>	<b>9.87281</b>	<b>1.53437</b>	<b>6.20616</b>	<b>2.95836</b>
R-squared	0.4312		0.4382		0.4111	
Number of observations	230930		230930		253441	

**Table 4: PISA 2006: Effect of post-communism and other factors on student achievement in mathematics**

(Coefficients significant at the 5% level in bold)

	<b>Model 1</b>		<b>Model 2</b>		<b>Model 3</b>	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>						
Constant	<b>447.70284</b>	<b>3.11133</b>	<b>430.33482</b>	<b>4.64777</b>	<b>457.52275</b>	<b>2.20776</b>
Student is male	<b>14.52186</b>	<b>0.92833</b>	<b>14.56802</b>	<b>0.92831</b>	<b>15.22983</b>	<b>0.99756</b>
Student is native	<b>21.99379</b>	<b>2.56449</b>	<b>21.79387</b>	<b>2.58573</b>	<b>19.10361</b>	<b>2.12014</b>
Index of socio-economic status	<b>14.62980</b>	<b>0.37773</b>	<b>14.63443</b>	<b>0.37786</b>	<b>13.49869</b>	<b>0.37243</b>
<b>School factors</b>						
Average socio-economic status	<b>56.30178</b>	<b>1.45986</b>	<b>56.57479</b>	<b>1.18500</b>	<b>52.36244</b>	<b>1.01975</b>
<b>Educational spending</b>						
In USD per student, primary education	<b>0.00155</b>	<b>0.00043</b>				
As a percentage of GDP per capita, primary education			<b>1.31455</b>	<b>0.21934</b>		
Quality of schools' educational res.					<b>4.81516</b>	<b>0.93942</b>
<b>Country is post-communist</b>	<b>24.76618</b>	<b>1.82196</b>	<b>19.70100</b>	<b>1.43672</b>	<b>6.89823</b>	<b>2.33726</b>
R-squared	0.3687		0.3694		0.365	
Number of observations	312544		312544		364555	

**Table 5: PISA 2000: Effect of post-communism and other factors on student achievement in reading, science and problem solving**

(Coefficients significant at the 5% level in bold)

	<b>Reading</b>		<b>Science</b>	
	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>				
Constant	<b>208.97930</b>	<b>6.69797</b>	<b>214.79523</b>	<b>7.96060</b>
Student is male	<b>-22.53971</b>	<b>1.08326</b>	<b>2.60777</b>	<b>1.35826</b>
Student is native	<b>6.29364</b>	<b>2.27935</b>	<b>3.48656</b>	<b>2.82079</b>
Index of socio-economic status	<b>0.69682</b>	<b>0.03596</b>	<b>0.63263</b>	<b>0.04485</b>
<b>School factors</b>				
Average socio-economic status	<b>4.92361</b>	<b>0.13607</b>	<b>4.57922</b>	<b>0.16268</b>
<b>Educational spending</b>				
Quality of schools' educational res.	<b>9.40919</b>	<b>1.21102</b>	<b>9.14401</b>	<b>1.48331</b>
<b>Country is post-communist</b>	<b>-24.56763</b>	<b>2.42856</b>	<b>7.71564</b>	<b>2.85872</b>
R-squared	0.3388		0.2709	
Number of observations	201238		111867	

**Table 6: PISA 2003: Effect of post-communism and other factors on student achievement in reading, science and problem solving**

(Coefficients significant at the 5% level in bold)

	<b>Reading</b>		<b>Science</b>		<b>Problem Solving</b>	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>						
Constant	<b>472.06149</b>	<b>2.94874</b>	<b>466.39265</b>	<b>3.00838</b>	<b>462.25868</b>	<b>2.96584</b>
Student is male	<b>-26.88907</b>	<b>0.91714</b>	<b>8.87608</b>	<b>1.10018</b>	<b>2.93256</b>	<b>1.06396</b>
Student is native	<b>33.77829</b>	<b>2.89171</b>	<b>27.73945</b>	<b>2.88472</b>	<b>26.02012</b>	<b>2.95154</b>
Index of socio-economic status	<b>15.71487</b>	<b>0.57688</b>	<b>17.22835</b>	<b>0.66398</b>	<b>16.00666</b>	<b>0.50959</b>
<b>School factors</b>						
Average socio-economic status	<b>59.79314</b>	<b>1.26863</b>	<b>59.20723</b>	<b>1.33665</b>	<b>68.30304</b>	<b>1.38786</b>
<b>Educational spending</b>						
Quality of schools' educational res.	<b>2.59038</b>	<b>0.79880</b>	<b>2.53251</b>	<b>1.03487</b>	<b>3.37139</b>	<b>1.02595</b>
<b>Country is post-communist</b>	<b>-23.57069</b>	<b>2.48937</b>	<b>6.60615</b>	<b>2.73597</b>	<b>6.96081</b>	<b>2.93672</b>
R-squared	0.3589		0.3433		0.3963	
Number of observations	253441		253441		253441	

**Table 7: PISA 2006: Effect of post-communism and other factors on student achievement in reading, science and problem solving**

(Coefficients significant at the 5% level in bold)

	<b>Reading</b>		<b>Science</b>	
	Coeff.	S.E.	Coeff.	S.E.
<b>Individual factors</b>				
Constant	<b>479.37394</b>	<b>2.11942</b>	<b>465.45924</b>	<b>2.57676</b>
Student is male	<b>-30.53719</b>	<b>1.15682</b>	<b>5.96357</b>	<b>0.95949</b>
Student is native	<b>26.52575</b>	<b>2.08207</b>	<b>25.35036</b>	<b>2.48016</b>
Index of socio-economic status	<b>10.72275</b>	<b>0.44892</b>	<b>14.89590</b>	<b>0.41094</b>
<b>School factors</b>				
Average socio-economic status	<b>52.65394</b>	<b>1.24430</b>	<b>50.78107</b>	<b>1.10578</b>
<b>Educational spending</b>				
Quality of schools' educational res.	<b>6.48939</b>	<b>0.87397</b>	<b>5.24401</b>	<b>0.76656</b>
<b>Country is post-communist</b>	<b>-25.47771</b>	<b>2.30085</b>	-1.06042	2.22112
R-squared	0.3311		0.3502	
Number of observations	359219		364555	

**Table 8: PISA 2006 science achievement: Impact of the Fay's variant of the BRR method for calculating standard errors**

(Standard errors for coefficients significant at the 5% level in bold)

	Coefficient	Fay's method S.E.	Standard method S.E.
<b>Individual factors</b>			
Constant	465.45924	<b>2.57676</b>	<b>0.71601</b>
Student is male	5.96357	<b>0.95949</b>	<b>0.28260</b>
Student is native	25.35036	<b>2.48016</b>	<b>0.71262</b>
Index of socio-economic status	14.89590	<b>0.41094</b>	<b>0.17658</b>
<b>School factors</b>			
Average socio-economic status	50.78107	<b>1.10578</b>	<b>0.26134</b>
<b>Educational spending</b>			
Quality of schools' educational res.	5.24401	<b>0.76656</b>	<b>0.13403</b>
<b>Country is post-communist</b>	-1.06042	2.22112	<b>0.40862</b>
R-squared	0.3502		
Number of observations	364555		



# Bibliography

- [1] Alesina, A. and Tabellini, G. 1990. "A positive theory of fiscal deficits and government debt." *Review of Economic Studies* 57: 403-414.
- [2] Brueckner, M. 2001. " Strategic delegation and international capital taxation." Working paper no. B22, Center for European Integration Studies, Bonn.
- [3] Bucovetsky, S. 1991. "Asymmetric tax competition." *Journal of Urban Economics* 30: 167-181.
- [4] Coate, S. and Morris, S. 1999. "Policy persistence." *The American Economic Review* 89: 1327-1336.
- [5] Edwards, J. and Keen, M. 1996. "Tax competition and Leviathan." *European Economic Review* 40: 113-134.
- [6] Fernandez, R. and Rodrik, D. 1991. "Status quo bias in the presence of individual-specific uncertainty." *The American Economic Review* 81: 1146-1155.
- [7] Gans, J. S. and Smart, M. 1996. "Majority voting with single-crossing preferences." *Journal of Public Economics* 59: 219-237.
- [8] Huizinga, H. and Nicodeme, G. 2006. "Foreign ownership and corporate income taxation: An empirical evaluation." *European Economic Review* 50: 1223-1244.
- [9] Kanbur, R. and Keen, M. 1993. "Jeux sans frontières: Tax competition and tax coordination when countries differ in size." *The American Economic Review* 83: 877-892.

- [10] Kehoe, P. 1989. "Policy cooperation among benevolent governments may be undesirable." *Review of Economic Studies* 56: 289-296.
- [11] Krogstrup, S. 2002. "What do theories of tax competition predict for capital taxes in EU countries?" Working paper no. 05/2002, The Graduate Institute of International Studies, Geneva.
- [12] Krogstrup, S. 2004. "A Synthesis of recent developments in the theory of capital tax competition." Working paper no. 2004-02, Economic Policy Research Unit, Copenhagen.
- [13] Nicodeme, G. 2006. "Corporate tax competition and coordination in the European Union: What do we know? Where do we stand?" *European Economy* 250.
- [14] OECD. 2002. *Manual for the PISA 2000 Database*. Paris: OECD.
- [15] OECD. 2002. *PISA 2000 Technical Report*. Paris: OECD.
- [16] OECD. 2003. *PISA 2003 Assessment Framework*. Paris: OECD.
- [17] OECD. 2004. *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.
- [18] OECD. 2005. *PISA 2003 Data Analysis Manual*. Paris: OECD.
- [19] OECD. 2005. *PISA 2003 Technical Report*. Paris: OECD.
- [20] OECD. 2006. *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD.
- [21] OECD. 2007. *PISA 2006: Science competencies for tomorrow's world*. Paris: OECD.
- [22] Olsen, R., Kjærnsli, M. and Lie, S. 2005. "Similarities and differences in countries' profiles of scientific literacy in PISA 2003." Paper presented at the 5th international conference of the European Science Education Research Association. Barcelona.
- [23] Peralta, S. and van Ypersele, T. 2005. "Factor endowments and welfare levels in an asymmetric tax competition game." *Journal of Urban Economics* 57: 258-274.

- [24] Persson, T. and Svensson, L. 1989. "Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences." *The Quarterly Journal of Economics* 104: 325-345.
- [25] Persson, T. and Tabellini, G. 1992. "The politics of 1992: Fiscal policy and European integration." *The Review of Economic Studies* 59: 689-701.
- [26] Persson, T. and Tabellini, G. 1994. "Representative democracy and capital taxation." *Journal of Public Economics* 55: 53-70.
- [27] Wilson, J. D. 1991. "Tax competition with interregional differences in factor endowments." *Regional Science and Urban Economics* 21: 423-451.
- [28] Zabulionis, A. 2001. "Mathematics and Science Achievement of Various Nations." *Education Policy Analysis Archives* 9: 33.