Quintuple Inquiries into Public Economics and Political Economy

Martin Gregor

Doctoral Dissertation

2006

Institute of Economic Studies
Faculty of Social Sciences, Charles University, Prague

CONTENTS

1. INTRODUCTION

The dissertation thesis is a collection of essays on topics in public economics and political economy, attempting to reveal complexities and heretofore neglected or paradoxical effects of rules, procedures, and policies. Richness of contents (fiscal policy, centralization, evolution of preferences, banking socialism, and public sector efficiency) is unified by the underlying approach of classic and evolutionary game theory. Each essay addresses a particular problem, typically conjoint of the political economy and public economics.

The main motivation for the thesis is recent proliferation of political economy, which, along with the currently widespread use of game theory, provides new perspectives on public economics. By linking public economics and political economy, the chapters seek to identify new trade-offs in policy-making, to be recognized and addressed by normative public economics.

Overall, the dissertation reveals how procedural fiscal rules may harm, why centralization may matter even for complementary public goods, how banking socialism arises with retrospective voting, why intolerance is constitutionally tolerable, and concludes that public sector efficiency measures are not robust for the Czech Republic.

Chapter 2 points to heretofore neglected costs of fiscal governance. The clue is that although fiscal governance reduces excessive spending and deficit bias, it also protects executive budget proposals from concessions in legislative bargaining. In political economy with heterogenous legislators, we examine properties of three procedural fiscal rules: timing of a vote on the budget size, spending cap, and deficit cap. We recognize three reverse sides of fiscal governance: nominal constraints pronounce volatility of fiscal policy, timing of vote on budget size may be irrelevant or leads to unpredictable changes, and the constraints increase political polarization by inducing conservative delegation of voters.

Chapter 3 explores strategic delegation of voters when complementary local public goods, aggregated by the symmetric weakest-link technology, are provided in decentralization and centralization. We show that asymmetric access, asymmetric costs, and existence of penalty for the strictly weakest contribution can lead to conservative bias in delegation. Non-cooperative centralization is shown to reduce strategic delegation in all asymmetries.

Chapter 4 formalizes concerns whether intolerance cannot eliminate efficient, yet tolerant confessions in a liberal society. To address this issue, we build an evolutionary model of confessions competing on a market with free entry. We define intolerance as propensity to reciprocal social behavior, rewarding homogeneity and penalizing heterogeneity. No efficient taboo is found absent in evolutionarily stable states, but reciprocity may render some efficient no-taboos extinct. If truth is metaphysical, the only constitutional rule to prescribe is to preserve free entry into market with confessions, and forbid protection of either taboos or no-taboos.

Chapter 5 provides an explanation why a government in a transition economy may postpone bank privatization to keep direct control over credit. When nominal wage rigidity is large, and banking sector oligopolistic, the government may prefer to regulate interest rates to boost labor demand, unless inefficiencies of state ownership are prohibitive. We model a transition economy where the government initially owns enterprises as well as banks. The economy features constant wage, and strong market power of banks. We identify when the government has incentive to privatize enterprises and/or banks. We derive conditions under which the banking socialism (the government owns banks, but privatizes enterprises) dominates other institutional modes: socialism, industrial socialism, and capitalism.

Chapter 6 analyzes the performance and efficiency of the public sector in the new EU member states. We construct a composite indicator and seven sub-indicators of performance, and measure input and output efficiency. We observe the highest public sector performance in Slovenia, and robustly superior performance in Cyprus, Czech Republic and Hungary. Contrariwise, Poland and Slovakia are at the bottom of the ranking. Considering efficiency, Free Disposable Hull method reveals that Slovenia keeps the best position along with Lithuania, while Slovakia is again the worst. Using sector-specific efficiency measures, Cyprus and Hungary move up, while Slovakia and Poland turn to be the worst. The cases of Cyprus and Slovenia show that superior performance can be achieved even with low average costs. Slovakia and Poland, in contrast, show that even a large public sector may produce a low public sector output. Relative efficiency of the Czech Republic is ambiguous due to insensitivity to the index construction.

Chapter 2 draws from an early and more extensive text, presented at the 1st IES Young Scholars Conference (September 2005), and EUROPAEUM Economics Workshop at University of Bologna (October 2005). The paper has been published as IES Working Paper 88/2005, and received 2005 Karel Engliš Prize of the Czech Economic Association. The second version has been presented at the European Public Choice Society Conference at Turku University, Finland (April 2006).

Chapter 3 draws from joint work with Petr Tuchyňa (IES Working Paper 81/2005). This version will be presented at EUROPAEUM Economics Workshop at University of Helsinki (October 2006), and at the 2nd IES Young Scholars Conference (September 2006).

Chapter 4 is an update of the IES Working Paper 72/2005 and the Prague Social Science Studies, EC-005/2005, presented at the Prague Conference on Political Economy (April 2005), and the 5th International Conference for Doctoral Students in Miskolc, Hungary (August 2005). Thrust of the theory ("The Rise and Fall of Confessions: An Evolutionary Game Theory Approach") has been published in the proceedings to the latter conference, Lehoczky, L. & L. Kalmar, eds. (2005), *5th International Conference of PhD Students*, University of Miskolc, Miskolc (ISBN: 963-661-6760). In a slightly longer version, it is to be presented at the conference "Reciprocity: Theories, Facts, Issue" at University of Milan (February 2007).

Chapter 6 has been published in the working paper series of the Prague Social Science Studies, PPF-005/2006.

The doctoral seminar Economic Theory of Political Markets of the Institute of Economic Studies, under auspices of Professor Turnovec, served as a perfect place to present and develop all core ideas of the chapters.

2. THE REVERSE SIDE OF FISCAL GOVERNANCE

2.1 Introduction

One of the core topics of positive political economics as well as normative public finance is the biased budgetary process in democracies, featuring excessive spending and deficits due to common-pool problem (Velasco 1999, von Hagen 2005). Numerous nominal limits, procedural rules, and institutional arrangements on budgetary policy have been designed to accommodate the inefficiencies (Wyplosz 2005), and their richness gave rise to comprehensive indices of fiscal governance, serving since early 1990s as important policy devices.

The fiscal governance indices trace origin from von Hagen (1992), who revived interest in measurement of legislative oversight in Oppenheimer (1983) and incorporated hierarchical procedures within the executive. Measures of fiscal constraints later proliferated in Poterba and von Hagen (1999), Gleich (2003), Yläoutinen (2004), and resulted in OECD/World Bank Budget Practices and Procedures Survey (OECD 2002, OECD/WB 2003), used, among others, by Filc and Scartascini (2004). In parallel, legislative studies of Krafchik and Wehner (1998) as well as Schick (2002) advanced the analysis of separation of the legislature from the executive, with recent contributions by Linert (2005) and Wehner (2005). Additional interest in fiscal governance indices can also be recognized in areas of public sector management, or generally in governance literature.

A very extensive set of rules is typically included in the indices (Filc and Scartascini 2004), ranging from fiscal limits (nominal fiscal rules, targets or ceilings for expenditures, borrowing limits, permitted uses of the budget reserves), hierarchical procedures (power of the Finance Minister, amendments in the Legislature, cash management), and transparency (extra budgetary funds, independent forecasts). The abundance nevertheless brings about weaknesses and pitfalls. First of all, weights in composite indices are difficult to determine. Moreover, although the first-generation research by von Hagen (1992), replicated by Hallerberg (2004), demonstrated significant impact of fiscal procedures on deficit and public debt, later studies are far less conclusive. De Haan, Moessen and Volkerink (1999, p. 284) argue that "... budget institutions affect fiscal policy outcomes, but the effect is quite small." Filc and Scartascini (2004) in a sample of Latin American countries recognize that hierarchical procedures are much more significant than nominal rules and levels of transparency. Hallerberg and Meier (2004) found that only when personal vote is at stake, hierarchical procedures are relevant. This reveals some likely omissions and blind spots in the theory of fiscal governance.

The unifying approach toward fiscal rules, especially regarding procedural and nominal limits, largely refers to common-pool resource incentives, which arise when policy-makers bear only part of tax burden but full benefits of spending on their constituencies. Any coordination, remedying this competitive negative externality game, is thus considered socially optimal (von Hagen 2005). However, as Persson and Tabellini (2000, p. 164) observe, "... one of the underlying problems that 'stricter' budgetary procedures are supposed to solve, namely the common-pool problem, also distorts the level of spending." Ehrhardt, K. *et al.* (2000) receive both theoretically and experimentally that top-down budgeting may result in excessive spending rather than bottom-up procedures. Niepelt (2006) discerns that balanced budget requirement not necessarily induces the government to spend less.

A related stream of literature investigates the role of the U.S. Congressional Budget Act 1974, which established a Budget Committee determining total spending before the composition is set in the Appropriations Committee. Ferejohn and Krehbiel (1987) and in a more general setup also Serrlitzew (2005) found that initial vote in spending often stimulates further spending. This is contested by Dharmapala (2003). The differences stem from defining key actors in the legislative bargaining; in the former, policy-seeking legislators sequentially vote on items, in the latter, interest groups lobby office-seeking legislators under resource constraints.

In this chapter, we attempt to shed new light on fiscal governance, using the former approach of policy-seeking legislators. We analyze sequential budgeting featuring contest of legislators with different priorities on total budgets, when the tragedy of budgetary commons is put aside. Specifically, we study structure-induced equilibria under the assumption that parties in the executive coalition cannot commit their legislators.

The leading idea is that fiscal governance pronounces fiscal volatility stemming from shocks into composition of executive coalitions. Without fiscal governance, the median legislator is relatively strong in balancing the shocks and preserving her optimum in legislative bargaining. Once governance is in place, conservative coalitions restrict power of the median and public-good loving legislators, and protect their budgetary proposals from concessions. The effect of protection is twofold: lower average spending, and higher volatility of tax and spending.

We construct a tractable model, where on the set of feasible allocations, utility is monotonous in spending level in the first period, which allows us to introduce a few useful concepts and receive properties of fiscal rules. We focus on volatility induced by fiscal rules, preciseness of fiscal governance indices, the role of broadness of coalitions (including the role of fragmentation) on the use of fiscal rules, and analyze the impact of fiscal rules on behavior of voters. Having constructed budgeting setup in Section 2.2, we proceed to the analysis in Section 2.3. The last and the least Section 2.4 concludes.

2.2 Setup

2.2.1 Citizens

Assume a population of citizens N in an economy without production. The population lives in two periods, $t \in \{1, 2\}$. Denote exogenous interest rate r. Without loss of generality, assume that each individual is in each period endowed with income constant in real terms across both periods. Hence, denoting the income of the individual $i \in N$ in period t as $y_{i,t}$, we have $y_{i,2} = y_{i,1}(1 + r)$. Let $\bar{y}$ be the average income in the first period.

In each period, incomes are taxed by flat tax $\tau \in [0, 1]$ and the citizens use the after-tax income only for private consumption in that period. Tax revenues of both periods cover production of a public good in both periods, where public expenses for the public good in period t write as $g_t \geq 0$.

We examine the simple quasi-linear utility function, where citizens differ by propensity to consume private good α_i, and where income effects don't exist. The instantaneous utility of each $i \in N$ in period $t \in \{1, 2\}$ writes as

$$u_{i,t} := \alpha_i \ln(1 - \tau)y_{i,t} + g_t, \qquad \text{where} \quad \alpha_i \in (0, 1).$$

Individuals are assumed to be of different age, which is reflected in the probability of surviving the second period, p_i, where $p_i \in \langle \frac{1}{2}, 1)$.[1] Assume a homogenous discount rate equal to banker's interest rate $(1+r)^{-1}$. Now, we

[1] Since individuals exit probabilistically (which can be interpreted as gradual exit), there is a legitimate question what happens to private assets left for private consumption but unconsumed. We assume that they entirely disappear from the economy. For instance,

can write the lifetime utility as a present value of discounted utilities over two periods,

$$U_i := \sum_t u_{i,t} \left(\frac{p_i}{1+r}\right)^{t-1} = u_{i,1} + u_{i,2}\frac{p_i}{1+r}.$$

2.2.2 Budget

The government has free access to financial markets, provided that the budget satisfies the transversality condition with interest rate r,

$$g_1 + \frac{g_2}{1+r} = \tau\bar{y} + \frac{\tau\bar{y}(1+r)}{1+r} = 2\tau\bar{y}. \tag{2.1}$$

Unlike the government, we assume that citizens have no access to credit and no use for savings, so the present value of private good consumption equals in both periods. Individuals cannot make intertemporal optimization when the government induces intertemporal redistribution. This is assumed only to avoid here unimportant effects of Ricardian equivalence.

In this framework, balanced-budget requirement in (2.1) allows to have public deficit in the first period as long as it is ultimately balanced. To capture how total tax revenues are distributed, we introduce deficit as the proportion of the first-period consumption to national income not covered by the first-period revenue, $\beta := (g_1 - \tau\bar{y})(\bar{y})^{-1}$, where $\beta \in \langle -\tau, \tau\rangle$. In other words, zero absolute deficit implies no use of the financial market, minimal deficit $\beta = -\tau$ (maximal surplus) occurs for zero public good in the first period, and maximal deficit $\beta = \tau$ is associated with zero public good provision in the next period.

Further, define spending to be the proportion of the first period spending to national income, $\gamma := g_1\bar{y}^{-1}$, yielding $\gamma = \beta + \tau$. With this notation, budgetary proposals can be written either as (τ, γ), or (τ, β). Moreover, the distribution of tax revenue across periods g_1/g_2 is either the function $g_1/g_2(\beta, \tau)$, or $g_1/g_2(\tau, \gamma)$. In the former case, we get that public spending levels write as $g_1 = (\tau + \beta)\bar{y}$ and $g_2 = (1 + r)(\tau - \beta)\bar{y}$; in the latter, $g_1 = \gamma\bar{y}$ and $g_2 = (1 + r)(2\tau - \gamma)\bar{y}$, hence

$$\frac{g_1}{g_2} = \frac{\tau + \beta}{(1+r)(\tau - \beta)} = \frac{\gamma}{(1+r)(2\tau - \gamma)}. \tag{2.2}$$

all bequests fall into an external pool such as foreign-aid charity. Or, the bequests are transferred to descendants who are not entitled to vote in our political economy. Otherwise we would have to specify how bequests are distributed, which affects equilibrium tax and deficit levels. We put this effect aside for an extended version of the text, although we expect that the new preferences still keep favorable properties.

2.2.3 Legislature

Let L be a set of legislators. The legislators are purely policy-seeking citizen-candidates, $L \subset N$. Suppose N is partitioned into subsets N_k, where $k = 1, \ldots, |L|$, and and each N_k (to be interpreted as district) elects single legislator. Within each district, Condorcet winner is elected (existence is shown in each institutional configuration). In the first part of chapter, we treat L exogenous; extension into endogeneity follows in Section 2.3.6.

The reason to follow the spatial approach with citizen-candidates is manifold. When conflicts on aggregate spending level and tax rate are investigated, group-specific spending is irrelevant, thus no element of common pool problem exists. Equally, there are no 'special' interests per se; the interests conflicting spread across all voters. In such setup, it is reasonable to assume that voters either bind elected legislators to announced policies, or that legislators are simply citizen-candidates with intrinsic preferences. Another important feature is that details of electoral competition are omitted, and, unlike in Terai (2003), electoral alliances cannot emerge. Unconstrained voting in the Legislature thus produces the median voter outcome.

Let q, where $q > \frac{|L|}{2}$, be the majority quota for voting in the legislature. Any subset of legislators $C \subseteq L$, where $|C| \geq q$, we shall call a winning coalition. Assume all winning coalitions are equiprobable. (The winning coalitions are not necessarily minimum-winning, or connected, i.e. bliss points of non-members may belong to a minimum convex set including all bliss points of coalition members.) The assumption of equiprobability may look contradictory, yet it reflects the main purpose—to model how exogenously given coalitions affect the budgets. The source of shock into composition of coalitions may rest with personal antipathies, trust, secret information, or reelection considerations of the President who designates the prime minister.

The coalition is assumed to maximize by standard utilitarian yardstick. Hence, for any winning coalition C, the first-best outcome of bargaining maximizes the sum of lifetime utilities, $(\tau_C^*, \gamma_C^*) = \text{argmax} \sum_{j \in C} U_j(\tau, \gamma)$. The coalition however submits a budget proposal (τ_C, γ_C), which anticipates constraints in the legislative bargaining, thus may differ from (τ_C^*, γ_C^*). The precise derivation of the proposals follow in Propositions 1, 2, and 3.

2.2.4 Fiscal governance

Before outlining the budget process in detail, we describe budgetary rules. We draw on rules examined in a fiscal governance index by De Haan, Moessen and Volkerink (1999), who re-constructed a seminal index by von Hagen (1992). Of the index, we concentrate on two items, the timing of vote on budget size and the presence of explicit fiscal constraints (ceilings/limits), as investigated by De Haan et al. (1999, p. 286–7) in the two survey questions:

5.(e) Is there a global vote on total budget size?

Final only
Initial

6. Could you please indicate whether the government is bound by some general constraint?

No constraint
...
Government spending to GDP or Golden Rule
Government spending to GDP & Deficit to GDP

Clearly, these two questions focus on three procedural rules: *Vote on budget size* (initial/final), *Spending constraint* (present/absent), and *Deficit constraint* (present/absent). The first rule is identical to the U.S. 1974 Congressional Budget Act, and the second is close to its supplement in the form of statutory spending limits in the Budget Enforcement Act of 1997. OECD/WB (2003) and Filc & Scartascini (2004) investigate equivalent rules in Questions 2.2.b.1 and 2.7.e.

2.2.5 Budget process

Members of the executive coalition, C, submit the budgetary proposal, (τ_C, γ_C). The optimal proposal is intimately related to the set of feasible outcomes, which is determined by fiscal governance. In one extreme, when the proposal can't affect incentives of legislators, it is cheap talk, and the set of enforceable outcomes reduces to single point. In the other extreme, elements from a large set of proposals are enforceable in the legislature.

The budgetary process can be split into cabinet and legislative phase.

1. Fiscal targets. If *Spending constraint* applies, the coalition C imposes an enforceable constraint $\gamma \leq \gamma_C$. At the same time, if *Deficit constraint* applies, it sets that $\gamma - \tau = \beta \leq \beta_C = \gamma_C - \tau_C$.

Then, the budget proposal proceeds to the Legislature.

1. Initial vote. If initial *Vote on budget size* applies, there is simple-majority vote on the budget size g_1 (or, γ).

2. Revenue side. There is simple-majority vote on tax rate τ.

3. Final vote. If final *Vote on budget size* applies, there is simple-majority vote on the budget size g_1 (or, γ).

Three binary variables (Vote on budget size, Deficit limit, Spending limit) offer eight decision-making configurations. Table 1 describes how spending and revenues are set in the presence of constraints present in each stage, when either initial or final Vote takes place, all with exception of obvious constraints of $\tau \in \langle 0,1 \rangle$ and $\gamma \in \langle 0, 2\tau \rangle$. Figure 1 depicts constraints in the space $\tau \times \gamma$.

Tab. 2.1: Institutional configurations (Initial/Final vote)

Vote	Constraints		Stage 1	Stage 2
Initial	No		γ	$\tau \geq \frac{\gamma}{2}$
	Spending	γ_C	$\gamma \leq \gamma_C$	$\tau \geq \frac{\gamma}{2}$
	Deficit	β_C	$\gamma \leq \beta_C + 1$	$\tau \geq \max\{\frac{\gamma}{2}, \gamma - \beta_C\}$
	Both	γ_C, β_C	$\gamma \leq \min\{\beta_C + 1, \gamma_C\}$	$\tau \geq \max\{\frac{\gamma}{2}, \gamma - \beta_C\}$
Final	No		τ	$\gamma \leq 2\tau$
	Spending	γ_C	τ	$\gamma \leq \min\{2\tau, \gamma_C\}$
	Deficit	β_C	τ	$\gamma \leq \min\{\frac{\tau}{2}, \tau + \beta_C\}$
	Both	γ_C, β_C	τ	$\gamma \leq \min\{\frac{\tau}{2}, \beta_C, \gamma_C\}$

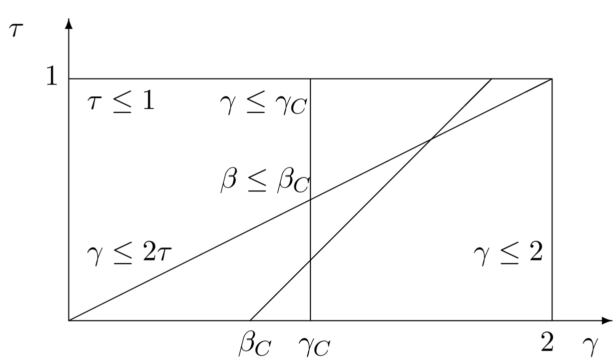

Fig. 2.1: Constraints on budget proposals

2.3 Analysis

From the lifetime utility, $U_i = y_{1,i}[\alpha_i(1+p_i)\ln(1-\tau)+(1-p_i)\gamma+2p_i\tau$, we derive Lemmas 1 and 2.

Lemma 1 (Bliss point): The first-best optimum is

$$(\tau_i^*, \gamma_i^*) = \left(1 - \frac{\alpha_i(1+p_i)}{2}, 2-\alpha_i(1+p_i)\right). \tag{2.3}$$

Proof First, $\frac{\partial U_i}{\partial \gamma} = y_{i,1}(1-p_i) > 0$ and by the intertemporal budgetary balance in (2.1), $\gamma \leq 2\tau$ holds, hence we have $\gamma_i^* = 2\tau_i^*$. For $y_{1,i} > 0$, we have $\frac{\partial U_i}{\partial \tau}|_{\gamma=2\tau} = 0 \Rightarrow \tau_i^* = 1 - \frac{\alpha_i(p_i+1)}{2}$. Notice that by assumption, $p \in \langle \frac{1}{2}, 1)$ and $\alpha_i \in (0,1)$, therefore $\tau_i^* \in (0,1)$. $\square$

We find it useful to introduce induced optima: optimal spending conditional on fixed tax, $\gamma_i(\tau_C)$, optimal tax conditional on fixed spending, $\tau_i(\gamma_C)$, optimal tax conditional on fixed deficit, $\tau_i(\beta_C)$, and optimal spending conditional on fixed deficit, $\gamma_i(\beta_C)$.

Lemma 2 (Induced optima): The conditional optima are as follows:

$$\tau_i(\gamma_C) = \begin{cases} \tau_i^* & \text{for} \quad \gamma_C \leq \gamma_i^* \\ 2\gamma_C & \text{for} \quad \gamma_C > \gamma_i^* \end{cases} \tag{2.4}$$

$$\gamma_i(\tau_C) = 2\tau_C \tag{2.5}$$

$$\tau_i(\beta_C), \gamma_i(\beta_C) = \begin{cases} 1-\alpha_i, 1-\alpha_i+\beta_C & \text{for} \quad \beta_C \leq 1-\alpha_i \\ \beta_C, 2\beta_C & \text{for} \quad \beta_C > 1-\alpha_i \end{cases} \tag{2.6}$$

Proof To get (2.4), we use that for $\tau_i(\gamma_C) = \tau_i^*$, so $\frac{\partial U_i}{\partial \tau} = 0$. However, this tax rate is feasible only for $\gamma_C \leq \gamma_i^*$; once $\gamma_C > \gamma_i^*$, we would have $\gamma_C \geq 2\tau_i^* = 2\tau_i(\gamma_C)$, and the bliss-point tax rate would be insufficiently low to cover fixed expenses γ_C. As $\frac{\partial U_i(\tau)}{\partial \tau} = y_{i,1}[\frac{\alpha_i(p_i+1)}{\tau-1}+2p] < 0$ for any $\tau > \tau_i^*$, the tax rate must be as low as possible, namely $\tau_i(\gamma_C) = 2\gamma_C$.

In the case of (2.5), we simply have $\frac{\partial U_i}{\partial \gamma} = y_{i,1}(1-p_i) > 0$. Corner solution $\gamma_i(\tau_C) = 2\tau_C$ applies.

We rewrite $U_i(\tau, \beta_C) = y_{i,1}[\alpha_i(p_i+1)\ln(1-\tau_i)+(\beta_C+\tau)+p_i(\tau-\beta_C)]$ in order to derive (2.6). For sufficiently low $\beta_C \leq 1-\alpha_C$, we can use that $\frac{dU_i}{d\tau}|_{\beta=\beta_C} = 0 \Rightarrow \tau_i(\beta_C) = 1-\alpha_i$. In that case, we apply that $\beta = \beta_C$ along with definition $\gamma - \tau = \beta_C$ yield $\gamma_i(\beta_C) = \tau_i(\beta_C)+\beta_C = 1-\alpha_i+\beta_C$.

This allocation is however unfeasible for $\beta_C > 1-\alpha_C$, since we would have $1-\alpha_i+\beta_C > 2(1-\alpha_i)$, and the constraint $\gamma \leq 2\tau$ would be violated. We use that $\frac{dU_i(\tau)}{d\tau}|_{\beta=\beta_C} = y_{i,1}\left[\frac{\alpha_i(p_i+1)}{\tau-1}+1+p_i\right] < 0$ for any $\tau > 1-\alpha_i$, so we have to have tax rate as low as possible, namely $\tau = 2\gamma$. As a result, $\tau_i(\beta_C) = \beta_C$ and $\gamma_i(\beta_C) = 2\beta_C$ are conditional optima for $\beta_C > 1-\alpha_C$. $\square$

First, notice that spending conditional on tax is maximal, i.e. $\gamma_i(\tau_C) = 2\tau_C$, which largely helps in analyzing solutions. The reason is that the quasi-linear utility function induces maximal deficit when tax rate is fixed, $\frac{\partial U(\beta,\tau)}{\partial\beta} > 0$. More specifically, $p < 1$ and linearity in public consumption mean that the present value of the future public consumption has always lower marginal utility than current public consumption, so individuals maximize deficits. For a fixed tax rate, this conditional optimum entails maximum spending in the first period.

The ordering of individuals, when ranked by various conditional optima, may change. For example, a legislator whose first-best tax rate is median, not necessarily remains median when legislature optimizes on the fixed deficit. Therefore, we introduce new notation, capturing median taxes for several cases of conditional optima. Generally, denote median value of x_i as $\tilde{x}$. We introduce median first-best tax rate, τ_M, median conditional-spending tax rate, τ_S, and median conditional-deficit tax rate, τ_B, such that

$$\tau_M := \tilde{\tau}_i^* = 1 - \frac{\widetilde{\alpha_i(p_i+1)}}{2}, \quad \tau_S := \tilde{\tau}_i(\gamma_C) = 1 - \frac{\widetilde{\alpha_i(p_i+1)}}{2p_i}, \\ \tau_B := \tilde{\tau}_i(\beta_C) = 1 - \tilde{\alpha}_i.$$

Lemma 3: For any $\mathbf{p}$, α, and $\mathbf{y}$, where $p_i \in (\frac{1}{2}, 1), \alpha_i \in (0,1), y_i > 0$ for all i, we have $\tau_S \leq \tau_B \leq \tau_M$.

Proof By contradiction, we prove $\tau_S \leq \tau_B$. For $p_i \in (\frac{1}{2}, 1)$ and $\alpha_i \in (0,1)$, we always have $1 - \frac{\alpha_i(p_i+1)}{2p_i} < 1-\alpha_i$. Suppose $\tau_B < \tau_S$. Define subset $J = \{i : 1-\alpha_i \leq \tau_B\}$. It must be that the size of J is at least one half of the original set. For all $i \in J$, we have $1 - \frac{\alpha_i(p_i+1)}{2p_i} < 1-\alpha_i \leq \tau_B < \tau_S$. Therefore, for at least one half of individuals, we have $1 - \frac{\alpha_i(p_i+1)}{2p_i} < \tau_B$, which implies $\tau_S \leq \tau_B$. This contradicts $\tau_S > \tau_B$. The proof of $\tau_B \leq \tau_M$ is done by analogy. We only use that for $p_i \in (\frac{1}{2}, 1)$ and $\alpha_i \in (0,1)$, we always have $1-\alpha_i \leq 1 - \frac{\alpha_i(p_i+1)}{2}$. $\square$

Lemma 3 is helpful in analyzing the subgame-perfect Nash equilibria. We drive the solutions by backward induction, starting in the final stage when constraints are imposed, and getting optimum under these constraints. Regardless of institutional configuration, the optimum is aligned with either of median players, i.e. legislators whose respective optima correspond to τ_M, τ_S, and τ_B . In the preceding step, the individuals are again sorted, but now by preferences over constraints. We again get that decisive player is either M, S, or B. Finally, coalitional decision-making about spending and/or deficit constraints is driven by maximization of transferrable utility.

Proposition 1 (No constraint): Without nominal constraints, the Legislature adopts the budget $(\tau, \gamma) = (\tau_M, 2\tau_M)$, regardless of initial/final Vote on budget size.

Proof In initial vote, Legislature votes γ in Stage 1 and τ in Stage 2. In Stage 2, the median proposal is $\tilde{\tau}_i(\gamma)$, which is τ_S for $\gamma \leq 2\tau_S$ (subset Θ_1) and $\frac{\gamma}{2}$ for $\gamma \geq 2\tau_S$ (subset Θ_2). In Stage 1, the conditional median outcome is anticipated, so the legislators optimize over $\Theta_1 \cup \Theta_2$. Optimization over Θ_1, where tax is fixed, yields always maximal spending, so $\gamma = 2\tau_S$, which is also an element of Θ_2. Therefore, the problem reduces to optimization in Θ_2. Individuals optimize over corner solutions, where they have quasiconcave preferences with maximum in τ_i^*. The median value is τ_M, so it is the candidate for solution. The only difference to unconstrained optimization is the limit $\tau \geq \tau_S$. However, this doesn't affect quasiconcavity of preferences of any individual, even of those with $\tau_i^* < \tau_S$. Moreover, the median is attainable, as $\tau_M > \tau_S$. As a result, we have $(\tau, \gamma) = (\tau_M, 2\tau_M)$.

In final vote, Legislature votes τ in Stage 1 and γ in Stage 2. In Stage 2, everyone prefers maximal spending, as $\gamma(\tau_C) = 2\tau_C$. This nonconflicting voting is anticipated in Stage 1. Therefore, the problem reduces to optimization over the corner solutions. Again, individuals have quasiconcave preferences, so τ_M is median value. In the absence of limit, we must have $(\tau, \gamma) = (\tau_M, 2\tau_M)$. $\square$

Unconstrained voting leads to an outcome identical to veto of the (median) legislator M. The outcome is stable regardless of coalition C, because the executive coalition has no device to bind the legislators. In the absence of fiscal governance, executive coalitions in our simplified model can't affect fiscal variables, and neither coalitional instability nor majority quota affect fiscal policy, as long as the set L is exogenous.

2.3.1 Spending limit

When Spending constraint is present and cannot be abolished or relaxed, the main difference is that γ_C may affect incentives in both subsequent stages. Therefore, we need to derive at least some properties of a jointly maximized coalitional utility function.

Lemma 4: The coalition C prefers the budget

$$(\tau_C^*, \gamma_C^*) = \left(1 - \frac{\sum_{j\in C} y_{j,1}\alpha_j(1+p_j)}{2\sum_{j\in C} y_{j,1}}, 2 - \frac{\sum_{j\in C} y_{j,1}\alpha_j(1+p_j)}{\sum_{j\in C} y_{j,1}}\right).$$

Proof Coalitional bargaining is driven by maximization of lifetime utilities. As $\frac{\partial \sum_{j\in C} U_j(\tau,\gamma)}{\partial\gamma} = \sum_{j\in C}\left(\frac{\partial U_j(\tau,\gamma)}{\partial \gamma}\right) > 0$, we optimize on corner solutions only, $\gamma = 2\tau$. We get $\frac{d\sum_{j\in C} U_j(\tau,2\tau)}{d\tau}|_{\gamma=2\tau} = y_{1,1}\left[-\frac{\alpha_1(1+p_1)}{1-\tau} + 2\right] + \ldots + y_{|C|,1}\left[-\frac{\alpha_{|C|}(1+p_{|C|})}{1-\tau} + 2\right] = 2\sum_{j\in C} y_{j,1} - \frac{\sum_{j\in C} y_{j,1}\alpha_j(1+p_j)}{1-\tau}$. Only for $\tau_C^* = 1 - \frac{\sum_j y_{j,1}\alpha_j(1+p_j)}{2\sum_j y_{j,1}}$, the F.O.C. $\frac{d\sum_{j\in C} U_j(\tau,2\tau)}{d\tau}|_{\gamma=2\tau} = 0$ is satisfied. □

Each coalition may set the constraint equal to the coalitional optimum, $\gamma_C = \gamma_C^*$. Nevertheless, if $\gamma_C \geq \gamma_M = 2\tau_M$, the constraint is inactive, because only a minority consisting of extreme public-good lovers, with $\tau_i^* \geq \tau_M$, are restricted, and median is not affected. The constraint affects result only for $\gamma_C < 2\tau_M$, because majority of legislators prefer then $\gamma \geq 2\tau_M$, but are bound by spending constraint, $\gamma \leq \gamma_C < 2\tau_M$. At the same time, we can see that no coalition, including extremely conservative ones, has any incentive to set too low spending constraint; for any $\gamma_C < 2\tau_S$, the Legislature would adopt τ_S, but for this "minimal tax", everybody—including coalition members—prefer maximal spending $\gamma_S = 2\tau_S > \gamma_C$. This leads our intuition towards the finding that the conservative coalitions apply the constraint to obtain conservative outcomes, but not the excessively conservative ones.

Proposition 2: The existence of Spending constraint allows the coalition C secure any outcome (τ, γ), where $\tau_S \leq \tau \leq \tau_M$ and $\gamma = 2\tau$. The coalition opts for the following Spending constraint irrespective of initial/final Vote on budget size: $\gamma_C = 2\min\{\max[\tau_S, \tau_C^*], \tau_M\}$.

Proof We divide the proof in two parts; in the first part, we derive which allocations can emerge from the use of Spending constraint. In the second part, we determine which value of the constraint the ruling coalition selects.

Start with the initial vote, where Legislature votes γ in Stage 1 and τ in Stage 2. In Stage 2, the median proposal is $\tilde{\tau}_i(\gamma)$, which is τ_S for $\gamma \leq 2\tau_S$ (subset Θ_1) and $\frac{\gamma}{2}$ for $\gamma \geq 2\tau_S$ (subset Θ_2). In Stage 1, the conditional median outcome is anticipated, so the legislators optimize over $\Theta_1 \cup \Theta_2$. Like in proof of (1), the problem reduces to optimization in Θ_2 (the product of the two dominates any other allocation in Θ_1). Individuals thus optimize over corner solutions (on the interval $\tau \in \langle \tau_S, \tau_C \rangle$), where they have quasiconcave preferences with maximum in τ_i^*. The median value is τ_M, so it is the candidate for solution. The only difference to situation without any constraint is the Spending limit $\gamma \leq \gamma_C$ (or $\tau \leq \tau_C = \frac{\gamma_C}{2}$). On one hand, this constraint doesn't affect quasiconcavity of preferences of any individual, so the median voter M is again decisive. On the other hand, once $\tau_C < \tau_M$, the median voter faces a binding upper constraint and the Legislature adopts the upper bound, $(\tau, \gamma) = (\tau_C, 2\tau_C)$.

In final vote, Legislature votes τ in Stage 1 and γ in Stage 2. In Stage 2, everyone prefers maximal spending. This nonconflicting voting is anticipated in Stage 1, and the problem reduces to optimization over corner solutions on the interval $\tau \in \langle \tau_S, \tau_C \rangle$. This yields identical result as for the initial vote case.

We have seen that the coalition always sets $\tau_C \geq \tau_S$, even for $\tau_C^* < \tau_S$. Secondly, we proved that once $\tau_C > \tau_M$, the constraint is inactive, and Legislature adopts τ_M. If $\tau_C \leq \tau_M$, the constraint is active and Legislature adopts the maximal feasible tax rate, τ_C. All in all, we get that the coalition can use the constraint to force the Legislature to select tax rate from the interval $\tau_C \in \langle \tau_S, \tau_M \rangle$. Since preferences are quasiconcave over corner solutions, we have that public-good loving coalitions, for which $\tau_C^* > \tau_M$, have no reason to use Spending constraint actively, and in optimum weakly prefer $\tau_C = \tau_M$. Conservative coalitions, for which $\tau_C^* < \tau_M$, select Spending constraint actively. Specifically, excessively conservative coalitions, for which $\tau_C^* < \tau_S$, select the lower bound, $\tau_C = \tau_S$, and each moderately conservative coalition, for which $\tau_S \geq \tau_C^* < \tau_M$, is able to secure the coalitional optimum, $\tau_C = \tau_C^*$. $\square$

If the coalition consists of public-good lovers, $\tau_C \geq \tau_M$, the solution is identical to the case with no constraint. In other words, public-good lovers cannot use the Spending constraint to promote their preferences in subsequent legislative votes. Extreme conservatives can only shift the outcome towards its optimum, while moderate conservatives are able to use the constraint to secure their own coalitional optimum. This may explain why constitutional preferences for spending rules stem from moderate right wing parties rather than extreme rightist parties, not to speak about left wing parties.

2.3.2 Deficit limit

Interestingly, application of Deficit constraint is but a subcase of the former constraint. Again, a too soft budget constraint, $\beta_C = \gamma_C - \tau_C > \gamma_M - \tau_M$, doesn't affect what the Legislature passes. And, again, a too hard constraint, here $\gamma_C - \tau_C < \gamma_B - \tau_B$, is counterproductive because the Legislature would adopt τ_B, and each party would prefer maximal spending, namely $2\tau_B$. Like in Spending constraint, the coalition can secure a non-empty set of outcomes, but the interval is narrower, because by Lemma 3, $\tau_S < \tau_B < \tau_M$.

Proposition 3: The existence of Deficit constraint allows the coalition to secure any outcome (τ, γ), where $\tau_B \leq \tau \leq \tau_M$ and $\gamma = 2\tau$. The coalition opts for the following Deficit constraint irrespective of the initial/final Vote on budget size, $\beta_C = \min\{\max[\tau_B, \tau_C^*], \tau_M\}$, and the budgetary outcome is $(\tau, \gamma) = (\beta_C, 2\beta_C)$.

Proof To understand essentials of the proof, recall conditional optimum in (2.6). The shape of the curve of the optimum looks like the curve of the conditional optimum in (2.4), only τ_B replaces τ_S. Therefore, voting problem is analogous to the case with Spending constraint, with τ_B replacing τ_S. □

2.3.3 Public sector size/fiscal volatility trade-off

The key finding of the model is that there is a trade-off between average tax level and volatility of taxes. The stricter institutional configuration implies higher volatility. This is stated in Proposition 4 and illustrated on Figures 2.2 and 2.3. Distribution of tax proposals is denoted by F, with superscript corresponding to each institutional configuration. Distribution of coalitional bliss points is denoted by $F(\tau_C^*)$. Since in our simple model, in the optimum always $\gamma = 2\tau$, the finding equally applies to the level and volatility of spending, and it captures fiscal volatility as such.

Proposition 4: For any exogenous distribution of C, expected taxes are lowest with the Spending constraint and highest with No constraint. Volatility in taxes with either Spending or Deficit constraint exceeds volatility with No constraint. Using corresponding superscripts, $E^S(\tau) \leq E^B(\tau) \leq E^N(\tau) = \tau_M$, $\sigma^S(\tau) \geq \sigma^N(\tau) = 0$, and $\sigma^B(\tau) \geq \sigma^N(\tau) = 0$.

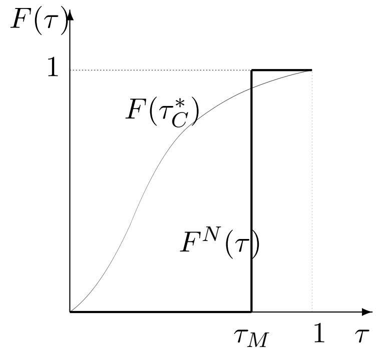

Fig. 2.2: Distribution of taxes (No constraint), $F^N(\tau)$

Proof From Proposition 1, we have $\tau = \tau_M$ for all C, hence $E^N(\tau) = \tau_M$ and $\sigma^N(\tau) = 0$. Divide the set of all coalitions into four subsets, $\mathbb{C}^1 = \{C : \tau_C^* < \tau_S\}$, $\mathbb{C}^2 = \{C : \tau_S \leq \tau_C^* < \tau_B\}$, $\mathbb{C}^3 = \{C : \tau_B \leq \tau_C^* < \tau_M\}$, and $\mathbb{C}^4 = \{C : \tau_M \leq \tau_C^*\}$. Denote $w_1, \ldots, w_4$ the ratio of number of coalitions in each group to the number of all coalitions. We can write $E^S(\tau) = \sum_{i=1}^4 w_i E^S(\tau, \mathbb{C}^i)$ and $E^B(\tau) = \sum_{i=1}^4 w_i E^B(\tau, \mathbb{C}^i)$. From Propositions 2 and 3, we have $E^S(\tau, \mathbb{C}^1) = E^B(\tau, \mathbb{C}^1) = \tau_S$, $E^S(\tau, \mathbb{C}^2) < E^B(\tau, \mathbb{C}^2) = \tau_B$, $E^S(\tau, \mathbb{C}^3) = E^B(\tau, \mathbb{C}^3)$, and $E^S(\tau, \mathbb{C}^4) = E^B(\tau, \mathbb{C}^4) = \tau_M$. Therefore, $E^B(\tau) - E^S(\tau) = w_2[E^B(\tau, \mathbb{C}^2) - E^S(\tau, \mathbb{C}^2)]$. For $w_2 = 0$, we have $E^B(\tau) = E^S(\tau)$ and for $w_2 > 0$, we have $E^B(\tau) > E^S(\tau)$. Since $w_i \in \mathbb{N}$, we always have $E^B(\tau) \geq E^S(\tau)$.

In a limit case when $w_4 = 1$, we have for all C, $\tau_C^* = \tau_M$, hence $E^B(\tau) = E^S(\tau) = E^N(\tau)$ and $\sigma^S(\tau) = \sigma^B(\tau) = \sigma^N(\tau) = 0$. If $w_4 < 1$, we always have to have $\sigma^S(\tau) > 0$ and $\sigma^B(\tau)$. Formally, there exists C', for which $\tau_C^* < \tau_M$, and $C'' \neq C'$, for which $\tau_C^* \geq \tau_M$. As a result, $\sigma^S(\tau) > 0$ and $\sigma^B(\tau) > 0$. □

We have derived a standard result, that spending and deficit constraints reduce public spending. A less standard result is that timing of the Vote on budget size has no effect, which is however only a consequence of optimizing on corner solutions. Even less obvious result is that the deficit limit is a generally weaker constraint than the spending limit. Again, this stems from the fact that bliss points are located on the boundary where spending level

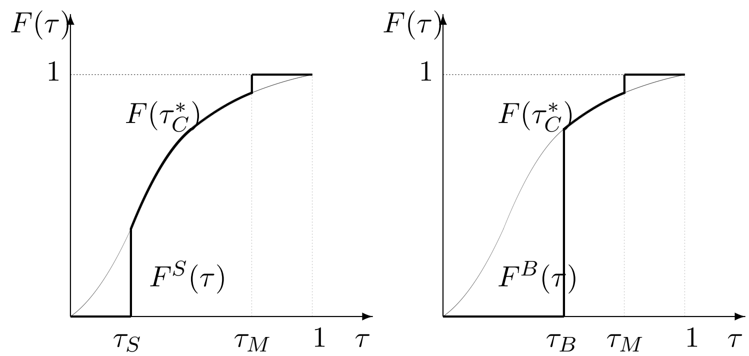

Fig. 2.3: Distribution of taxes for Spending constraint ($F^S(\tau)$) and Deficit constraint($F^B(\tau)$)

is maximal, so the present public good consumption is always preferred to public good consumption in the future.

A novelty of this approach is to point to volatility, induced by the procedural rules. The procedural fiscal rules protect budgetary proposals of moderate conservative coalitions, so if the distribution of coalitions is exogenous, we observe higher volatility of taxes and spending levels. Distribution of taxes, in comparison with distribution of individually optimal taxes, is more skewed in the presence of deficit and especially spending constraint. The skewness is related to the fact that there are more public-loving coalitions that cannot protect their proposals than conservative coalitions.

2.3.4 Fiscal governance index

Fiscal governance indices may not appropriately reflect constraining limits of the procedural rules, as the political economy with quasi-linear utility shows. To demonstrate this property, we construct *apriori index* based on the presence of rules like in the empirical literature, and the *real index*, reflecting the extent to which the governance mode reduces spending in the model. Obviously, using both indices is normatively useful only on presumption that the median behavior exhibits spending bias, which is however not modeled here.

Apriori index, normalized for minimum zero and maximum six points, is constructed such that two points are given for each procedural rule. Initial vote rule means two points (Final vote zero points), and presence of Spending constraint and Deficit constraint give two additional points each. Real index, normalized for minimum zero and maximum six points, counts results, not procedures. If the median tax τ_M is never reduced, we impose zero points. If the executive coalition can select from a narrow interval, $\tau \in \langle \tau_B, \tau_M \rangle$, we count three points. If the coalition optimizes on a larger interval, $\tau \in \langle \tau_S, \tau_M \rangle$, we count six points. Correlation coefficient of these indices, when each configuration weights equally, is 0.69. For comparison of the two indices, see Table 2.2.

Tab. 2.2: Normalized Apriori/Real index values

	No	Spending	Deficit	Both
Initial	2/0	4/6	4/3	6/6
Final	0/0	**2/6**	2/3	4/6

One deviation of real index from apriori index deserves special attention: on one hand, the apriori index underestimates disciplining effect of the Spending constraint with Final vote, and on the other hand overestimates effect of both constraints in Initial mode. The reason is threefold here: First, Initial vote on the budget size in fact brings no additional discipline in this model. Second, when Spending constraint is present, Deficit constraint is redundant. And last, which is the reverse of the latter, Spending constraint is stronger than Deficit constraint.

2.3.5 Majority quota

Decrease in majority quota q lowers the average number of coalition members and increases the total number of winning coalitions. New budgetary proposals emerge, both compromising ones and extreme ones. The fact that there exist new coalitions in which the most extreme legislator need not to make large concessions, makes distribution of proposals a less skewed. As a result, relative more extremely conservative budgetary proposals can be protected by imposing spending or deficit constraint.

With a larger number of coalitions, procedural constraints thus become relatively more important. The stronger pro-conservative effect of constraints however doesn't imply that also the average budget declines. In cases when individuals with extremely conservative preferences obtain very large incomes, the lower quota rather diminishes their bargaining power, which occurs for two reasons. First, for lower q, relatively many new coalitions, where the extreme individuals are not included, emerge.[2] Second, the existence of lower bound, $\tau \geq \tau_S$, allows to protect moderately conservative coalitions better than than extremely conservative coalitions; the new extreme coalitions have no better choice but the lower bound, which exceeds their optimum. Hence, lower quota may work in the same way as extension of franchise to poor voters. Decrease in quota may be in the interests of big spenders, as the size of government becomes inverse to broadness of coalitions. Recall that we achieved that without any analysis of coordination of budgetary commons, where narrower coalitions are associated with lower ability to internalize fiscal externalities, and the high externality gives rise to excessive spending.

Another likely implication concerns classic "policy-seeking vs. office-seeking" trade-off between re-election and rent. The tradeoff may be endogenous to the level of quota. Lower quota and narrow coalitions redistribute more rents per member, and use fiscal constraint more pro-conservatively. The motivation of conservative legislators for lower quota (in the form of majoritarian elections, for example), may reflect not office-seeking, but policy-seeking motives.

2.3.6 Endogenous legislators

By Proposition 4, the distribution of taxes in the presence of Deficit and especially Spending constraint is more skewed in comparison with distribution of individually optimal taxes. We may expect that the skewness, namely the fact that a large number of coalitions cannot protect their proposals, affects selection of legislators. The voters may resort to strategic electoral behavior, known also as strategic delegation.

Extension towards endogenous coalitions can proceed in two ways—by making the set legislators L endogenous, and by abandoning the equiprobability assumption. We follow the first way as suggested in Section 2.2.3. In the analysis, we need not to derive all subgame-perfect Nash equilibria of the game. We only investigate the case of sincere delegation. By sincere delegation, denote a strategy profile when each district nominates a legislator on the basis of preferences of median voter within the district.

Proposition 5: Sincere delegation is a subgame-perfect Nash equilibrium when

[2] Consider legislature with $\tau^* = (1, 2, 3, 4, 5)'$, quota $q^1 = 3, q^2 = 4$, spending constraint and $\tau_S = 2 < \tau_M = 3$. When total income 60 is distributed by $\mathbf{y} = (56, 1, 1, 1, 1)'$, we have $E_1(\tau) \doteq 2.31 > 2.17 \doteq E_2(\tau)$. When the income is more equally distributed, by $\mathbf{y} = (9, 1, 40, 9, 1)'$, we get $E_1(\tau) \doteq 2.75 < 2.79 \doteq E_2(\tau)$.

nominal fiscal constraints are absent. With the Deficit and Spending constraint, sincere delegation is not a Nash equilibrium.

Proof We examine stability of sincere voting. In case without constraint, the median district elects the citizen-candidate whose first-best optimum is identical with the budgetary outcome. Consider deviations from sincere voting in any conservative district. To vote strategically would change the budget only if the elected citizen-candidate were more public-good-loving than the median legislator. The budget might become more expansionary. However, a candidate promoting this change cannot win in a pairwise vote with candidate who represents the median in a conservative district. Quite similarly, any public-good-loving district can only reduce the budget, by delegating more conservatively that the median legislator. This nomination again loses in a pairwise vote with median outcome in a conservative district. To sum up: when the budgetary outcome is also Condorcet winner within the district, there is no incentive for other but sincere delegation.

With Spending constraint, consider any extremely conservative district j, with the Condorcet-winning allocation with district $\tilde{\tau}_j < \tau_S$. Denote optimum of the sincerely delegated candidate $\tau_j^* = \tilde{\tau}_j$. By voting more conservatively $\tau_j^{*\prime} < \tau_j^*$, we have $\forall C, j \in C : \tau_C^{*\prime} < \tau_C^*$. By Proposition 2, we have for majority of voters $\tau_C^{*\prime} < \tau_C^* \implies \tau_S \leq \tau_C^{\prime} \leq \tau$. Within the extremely conservative district, we have $\tau' \succeq \tau$, so the more conservative candidate (with the more conservative outcome) wins in a pairwise vote with the delegate who represents the median voter of the district. $\square$

Without nominal fiscal rules, explanation is obvious. The median district preference is assured by Proposition 1, so whenever the median district votes sincerely and remaining districts by vote more extremely, thereby pronouncing the difference to median district, only τ_M will be adopted.

With nominal constraints, sincere voting is suboptimal for extremely conservative districts (i.e. where median voter has preference below the interval in which proposals are secured). By voting more extremely, the district shifts all compromises in which his legislator takes part towards the optimum of the district median. This is costless, because compromises with even more extreme legislators are not affected. Therefore, the best response is voting more extremely. The presence of fiscal constraints in political markets with costless voting and rational voters thus distorts elections, and further affects volatility of tax rates and spending levels.

2.3.7 Extensions

The Cobb-Douglas variant of the model in Gregor (2005) extends the considerations in two ways. First, volatility and skewness of distribution, when induced by constraints, relates differently to taxes, spending, and deficit levels. Second, the presence of initial vote changes equilibrium budget, but not in any predictable direction.

Properties of the Spending limit are very similarly to the model with quasi-linear utility: the limit reduces both tax and spending levels, and deficits are unchanged (final Vote), or changed to a low extent (initial Vote). Spending constraint is stronger than Deficit limit, because it binds not only spending, but also tax rates.

Relative deficit limit works differently, which is especially evident if a new type of deficit constraint (cap on deficit expressed in percentage of the budget, not GDP) is taken into account. Then, tax rate remains unchanged, and the limit only changes the proportion of first-period to second-period spending. Thus, it surprisingly induces no change in taxes.

Again, fiscal governance indices may provide insufficient information. First, the initial Vote not necessarily decreases spending level, or tax rate. Second, each nominal limit addresses a different policy bias. Deficit limit helps to reduce first-period spending, but not taxes, so it may be help in restoring intertemporal social optimum. In Cobb-Douglas Model, however, it is toothless for intratemporal inefficiency such as common-pool problem, because it imposes no constraint on excessive public sector. Spending limit, in contrary, reduces both spending and taxes, so it solves for excessive spending, but not for intertemporal inefficiency. The limits are not equivalent; a theoretical purist wouldn't mix them in a single index.

For quasi-linear utility, the caps were used only by moderately conservative coalitions. In Cobb-Douglas Model, the use of limits is more sophisticated. Spending limit is more likely used in a coalition of conservatives, regardless of probability of survival (indicated, for example, by age). We may say that Spending limit is in common interest of young and old conservatives. The Relative deficit limit is more used by coalitions of legislators with large probability of survival, regardless of preference for total size of the budget. In other words, this limit is in common interest of young legislators, both conservative and public-good loving ones. Thus, young conservatives are those who tend to use the constraints at most, but which constraint they use depends on with whom they create a coalition.

2.4 Conclusion

Because the public good is single and homogenous, the legislative process we have modeled does not suffer from allocation inefficiencies of the common-pool type. The problem is the pure conflict of interest that cannot be overcome under given flat tax system and in the absence of compensations.

We have shown that this conflict is settled in the absence of procedural fiscal rules and enforceable coalitional commitments. The outcome is stable regardless of who holds executive power, because in the legislature, legislators in opposition can perfectly exploit differences within the ruling coalitions when voting in separate stages. When some nominal fiscal rule is present, the outcome is coalition-dependent. Spending limit groups interests of conservative legislators, and Deficit limit induces future-oriented legislators cooperate. Shocks into composition of spending thus translate into shocks into fiscal policy. When coalitions appear randomly (we don't have a dynamic model allowing for endogenous coalition-making), the fiscal governance implies macroeconomic volatility. Moreover, the volatility may not decrease when broader coalitions (by increased majority quota) are required. And last, but surely not least, fiscal governance invites strategic delegation effect.

Policy implications of fiscal governance are thus twofold. In common-pool problems, fiscal governance works as a coordination device; in pure conflicts of fiscal preferences, it is a protection device. This ambiguity implies that strong fiscal governance may both eliminate and magnify fiscal costs of political fragmentation. The fact that multi-party governments find it easier to rely on Commitment mode, as Hallerberg (1999) observes, not necessarily means motivation for efficiency, but for mutual protection. Also notice that common measures of fiscal governance may be misleading, because the limits solve different biases in different specifications.

To conclude: in political economy where coalitions cannot commit legislators, higher fiscal governance improves fiscal policy only if macroeconomic volatility is not too costly, and if shocks into composition of coalitions are not too large.

References

[1] Balassone, F. and R. Giordano (2001), "Budget Deficits and Coalition Governments", *Public Choice*, 106, 327–349.

[2] Baron, D. & J. Ferejohn (1989). "Bargaining in Legislatures", *American Political Science Review*, 83, 1181–1206.

[3] De Haan, J., Moessen, W., and B. Volkerink (1999), "Budgetary Procedures—Aspects and Changes: New Evidence for Some European Countries", in Poterba, J. M. and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago and London: The University of Chicago Press, 265-299.

[4] Dharmapala, D. (2002). "The Congressional Budget Process and the Aggregate Level of Spending", University of Connecticut Working Paper 2002-13.

[5] Dharmapala, D. (2003). "Budgetary Policy with Unified and Decentralized Appropriations Authority", *Public Choice*, 115, 347–367.

[6] Ehrhardt, K. et al. (2000). "Budget Process: Theory and Experimental Evidence", *ZEI Working Paper*, B 18/2000.

[7] Ferejohn, J., and K. Krehbiel (1987), "The Budget Process and the Size of the Budget", *American Journal of Political Science*, 31(2), 296-320.

[8] Filc, G. & C. Scartascini (2004). "Budget Institutions and Fiscal Outcomes: Ten Years of Inquiry on Fiscal Matters at the Research Department". IADB Working Paper, Inter-American Development Bank, Office for Evaluation and Oversight.

[9] Gleich, H. (2003). "Budget institutions and fiscal performance in Central and Eastern European Countries", ECB Working Paper, European Central Bank.

[10] Gregor, M. (2005). "Committed to Deficit: The Reverse Side of Fiscal Governance," IES Working Paper 88/2005. Institute of Economic Studies, Charles University, Prague.

[11] Hallerberg, M. (2004). *Domestic Budgets in a United Europe: Fiscal Governance from the End of Bretton Woods to EMU.* Ithaca, NY: Cornell University Press.

[12] Hallerberg, M. & P. Meier (2004). "Executive Authority, the Personal Vote, and Budget Discipline in Latin American and Carribean Countries", *American Journal of Political Science.*

[13] Hallerberg, M. (1999), "Electoral Institutions, Cabinet Negotiations, and Budget Deficits in the European Union", in Poterba, J. M. and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago: The University of Chicago Press, 209-232.

[14] Imbeau, L. M. (2000), "The Political Economy of Public Deficits", in Imbeau, L. M. and F. Pétry (eds.), *Politics, Institutions, and Fiscal Policy: Public Deficits and Surpluses in Federated States*, Lanham, MD: Lexington Books, 1-19.

[15] Krafchik, W. & J. Wehner (1998). "The Role of Parliament in the Budgetary Process", *South African Journal of Economics*, 66, 4.

[16] Lienert, I. (2005). "Who Controls the Budget: The Legislature or the Executive?", IMF Working Paper WP/05/115.

[17] Niepelt, D. (2006). "Starving the beast? Intra-generational conflict and balanced budget rules", *European Economic Review*, forthcoming.

[18] Organisation for Economic Cooperation and Development (2002). "The OECD Budgeting Database", *OECD Journal on Budgeting*, 1 (3), 155–171.

[19] Organisation for Economic Cooperation and Development & World Bank (2003). "Results of the Survey on Budget Practices and Procedures". Online.

[20] Oppenhheimer, B. I. (1983). "How Legislatures Shape Policy and Budgets", *Legislative Studies Quarterly*, 8, 4.

[21] Osborne, M. J. & A. Rubinstein (1990). *Bargaining and Markets*. New York: Harcourt Brace Jovanovich, Academic Press.

[22] Persson, T. & G. Tabellini (2000). *Political Economy: Explaining Economic Policy*. Cambridge, Mass.: MIT Press.

[23] Poterba, J. M. & J. von Hagen, eds. (1999). *Fiscal Institutions and Fiscal Performance*. Chicago and London: The University of Chicago Press.

[24] Schick, A. (2002). "Can National Legislatures Regain an Effective Voice in Budget Policy?", *OECD Journal of Budgeting*, 1, 3.

[25] Serritzlew, S. (2005). "The Perverse Effect of Spending Caps", *Journal of Theoretical Politics*, 17,1,75–105.

[26] Terai, K. (2003). "Electoral Alliance and Implemented Redistribution: An Interpretation of Non-Competitive Politics of Japan", *Applied Economics Letters*, 10, 235–238.

[27] von Hagen, J. (1992). "Budgeting Procedures and Fiscal Performance in the European Community", Economic Paper No. 96, Commission of the European Communities.

[28] von Hagen, J. (2005). "Political Economy of Fiscal Institutions", GESY Discussion Paper No. 149, University of Mannheim.

[29] Wehner, J. (2005). *Cross-National Variation in Legislative Budgeting.* London: London School of Economics, forthcoming.

[30] Wyplosz, C. (2005). "Fiscal Policy: Institutions versus Rules", *National Institute Economic Review*, 191, 1, 64–78.

[31] Yläoutinen, S. (2004). "Fiscal Frameworks in the Central and Eastern European Countries", Discussion Paper 72, Ministry of Finance, Finland.

3. WEAKEST-LINK PUBLIC GOODS: STRATEGIC DELEGATION IN DECENTRALIZATION AND CENTRALIZATION

3.1 Introduction

Non-cooperative decentralization induces rich forms of strategic behavior. The very standard common-pool incentives arise when district representatives tend to exploit each other in a struggle for budgetary commons (Weingast, Shepsle & Johnsen 1981). Centralization of commons, through coordination or bargaining devices, may render governance immune to this exploitation. Besley and Coate (2003) nonetheless demonstrated that the remedy is often imperfect; even if representatives promote a cooperative solution, non-cooperative incentives remain between voters of the districts. Voters thus keep exploiting the other regions by misrepresenting preferences in an electoral game, which is known as the strategic delegation effect. Centralization then fulfills its premises incompletely.

The strategic delegation effect arises mostly when local public goods have spillovers (positive externalities) in other regions. A spillover emerges when public good in one region contributes to the public goods in the other regions, typically incompletely, i.e. one unit in the source region creates ψ units in the beneficiary region, where $\psi < 1$. Dur and Roelfsema (2005) found that for neutral public goods and strategic substitutes, this asymmetry invites misrepresentation of preferences, both in decentralized and centralized setting.

Not much interest has been paid to more extreme types of spillovers, however. Marginal spillovers may be discontinuous, as there can be various complementarities between local public goods. Public economics tackles this problem by constructing a total public good, aggregated from local public goods, or local contributions. There is a vast array of technologies of aggregation: besides substitutes (summation technologies), we may have best-shot, better-shot, weaker-link, and weakest-link technologies.

In this chapter, we keep focus on the weakest-link technologies. For the weakest-link technologies, marginal contributions are step-wise functions. Unless the amount of local public good reaches certain level, the marginal contribution is strictly positive and constant; once the threshold is passed, the marginal contribution falls to zero—complements are missing. In the fully symmetric case, the unique decentralized Nash equilibrium is Pareto-optimal, so the weakest-link aggregation loses theoretical appeal. The real issue comes on stage with asymmetries. Sandler and Vicary (2001) as well as Vicary and Sandler (2002) study how regions provide money and in-kind transfers in the case of asymmetries. Grossman (2003) interprets public good provision as implicit redistribution, and points to specific incentives for the provision of weakest-link public goods. Bardhan *et al.* (2002) analyze the effects of inequality in private inputs that are complementary in production with collective inputs. On one hand, higher inequality leads to a relative decrease in free-riding as one agent becomes the predominant provider, while on the other hand, higher inequality leads to lower social welfare due to the concavity of agents' preferences.

Herein, we focus not only on the welfare properties of equilibrium provision, but also on the incentives to delegate strategically. This is important for understanding of international or inter-regional policy games, and the dynamics of integration.

We devote Section 1 to argument for the relevance of weakest-link technologies; one of the reasons is that best responses in far more frequent weaker-link technologies may resemble best responses in weakest-link case, provided that pure self-interest is abandoned. Section 3.2 classifies asymmetries in weakest-link technologies and seeks empirical counterparts for each type of technology. In Section 3.3, we build up a framework for all types of asymmetries. Section 3.4 is devoted to the benchmark model with full symmetry, where no strategic effects exist. Sections 3.5 and 3.6 analyze the role of asymmetries in access and cost, and Section 3.7 examines the situation when regions with the strictly lowest contribution are penalized. The concluding Section 3.8 overviews strategic behavior in all models, and draws inferences on effects of non-cooperative centralization.

3.2 Motivation

3.2.1 Extremeness of the weakest link aggregation

Aggregation functions of local public goods can be interpreted as production functions; local public goods contribute to (virtual or real) total public good like factors of production contribute to total output. Constant-elasticity-of-substitution (CES) aggregation functions, established by Arrow and Enthoven (1961), are standard in this realm.

Definition 1 (CES production function): For vector of factors $\mathbf{q} = (q_i)$, where $i \in \{1, \ldots, n\}$, and for aggregation parameter v, where $v \in R$, the CES-production function writes

$$G = \left[\frac{1}{n}\sum_{i=1}^{n} q_i^v\right]^{\frac{1}{v}}.$$

The advantages of CES functions/aggregations are threefold: (i) CES technology is described only by single aggregation parameter, v, (ii) it exhibits constant returns to scale, and (iii) it has constant elasticity of substitution, $1/(1-v)$. For $v = 1$, we have the case of pure substitutes (summation technology), $-\infty < v < 1$ represent weaker-link technologies, and $1 < v < \infty$ represent better-shot technologies. Lemma 5 proves that weakest-link technology is an extreme case of $v = -\infty$.

Lemma 5: A CES aggregation technology with $v = -\infty$ is a weakest-link technology, $G = \min\{q_1, q_2, \ldots, q_n\}$.

Proof Let $\hat{g} := \min\{g_1, g_2, \ldots, g_n\}$. From Assumption 1 and $v = -\infty$,

$$G = \lim_{v\to-\infty}[q_1^v + \ldots + g_n^v]^{\frac{1}{v}} = e^{\lim_{v\to-\infty}[\frac{1}{v}\ln(q_1^v+\ldots+g_n^v)]}. \tag{3.1}$$

Rewriting the argument of the exponential function, we get

$$\lim_{v\to-\infty}\left[\frac{1}{v}\ln(q_1^v + \ldots + g_n^v)\right] \overset{\text{L'Hôpital}}{=} \lim_{v\to-\infty}\frac{\ln q_1 q_1^v + \ldots + \ln q_n q_n^v}{q_1^v + \ldots + q_n^v} =$$

$$\overset{\forall m\in\mathbb{N}}{=} \lim_{v\to-\infty}\frac{\ln q_1^{m+1} q_1^v + \ldots + \ln q_n^{m+1} q_n^v}{\ln q_1^m q_1^v + \ldots + \ln q_n^m q_n^v} =$$

$$= \lim_{v\to-\infty}\frac{\ln q_1\left(\frac{\ln g_1}{\ln\hat{g}}\right)^m\left(\frac{g_1}{\hat{g}}\right)^v + \ldots + \ln q_n\left(\frac{\ln g_n}{\ln\hat{g}}\right)^m\left(\frac{g_n}{\hat{g}}\right)^v}{\left(\frac{\ln g_1}{\ln\hat{g}}\right)^m\left(\frac{g_1}{\hat{g}}\right)^v + \ldots + \left(\frac{\ln g_n}{\ln\hat{g}}\right)^m\left(\frac{g_n}{\hat{g}}\right)^v} = \ln\hat{g}. \tag{3.2}$$

In the last part, we used that $\forall q_i \neq \hat{q} : \lim_{v\to-\infty} (g_i/\hat{g})^v = 0$ and $\forall q_i = \hat{q} : \lim_{v\to-\infty} (g_i/\hat{g})^v = 1$. Finally we put the solution in (3.2) into the exponent of (3.1) and get $G = \hat{g} = \min\{g_1, g_2, \ldots, g_n\}$. □

The fact that weakest-link aggregation is associated with the limit case of the aggregation parameter, $v = -\infty$, gives rise to intuition that this type of aggregation may be too extreme in the real world. Moreover, Barro & Sala-i-Martin (1999, p. 46) proved that the weakest-link aggregation used as a production function (known as the fixed-proportions technology, cf. Leontief 1941) may yield paradoxical results associated with the Harrod-Domar controversy.

Thirdly, Cornes (1993, p. 262) shows that for $-\infty < v \leq 1$, identical unique Nash equilibrium always occurs. Unlike that, the weakest-link technology gives multiple equilibria. It is a very extreme case of zero marginal products of certain local public goods; once strictly positive marginal product occur, all particular properties of weakest-link technology disappear and the identical Nash equilibrium is restored.

Even though, there are at least three tentative reasons why to focus on this technology:

1. If the Nash equilibrium is to be achieved in evolutionary setting with boundedly rational individuals, the problem with result by Cornes (1993) is that speed of adjustment towards the equilibrium may strikingly differ for each technology. This is especially when it is costly to play something else than imitation. The weaker is the marginal spillover (or marginal product), the closer are the best-response functions to matching strategy in n-dimensional space. The weakest-link case is then just a special situation when complete matching (imitation) is always optimal; for weaker links, incomplete matching is the best reply. For technologies with extremely "weak links" (low marginal product), temptation for deviation thus will be very weak and we may expect that out-of-equilibrium behavior preserves for longer time.

2. Moreover, if agents are to some extent altruistic (or if equity is argument of the utility function), temptation to free ride may disappear for technologies with aggregation parameters $v \in (-\infty, \overline{v}\rangle$, where $\overline{v} > -\infty$. The technologies produce imitation as an optimum best response, resembling thus weakest-link technology in purely self-interest setting.

3. Similarly, the weak temptation may disappear if the agents have non-zero propensity to reciprocate.

Any deviation from strict self-interest (altruism, satisfaction from reciprocation, cost of calculation) may significantly enlarge relevance of weakest-link technologies. Consider the following example:

CES-function with and without drift Assume the Cobb-Douglas utility function $U(\cdot) = y_i^{\beta}Q$, two players with initial endowments m_1, m_2, price p per unit of q_1 and q_2, and $v = 0$, i.e. $Q = \sqrt{q_1 q_2}$. The best responses are constant, $q_1^*(q_2) = q_2^*(q_1) = m_1/3p$. Suppose that the players exhibit "equity drift"; each player suffers penalty if $q_1 \neq q_2$. The penalty is linear in $|q_1 - q_2|$ and enters utility function as an equivalent of income loss; for convenience, say each unit of difference represents an income loss of $p/2$ (if the player raises q_1 above q_2, she pays unit-price $3p/2$, not only p). Hence, $U' = (y - p/2|q_1 - q_2|)^{\beta}Q$. To get best response function $q_1^*(q_2)$, we firstly get the best reply in subset of strategies $q_1 \leq q_2$. Here, $q_1^*|_{q_1 \leq q_2} = \min\{4m_1/9p - q_2/9, q_2\}$. In subset of strategies $q_1 \geq q_2$, we have $q_1^*|_{q_1 \geq q_2} = \max\{4m_1/10p - q_2/10, q_2\}$. This yields that in the neighborhood of $m_1/3p$, namely for $q_2 \in \langle m_1/4p, m_1/2p \rangle$, imitation is the best response: $q_1^*(q_2) = q_2$. Symmetrically, $q_2^*(q_1) = q_1$. As a result, we have infinite amount of equilibria, which is a major change to the case without "equity drift" with super-stable (constant) best responses, hence unique equilibrium. Figure 3.1 illustrates how drift in the utility function affects best responses.

3.2.2 Complementarities

A total public good, aggregated by weakest-link technology, is but a special type of complementarity, when local public goods are perfect complements. This needs not always be the case. Generally, we can classify complementarities between local public goods as follows: Can we argue that a single total good is produced, or, in fact, two different region-specific goods emerge from local public goods? For single total good, is access to its consumption symmetric or not? Is there a penalty for the provider of the strictly weakest link? Are costs of provision symmetric or not? These questions suggest that differences among the complementarities stem from four sources:

1. Single or region-specific good. A single good describes traditional complementary goods: each soldier needs only one kalashnikov, a university hosts only one university library, space ship can carry only a given number of astronauts, a good journal paper typically requires two

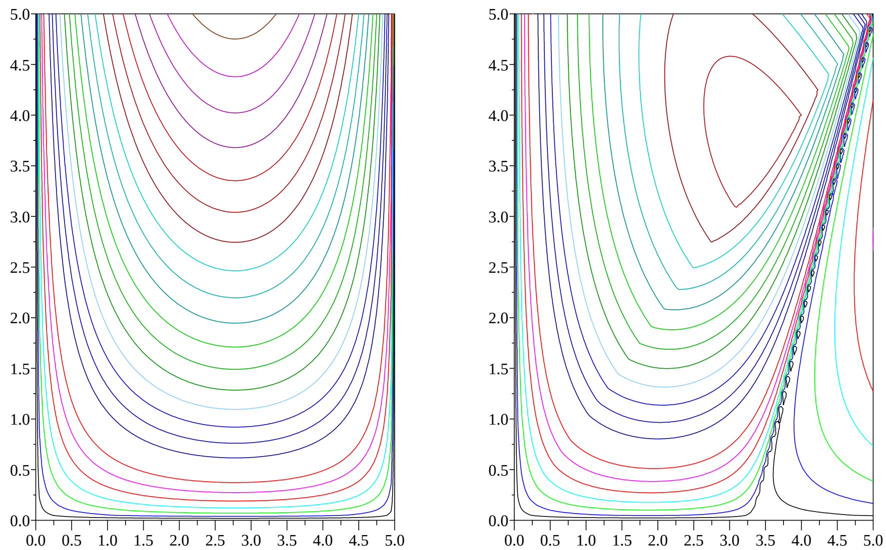

Fig. 3.1: CES function with and without drift $(v = 0, \beta = \frac{1}{5}, m_1 = 5, p = 1)$

authors, one theorist and one econometrician. Region-specific goods emerge for provisions against complementary bads, which are region-specific. An example of money laundering is the most illuminating: Suppose mafias in two countries need to launder money in a country with looser financial regulation, yet transfer of money across borders can be costly. Thus, mafia in Region 1 will stay in Region 1 unless regulation in Region 2 is sufficiently looser than regulation in Region 1. The reverse holds for mafia in Region 2. Penetration of mafias is region-specific, yet there exist complementarities.

2. Access. One of regions may be less able to use the total public good, either due to incompetence (small countries hardly exploit gains from space research), or due to delays of delivery when total public good has to be located in one place (transportation of patients is often costly for serious injuries).

3. Penalty. The region with the strictly lowest amount can suffer a penalty. This is best interpreted as a penalty for hosting a public bad. For floods, we can imagine that the first strike of water may be more damaging than if floods are coming from neighboring regions. For drug traffic, if drugs go from outside, the region with looser border control may face higher supply. For money laundering, if financial regulation in the region is sufficiently attractive for mafia from the other region, the region will face additional flow of black market activities.

4. Costs. Violation of law of one price mostly reflect immobility of costs of production. Extra surcharge may occur when the local good (expensive hospital equipment) has to be transported to the place of construction of total public good (hospital).

Asymmetries can draw from entirely different reasons than for pure substitutes. For substitutes, spillovers were interpreted by less frequent use, costlier use, or limited access, all given by spatial differences, or perhaps cultural or administrative barriers. Interpretation of asymmetries for weakest-link technologies is less obvious. In general, we have two possibilities. In the first case, we may suppose technologies leading to single total public good, and study differences stemming from different access to the total good, or different costs of provision. Also, there may be penalty for the strictly weakest-link. A classic case is dike maintenance (Hirshleifer 1983), concerning a circular region, of which perimeter is marked by a dike. Floods may penetrate the weakest point of the dike and cause damage equally throughout the region. Penalty occurs when the damage isn't equal.

In another case, the assumption of single total public good is abandoned. In fact, two different goods are produced, connected by complementarities. This corresponds to situation when regions struggle with public bads, located inwards (e.g., imperfectly mobile money laundering). We leave this interesting case of complementarities for further research.

Table 3.1 illustrates several cases of complementary aggregation. The examples indicate that complementarities need not to be associated with natural events, e.g. forest-fire suppression, or flue-disease. In contrary, it often corresponds to technologies of protection against common public bad that enters both regions via the weakest link (e.g., mafia smuggling drugs into free-border area, or terrorists loading contaminated bag into the least diligent air carrier, cf. Heal and Kunreuther 2005). Moreover, technology is often partly endogenous and weakest-link technologies difficult to trace. Out of four weakest-link technologies provided by Sandler (2006, p. 12), we can see that in at least three cases, technologies with different substitution (mostly weaker-link) can replace the weakest-link technology. Integrity of Internet network, for example, is enhanced by Internet protocols (e.g., TCP), which split information into packets and send the packets through different channels. Then, speed of connection is not necessarily determined by the weakest link. Moreover, for surveillance of regional disease outbreaks, we cannot say that risks of outbreak are fully correlated (disease is not traveling across countries, seeking for the least protected targets, like mafia, flood, or foreign army does). Also air traffic control is weakest-link technology only as long as risks are correlated, which is doubtful.

Tab. 3.1: Complementarities: selected examples

Total public good	Local public good	access	penalty
Sea flood protection	dikes	s	yes
Terrorism	intelligence	s	yes
Rapid force squad	soldier, weapon, ammo	s/a	no
Higher education	lecturer, building	s/a	no
Journal paper	theorist, econometrician	s	no
Space research	spacecraft, astronaut	s/a	no
Drug traffic	border controls	s	yes/no
Money laundering	financial regulation	s	yes

3.3 Model

3.3.1 Setup

Suppose Region 1 produces local public good x, where $x \geq 0$, and Region 2 provides local public good y, where $y \geq 0$. The public goods are complements, linked by a symmetric weakest-link technology, which yields total public good

$$G = \min\{x, y\}. \tag{3.3}$$

Access to total public good G in each region in general need not to be perfect. We express the rates of access by $\psi_1 \in [0, 1]$ and $\psi_2 \in [0, 1]$. Assume further that each region finances only his good using a lump sum tax, where the unit cost of x per citizen or Region 1 is c_x and the unit cost of y per citizen of Region 2 is c_y. Citizens differ only in propensity for public good consumption $\lambda > 0$, so propensity of citizen i from Region 1 (or citizen j from Region 2) is denoted λ_1^i (or λ_2^j). We impose that both λ_1 and λ_2 have symmetric distribution, normalized around 1. Voters with $\lambda = 1$ are, in the dimension λ, median voters; we shall use superscript m for their preferences and voting strategies (e.g., $\lambda^m = 1$). Utility of any citizen in any region is supposed to be quasi-linear, where $b(\cdot)$ is a standard twice differentiable concave function, $b(0) = 0, b' > 0, b'' < 0$. For any citizen i from Region 1 and any citizen j from Region 2, we write

$$U_1^i = \lambda_1^i b(\psi_1 G) - c_x x, \qquad U_2^j = \lambda_2^j b(\psi_2 G) - c_y y. \tag{3.4}$$

Their first-best (unconstrained) optima write as

$$(\hat{x}_1^i, \hat{y}_1^i) = \arg\max_{x,y \geq 0} U_1^i, \qquad \hat{G}_1^i := G(\hat{x}_1^i, \hat{y}_1^i), \tag{3.5}$$

$$(\hat{x}_2^j, \hat{y}_2^j) = \arg\max_{x,y \geq 0} U_2^j, \qquad \hat{G}_2^j := G(\hat{x}_2^j, \hat{y}_2^j). \tag{3.6}$$

The game proceeds as follows. In Stage 1, voters elect delegates, one per region. The delegates are citizen candidates with propensities for public good consumption denoted as λ_1^d and λ_2^d, hence are distinguished from other citizens by label d. We assume that election in each region always picks up the Condorcet citizen-candidate. In Stage 2, delegates determine x and y, by maximization of U_1^d and U_2^d, respectively (no political rent). The way local public good levels are determined is dependant on the mode of governance, decentralization or centralization.

The first-best optima in (3.5) and (3.6) can be used to derive Lemma 6, which intuitively states that a voter can affect the level of its local public good (and also G) by strategic delegation; of course, only as long as his delegate participates in the decision about that level.

Lemma 6: The more conservative delegate in Region 1 (or 2), the less local public good $\hat{x}_1^d$ (or $\hat{y}_2^d$) she prefers to provide, $d\hat{x}_1^d/d\lambda_1^d > 0$, or $d\hat{y}_2^d/d\lambda_2^d > 0$.

Proof This is given by the implicit function theorem from the following first-order conditions and requirements imposed on $b(\cdot)$:

$$b'(\psi_1\hat{x}_1^d) = \frac{c}{\lambda_1^d\psi_1} \Longrightarrow \frac{d\hat{x}_1^d}{d\lambda_1^d} = -\frac{b'(\psi_1\hat{x}_1^d)}{\lambda_1^d\psi_1 b''(\psi_1\hat{x}_1^d)} > 0,$$

$$b'(\psi_1\hat{y}_2^d) = \frac{c}{\lambda_2^d\psi_2} \Longrightarrow \frac{d\hat{y}_2^d}{d\lambda_2^d} = -\frac{b'(\psi_2\hat{x}_2^d)}{\lambda_2^d\psi_2 b''(\psi_2\hat{x}_2^d)} > 0. \quad \square$$

In similar setup, we can significantly simplify the game into the game of median voters of both regions. Utility function is quasiconcave in x (or y), and by Lemma 6, x is monotonic in λ_1 and y is monotonic in λ_2. Besley and Coate (2003) show that this makes conditions for the existence of Condorcet winner satisfied, and the Condorcet winners are median voters. As a result, in all what follows, we shall solve the voting in Stage 1 as a two-person subgame of median voter in Region 1 and median voter in Region 2.

3.3.2 Non-cooperative decentralization

Each delegate is allowed to set the level of public good in her region independently from the decision of the other delegate; decentralization is a classic non-cooperative game of two regions. As delegates are citizen candidates without access to any extra rent from power, they derive best responses $\tilde{x}(y)$ and $\tilde{y}(x)$ from optimization of utility with respect to decision of the other delegate

$$\tilde{x}(y) := \arg\max_{x\geq 0} U_1^d(x,y) = \arg\max \lambda_1^d b[\psi_1 G(x,y)] - c_x x,$$

$$\tilde{y}(x) := \arg\max_{y\geq 0} U_2^d(x,y) = \arg\max \lambda_2^d b[\psi_2 G(x,y)] - c_y y.$$

In this non-cooperative setup, a pair (x,y) is stable if and only if it is a Nash equilibrium, $x = \tilde{x}(y)$ and $y = \tilde{y}(x)$.

3.3.3 Non-cooperative centralization

In centralization, a centralized body decides on G. We consider only the simple case of centralization, in which Nature in the beginning of Stage 2 equiprobably selects one of the delegates to control the body. The ruling delegate has full control over both x and y, so she is able to secure her first-best optimum, $(\hat{x}_1^d, \hat{y}_1^d)$, or $(\hat{x}_2^d, \hat{y}_2^d)$. In this mode of governance, the non-ruling delegate has no control, and the ruler has no incentive to take interests of non-ruler into account. For only Nature determines the ruler, no strategic aspect among delegates is present.

3.3.4 Price elasticity of real public good demand

When access to total public good is imperfect, we can distinguish between demand for total amount of the public good, $\hat{G}^i$, and demand for the actually accessed share of the good, $\psi\hat{G}^i$. We call the former as *nominal* public good demand, and the latter *real* public good demand.

Price elasticity of the demand for real public good plays a crucial role when access to consumption is asymmetric. Consider two identical-type (λ^i) citizens coming from different regions, who enjoy asymmetric access but bear symmetric cost $(1 = \psi_1 > \psi_2 > 0$ and $c := c_x = c_y)$. The first-order conditions for the Region 1 citizen satisfy $b'(\hat{x}_1^i) = c/\lambda_1^i$ and $\hat{y}_1^i \geq \hat{x}_1^i$; for the latter, $b'(\psi\hat{y}_2^i) = c/\psi\lambda_2^i$ and $\hat{x}_2^i \geq \hat{y}_2^i$.

Optimal total public good levels satisfy $b'(\hat{G}_1^i) = \psi b'(\psi\hat{G}_2^i) = c/\lambda^i$. Now, the question is whether diminished access in Region 2 reduces or boosts nominal demand for the total public good. In other words, we ask whether $\hat{G}_2^i > \hat{G}_1^i$ or $\hat{G}_2^i < \hat{G}_1^i$. Lemma 7 explains that the answer depends on price elasticity of real public good demand. A higher than unit elasticity yields standard result, namely that worse access decreases (nominal) demand for total public good (here $\hat{G}_2^i < \hat{G}_1^i$), and vice versa. Contrarily, inelastic demand yields that diminishing access rather increases demand. The difference goes via substitution effect: in standard elastic case, substitution effect is stronger than income effect (e.g. in space research), whereas inelastic marginal benefit implies that income effect dominates substitution effect (e.g. in defense).

Lemma 7: For each citizen of type λ^i, if real demand for total public good $\psi\hat{G}^i(\cdot)$ is more than unit-elastic, worse access to total public good decreases nominal demand for total public good, $\hat{G}^i$. If otherwise, worse access to total public good increases demand for nominal total public good:

$$\forall \lambda^i : \begin{cases} \varepsilon(\psi \hat{G}^i, b'(\psi \hat{G}^i)) < -1 \iff \dfrac{d\hat{G}^i}{d\psi} > 0 \\ \varepsilon(\psi \hat{G}^i, b'(\psi \hat{G}^i)) = -1 \iff \dfrac{d\hat{G}^i}{d\psi} = 0 \\ \varepsilon(\psi \hat{G}^i, b'(\psi \hat{G}^i)) > -1 \iff \dfrac{d\hat{G}^i}{d\psi} < 0 \end{cases}$$

Proof For each citizen of type λ^i, and for all $\psi \in [0,1]$, we solve for individual optima in (3.5) and (3.6), and receive $\psi b'(\psi \hat{G}^i) = c/\lambda^i = \text{const.}$ Therefore, we have implicit nominal demand $\hat{G}^i(\psi)$ for all $\psi \in [0,1]$. By implicit function theorem

$$\frac{d\hat{G}^i}{d\psi} > 0 \iff -\frac{b'(\psi \hat{G}^i) + \psi \hat{G}^i b''(\psi \hat{G}^i)}{\psi^2 b''(\psi \hat{G}^i)} > 0 \iff b'(\psi \hat{G}^i) > -\psi \hat{G}^i b''(\psi \hat{G}^i)$$
$$\iff \frac{\frac{db'(\psi \hat{G}^i)}{d\psi \hat{G}^i}}{\frac{b'(\psi \hat{G}^i)}{\psi \hat{G}^i}} = \varepsilon(b'(\psi \hat{G}^i), \psi \hat{G}^i) > -1 \iff \varepsilon(\psi \hat{G}^i, b'(\psi \hat{G}^i)) < -1.$$

In the last equivalence, we use that for monotonic $y = y(x)$, $y'(x)x'(y) = 1$ and therefore $\varepsilon(y,x) = y'x/y = x/x'y = 1/\varepsilon(x,y)$. The other equivalences can be received by analogy. □

How to interpret the real public good demand and price elasticity of thereof? The citizen i with imperfect access has utility $U^i = \lambda^i b(\psi G) - cG$. Optimal amount of real public good always satisfies $\lambda^i b'(\psi G) = c/\psi$; in other words, marginal benefit of additional unit of real public good equals marginal cost. Decrease of ψ always increases marginal cost of real public good consumption. By assumption, $b'(\cdot) > 0$ and $b''(\cdot) < 0$, so real public good consumption ψG has to decrease. So, although imperfect access works as an additional tax on both nominal and real public good consumption, responses of nominal and real public good consumption may differ.

3.4 Full symmetry

The traditional weakest-link technology is defined by symmetric weakest-link technology, full access of both regions to total public good, and equal cost of provision. Both regions are symmetric in all respects in all stages, so we may call this full symmetry. Consider for illustration airport security controls aiming at illegal drug imports into Schengen area. If only the quality of controls matters for the smugglers, if security costs are roughly the same, and drugs within the Schengen area are perfectly mobile, assumptions of full symmetry are satisfied.

Definition 2 (Full symmetry): The fully symmetric weakest-link technology is defined by $G = \min\{x, y\}$, $\psi_1 = \psi_2 = 1$ and $c := c_x = c_y$.

We firstly derive the first-best optima for all citizens, using the weakest-link property. For citizen i from Region 1, the individual first-best optimum satisfies $b'(\hat{x}_1^i) = c/\lambda_1^i$ and $\hat{y}_1^i \geq \hat{x}_1^i = \hat{G}_1^i$. For citizen j from Region 2, $b'(\hat{y}_2^j) = c/\lambda_2^j$ and $\hat{x}_2^j \geq \hat{y}_2^j = \hat{G}_2^j$.

In decentralization, we need best responses of delegate in Region 1, i.e. $\tilde{x}(y)$. We apply the key property of the weakest-link technology: any player can always reduce the amount of total good below the level proposed by the other player. The amount of total good is determined by *the lowest common denominator* of both players. Here, Region 1 effectively chooses $\tilde{x}(y) \in \langle 0, y \rangle$. On this interval, he matches ($G = \tilde{x}(y) = y$) if y is too low, and sets optimal total public good $G = \tilde{x}(y) = \hat{x}_1^d = \hat{G}_1^d$ if y is high:

$$y < \hat{G}_1^d \Longrightarrow \tilde{x}(y) = y \Longrightarrow G = y \tag{3.7}$$

$$y \geq \hat{G}_1^d \Longrightarrow \tilde{x}(y) = \hat{x}_1^d \Longrightarrow G = \hat{x}_1^d \tag{3.8}$$

Due to full symmetry, best responses of delegate in Region 2 are symmetric:

$$x < \hat{G}_2^d \Longrightarrow \tilde{y}(x) = x \Longrightarrow G = x \tag{3.9}$$

$$x \geq \hat{G}_2^d \Longrightarrow \tilde{y}(x) = \hat{y}_2^d \Longrightarrow G = \hat{y}_2^d \tag{3.10}$$

In other words, if the other region provided less than the optimum, the delegate would imitate the proposal. Less would decrease the level of total public good (decrease of the weakest link), and more would imply wasteful expenditures. If the other region provided at least the optimum, the delegate's optimum could be secured by controlling the weakest link. The delegate would thereby determine her optimal level of local public good as well as total public good.

The symmetric best responses provide with multiple symmetric Nash equilibria in which $G = x = y$, where $G \in [0, \min\{G_1^d, G_2^d\}]$. Using pre-play announcement as a plausible refinement, only one (Pareto-efficient) equilibrium remains, in which the total public good is determined by the preference of the more conservative candidate,

$$x = y = G = \min\{G_1^d, G_2^d\}. \tag{3.11}$$

Voters, anticipating this particular equilibrium to come, consider benefits from deviations from sincere delegation. We use that median voters are Condorcet winners, and to investigate effect of strategic delegation, we only consider deviations of median voters in both regions. Accordingly, we solve a sub-game of two players with preferences normalized to 1, who decide on λ_1^d, respectively λ_2^d. The best-response correspondence for median voter in Region 1, $\tilde{\lambda}_1^d(\lambda_2^d)$, satisfies

$$\lambda_2^d < 1 \Longrightarrow \tilde{\lambda}_1^d(\lambda_2^d) \geq \lambda_2^d, \tag{3.12}$$

$$\lambda_2^d \geq 1 \Longrightarrow \tilde{\lambda}_1^d(\lambda_2^d) = 1. \tag{3.13}$$

Full symmetry gives analogical best responses of Region 2 median voter,

$$\lambda_1^d < 1 \Longrightarrow \tilde{\lambda}_2^d(\lambda_1^d) \geq \lambda_1^d, \tag{3.14}$$

$$\lambda_1^d \geq 1 \Longrightarrow \tilde{\lambda}_2^d(\lambda_1^d) = 1. \tag{3.15}$$

In short, each median voter is pondering the following: If the other region proposes a too conservative candidate, in our region we have to accept that the other region will control the weakest link. Any equally or less conservative candidate we delegate will have to imitate proposal of the more conservative candidate. Given that the weakest link is controlled by the other region, we can achieve no better but this (second-)best outcome. If, however, the other region proposes a candidate who is willing to spend even more than we require in our first-best optimum, we propose our optimal candidate, who becomes the more conservative of the two, and thereby keeps the weakest link under control. As a result, we arrive at our first-best optimum.

Like in subgame of delegates, this leads to multiple symmetric Nash equilibria, $\lambda^d = \lambda_1^d = \lambda_2^d$, where $0 < \lambda^d \leq 1$. Once again, only the upper bound equilibrium remains when pre-play announcement is allowed, so the subgame yields $\lambda_1^d = \lambda_2^d = 1$.

Proposition 1: In decentralization with full symmetry, no conflict of interests arises between median voters. By delegating sincerely, $\lambda_1^d = \lambda_2^d = 1$, median voters secure individual first-best optima $\hat{G}^m = \hat{G}_1^m = \hat{G}_2^m$.

Proof Solving Stage 2, we firstly derive best responses of delegates from the first-best optima, and the result is in (3.7) and (3.9). Equilibrium in Stage 2 is determined by the first-best optimum of the more conservative delegate, $x = y = G = \min\{G_1^d, G_2^d\}$. Anticipating this equilibrium, we solve backwards Stage 1. Best responses of median voters are in (3.12) and (3.14), and the multiple Nash equilibria satisfy $\lambda^d = \lambda_1^d = \lambda_2^d$, where $0 < \lambda^d \leq 1$. With pre-play announcement, we have the equilibrium with $\lambda_1^d = \lambda_2^d = 1$. □

In centralization, Stage 2 is trivial: each delegate, if chosen by Nature, promotes her first-best optimum, i.e. $\hat{G}_1^d = \hat{x}_1^d$ and $\hat{G}_2^d = \hat{y}_2^d$. In Stage 1, this implies sincere delegation. No incentive for misrepresentation occurs because the delegate is either fully powerful or fully powerless. If selected as a ruler, the delegate gains full control over public good spending, so the best strategy is to have a ruler with identical preference. If not selected, the preference of the delegate is irrelevant for the outcome, so nominating a delegate with identical preference is not costly.

Proposition 2: In centralization with fully symmetric technology, no conflict of interests arises between median voters. By delegating sincerely, $\lambda_1^d = \lambda_2^d = 1$, median voters secure individual first-best optima $G^m = G_1^m = G_2^m$. There may be incentive for overprovision, though, since $\hat{y}_1^m \geq G^m$ and $\hat{x}_2^m \geq G^m$.

Proof Before Nature determines the ruler, the expected total public good level writes $E(G) = \frac{1}{2}\hat{G}_1^d + \frac{1}{2}\hat{G}_2^d = \frac{1}{2}\hat{x}_1^d + \frac{1}{2}\hat{y}_2^d$. Consider utility of median voter in Region 1, given this expected level:

$$U_1^m = b[E(G)] - cE(x) = \frac{1}{2}[b(\hat{x}_1^d) - c\hat{x}_1^d] + \frac{1}{2}[b(\hat{y}_2^d) - c\hat{x}_2^d]$$

In this expression, the delegate of Region 1 controls only the first term, so the median voter's optimization over λ_1^d is equivalent to maximizing only the first term. The first order condition yields:

$$b'(x_1^d)\frac{\partial x_1^d}{\partial \lambda_1^d} - c\frac{\partial x_1^d}{\partial \lambda_1^d} = 0 \Longrightarrow \frac{\partial x_1^d}{\partial \lambda_1^d}\left[b'(x_1^d) - c\right] = 0$$

From Lemma 6, we know $\partial x_1^d / \partial \lambda_1^d > 0$, so we need $b'(x_1) = c$. This is the first-best optimum of delegate $\lambda_1^d = 1$, as is obvious from (3.5). We thus confirmed sincere delegation: the median voter chooses a citizen candidate with identical preference, $\lambda_1^d = 1 = \lambda_1^m$, who promotes the median voter's first-best allocation if elected. □

Propositions 1 and 2 prove that the mode of governance may be irrelevant for the level of total public good, when symmetry is complete, which is a traditional conclusion in the public economics literature. In addition, neither centralization nor decentralization creates incentives for deviation from sincere delegation. The likely overprovision effect is but an effect of delegates' indifference about wasteful spending in the other region. This effect vanishes if we explicitly assume that citizens, if indifferent, select an allocation that maximizes utility of the other region (making no "harm without reason"); alternatively, we might assume that citizen weigh utility in the other region by an infinitesimally small amount. In consequence of that, both decentralization and centralization yield the first-best optima of median voters.

3.5 Asymmetric access

We keep focusing on technologies of symmetric weakest-link technology, for which components of total goods are equally important and diminish neither in quantity, nor in quality, i.e. $G = \min\{x, y\}$. The novelty is allowing for $\psi_1 \neq \psi_2$ and/or $c_x \neq c_y$.

In the first case, asymmetry is due to spatial properties of the total public good. Consider foundation of elite university, when one regions contributes by physical capital and the other region delivers experts. Because of its immobility of physical capital, the university has to be located in the region endowed by physical capital. As a result, spillovers from education may be asymmetric; labor market in region with the university has larger inflow of graduates and local firms may be more able to cooperate with the academy. A related example may be joint military forces. Prompt and efficient use of the forces are often limited by transportation costs. When the forces are located only in one region (due to training, administration and other reasons), and transportation costs are impossible to share by building the base somewhere close to borders, the region with the base cashes in additional benefits.

Definition 3 (Asymmetric access): For the fixed asymmetric access, let $G = \min\{x, y\}$, $\psi_1 = 1 > \psi_2 > 0$ and $c = c_x = c_y$. The randomly asymmetric access is defined by $G = \min\{x, y\}$, $c = c_x = c_y$ and two equiprobable states of the world: $\psi_1 = 1 > \psi_2 > 0$ (state X) and $\psi_2 = 1 > \psi_1 > 0$ (state Y). The state of the world is determined by Nature between Stages 1 and 2.

The definition states that for random asymmetry, the state of the world realizes after the elections in Stage 1. Why not sooner or later? Had it happened before Stage 1, we would just get the case of the fixed asymmetric access. Realization of uncertainty after x and y were determined (Stage 2) would only introduce new elements of insurance; asymmetry would disappear. In this analysis, we focus rather on conflicting interests involved in asymmetries.

Another reason has to do with possible extensions of our game. Suppose assume that after Stage 1, we have finitely repeated Stage 2, i.e. a finitely repeated game of delegates. Stage 1 can be interpreted as elections, and the repeated Stage 2 represents the annual decision-making on budget bills. To assume that the relevant states of the world realize inbetween elections, but not before elections, is a very plausible assumptions—politicians often get elected to solve problems whose details become apparent later, but information reveals soon enough to be taken into account prior decision. Now, as the extended (and more realistic) game is finitely repeated, its subgame-perfect Nash equilibrium strategies consist from actions that would be Nash equilibrium strategies in Stage 2 even if played just once. Therefore, we can rest with our game and have reason to believe that outcomes become relevant also in extended games, where the assumption of randomness is more justified.

3.5.1 Fixed asymmetry

We begin by fixedly lower access of Region 2, and distinguish between elastic and inelastic marginal benefit from public good. By Lemma 7, we use that the elasticity determines whether diminished access to total public good G ($\psi < 1$) decreases or increases demand for public good.

For decentralization, best responses are as standard. The subgame of delegates yields Nash equilibria in which $G = x = y$, where $G \in [0, \min\{G_1^d, G_2^d\}]$, so by iterated weak-dominance, we have $x = y = G = \min\{G_1^d, G_2^d\}$. As before, the amount of total public good is determined by preferences of the more conservative delegate.

Voters' optimization is different, though. Because of asymmetry, median voters have conflicting interests. We show that the weakest-link property will be always to the benefit of the median voter with more conservative preferences, so the other one will have incentive for strategic delegation.

Starting with high price elasticity, from Lemma 7 we infer that delegate from Region 2 is more conservative than delegate from Region 1 if voting is sincere, and thereby controls the weakest link:

$$\lambda_1^d = \lambda_2^d = 1 \Longrightarrow \hat{G}_1^d = \hat{x}_1^d > \hat{G}_2^d = \hat{y}_2^d \Longrightarrow \min\{\hat{G}_1^d, \hat{G}_2^d\} = \hat{G}_2^d$$

Region 2 median voter has no incentive for strategic voting, for his delegate is decisive for G (controls the weakest-link) and acts in accordance with preferences of the median voter. In contrast, median voter of Region 1 can improve the (second-)best outcome by strategic misrepresentation. A more conservative delegate promotes more conservative policy with lower costs, without affecting either the weakest-link or the level of total public good G. The incentive to misrepresent goes as far as the conservative delegate regains control over the weakest link.

In the case of inelasticity, Lemma 7 conjectures that delegate from Region 1 is more conservative, despite better access to public good. Now, it is Region 2's median voter who has an incentive to vote strategically for the more conservative delegate. Notice that in both cases, median voters delegate in a way that optimal decisions of their delegates become identical, $\hat{G}_1^d = \hat{G}_2^d$.

Proposition 3: In decentralization for the fixed asymmetry in access and elastic real demand for public good, we have conservative delegation in Region 1 and sincere delegation in Region 2, $\lambda_1^d < \lambda_2^d = 1$ and $G = \hat{G}_2^m < \hat{G}_1^m$. For inelastic demand, we have the inverse case, $1 = \lambda_1^d > \lambda_2^d$ and $G = \hat{G}_1^m < \hat{G}_2^m$.

Proof From Lemma 7, sincere voting in elastic case implies $\hat{G}_1^m = \hat{G}_1^d > \hat{G}_2^d = \hat{G}_2^m$, and symmetric weakest-link technology thus yields $G = \hat{G}_2^m$. Using monotonicity and Lemma 6, we get that utility of Region 1's median voter on some neighborhood $\lambda_1^d \in (1-\varepsilon, 1+\varepsilon)$ features $\partial U_1^m/\partial\lambda_1^d = b'(G)\partial G/\partial\lambda_d^1 - c = -c < 0$, since Region 1's delegate has no control over the weakest link ($dG/d\lambda_2^d = 0$). This holds up to $\bar{\varepsilon} > 0$, for which the delegate regains control over the weakest-link, namely for $\lambda_1^d \in (0, 1-\bar{\varepsilon}]$, we have $\partial U_1^m/\partial\lambda_1^d = b'(G)\partial G/\partial\lambda_d^1 - c > 0$. Positivity is evident from definition of optimum $\hat{G}_1^m$. The Region 1's median voter therefore delegates $\lambda_1^d = 1 - \bar{\varepsilon} < 1 = \lambda^m$. In elastic case, we again apply Lemma 7: $\hat{G}_1^m = \hat{G}_1^d = \hat{x}_1^d < \hat{G}_2^d = \hat{G}_2^m$. By switching subscripts, we analogically derive that Region 2's median voter selects $\lambda_2^d < 1 = \lambda^m$. $\square$

In centralization, no strategic aspect is involved; each delegate, if chosen by Nature, selects her first-best optimum ($\hat{G}_1^d = \hat{x}_1^d$ and $\hat{y}_1^d \geq \hat{x}_1^d$ for delegate 1, and $\hat{G}_2^d = \hat{y}_2^d$ and $\hat{x}_2^d \geq \hat{y}_2^d$ for delegate 2). This wipes out the strategic delegation effect completely.

Proposition 4: In centralization with the fixed asymmetry of access, we always have sincere delegation, $\lambda_1^d = \lambda_2^d = 1$, and the total amount of public good is $E(G) = \frac{1}{2}\hat{G}_1^m + \frac{1}{2}\hat{G}_2^m$. For price elastic demand for real public goods $\hat{G}_1^m > \hat{G}_2^m$, and for price inelastic demand $\hat{G}_1^m < \hat{G}_2^m$.

Proof In Stage 2, the ruling delegate j sets her first-best optimum, $\hat{G}_j^d$, thus $d\hat{G}_j^d/d\lambda_{-j}^d = 0$. Hence, $\lambda_1^d = \arg\max_{\lambda_1} E(U_1^m) = \arg\max[U_1^m(\hat{G}_1^d) + U_1^m(\hat{G}_1^d)]/2 = \arg\max_{\lambda_1} U_1^m(\hat{G}_1^d) = \lambda_1^m = 1$. Finally, from $\lambda_1^d = \lambda_2^d = 1 = \lambda^m$ and Lemma 7, we have $\hat{G}_1^m > \hat{G}_2^m$ for price-elastic demand for real public good, and the reverse $\hat{G}_1^m < \hat{G}_2^m$ for price-inelastic demand for real public good. □

3.5.2 Random asymmetry

Random asymmetry involves, by Definition 3, two equiprobable states of the world: $1 = \psi_1 > \psi_2 > 0$ (producers of x have better access, to be called X) and $\psi_2 = 1 > \psi_1 > 0$ (producers of y have better access, to be called Y). Nature determines the state of the world prior elections/delegation. This is possible in reality when elections are infrequent, so details are future projects are unknown, or when expertise to determine the place of construction is heretofore unknown. In the realm of defense, for instance, military experts would decide whether the common defense is to be provided predominantly by aircraft or ground-missiles. The former would better protect the border region (enemy airplanes have limited scope, so prefer easier targets and attack accessible regions), while the latter would be to the benefit of core region (enemy missiles need not to return, and thus target core industrial constructions regardless of location).

The total amount of public good writes, depending on the state of the world, as $G|_X = \min\{G_1^d|_X, G_2^d|_X\}$ or $G|_Y = \min\{G_1^d|_Y, G_2^d|_Y\}$. These are given by weakest links in each of the states of the world. Apparently, voters can expect four theoretical cases:

$$G|_X, G|_Y = \begin{cases} G_1^d|_X, G_1^d|_Y & \text{Case 11} \\ G_1^d|_X, G_2^d|_Y & \text{Case 12} \\ G_2^d|_X, G_1^d|_Y & \text{Case 21} \\ G_2^d|_X, G_2^d|_Y & \text{Case 22} \end{cases}$$

Lemma 8 explains conditions of existence of each case.

Lemma 8: For price-elastic demand for real public goods, there exist threshold functions $\bar{\lambda}_2^d|_X := \bar{\lambda}_2^d(\lambda_1^d) > \lambda_1^d$ and $\bar{\lambda}_2^d|_Y := \bar{\lambda}_2^d(\lambda_1^d) < \lambda_1^d$, such that:

$$\lambda_2^d \geq \bar{\lambda}_2^d|_X \Longrightarrow \text{Case 11}$$
$$\bar{\lambda}_2^d|_X \geq \lambda_2^d \geq \bar{\lambda}_2^d|_Y \Longrightarrow \text{Case 21}$$
$$\bar{\lambda}_2^d|_Y \geq \lambda_2^d \Longrightarrow \text{Case 22}$$

For price-inelastic demand for real public goods, there exist threshold functions $\bar{\lambda}_2^d|_X := \bar{\lambda}_2^d(\lambda_1^d) < \lambda_1^d$ and $\bar{\lambda}_2^d|_Y := \bar{\lambda}_2^d(\lambda_1^d) > \lambda_1^d$, such that:

$$\begin{aligned} \lambda_2^d \geq \bar{\lambda}_2^d|_Y &\Longrightarrow \text{Case 11} \\ \bar{\lambda}_2^d|_Y \geq \lambda_2^d \geq \bar{\lambda}_2^d|_X &\Longrightarrow \text{Case 12} \\ \bar{\lambda}_2^d|_X \geq \lambda_2^d &\Longrightarrow \text{Case 22} \end{aligned}$$

Proof Begin with price-elastic demand. For $\lambda_1^d = \lambda_2^d$, we know by Lemma 7 that in state X, Region 2 has lower demand, thus controls the weakest link, $G|_X = G_2^d|_X$, while for state Y, the reverse holds, $G|_Y = G_1^d|_Y$. This is Case 21. From Lemma 6, we get that $G_2^d|_X$ grows in λ_2^d, while $G_1^d|_X$ is constant, so for a certain (threshold) value $\bar{\lambda}_2^d(\lambda_1^d) > \lambda_1^d$, we attain equality $G_2^d|_X = G_1^d|_X$, and for higher values of λ_2^d, Region 1 overtakes control over the weakest link, therefore $G|_X = G_1^d|_X$. Moreover, the increase in λ_2^d doesn't affect the Region 1's control over the weakest link of in state Y, rather further enlarges the difference $G_1^d|_Y - G_2^d|_Y$. As a result, we get Case 11.
Now, start again from Case 21 with $\lambda_1^d = \lambda_2^d$. By decreasing λ_2^d, we know by Lemma 6 that we decrease both $G_2^d|_X$ and $G_2^d|_Y$. For state X, it means that we further enlarge difference $G_1^d|_X - G_2^d|_X$, so Region 2 only confirms control over the weakest link. For state Y, however, Region 2 is approaching the threshold value $\bar{\lambda}_2^d(\lambda_1^d) < \lambda_1^d$, for which equality $G_2^d|_Y = G_1^d|_Y$ occurs. Below the threshold value, Region 2 overtakes the weakest link, $G|_Y = G_2^d|_Y$, and since the other weakest link is unaffected, we are in Case 22.
For price-inelastic demand, we again start with symmetric delegation $\lambda_1^d = \lambda_2^d$, when Case 21 is in place. The reason is that by Lemma 7, if state X occurs for price-inelastic demand, Region 2 has higher nominal demand, thus doesn't control the weakest link, and $G|_X = G_1^d|_X$. For state Y, the reverse holds, so $G|_Y = G_2^d|_Y$. The threshold values are reached analogically, because property in Lemma 6 is not affected by the price elasticity. □

Anticipating outcomes in Lemma 8 when delegating, voters recognize that Nash equilibrium cannot be in Case 11, or Case 22. Intuitively, if one of the regions is decisive regardless of the state of Nature (i.e. controls weakest-links in both states, X and Y), the other region has an opportunity to increase her utility by decreasing production, which doesn't affect the total public good level, but lowers costs. This is achieved by delegating more conservatively. Case 11 thus stimulates Region 2 to delegate lower λ_2^d, and also Case 22 is unstable from the perspective of Region 1, as it would tend to delegate lower λ_1^d. Both best-responses would restore the case when each region got control in exactly one of the states (Case 21 for price-elastic demand, and Case 12 for price-inelastic demand). Therefore, we can restrict search for the equilibrium to Case 21 and Case 12.

Proposition 5: Decentralization with the random asymmetry of access implies sincere delegation of both regions, $\lambda_1^d = \lambda_2^d = 1$.

Proof In Case 21, $E(U_1^m) = \frac{1}{2}b(G_2^d|_X) - \frac{1}{2}G_2^d|_X + \frac{1}{2}b(\psi_2 G_1^d|_Y) - \frac{1}{2}G_1^d|_Y$. The first-order condition yields $\partial E(U_1^m)/\partial\lambda_1^d = \frac{1}{2}[\psi b'(\psi G_1^d|_Y - c)]dG_1^d|_Y/d\lambda_1^d$. For $\lambda_1^d = 1$, we have $\psi b'(\psi G_1^d|_Y) = c$. In consequence of that, $\partial E(U_1^m)/\partial\lambda_1^d = 0$. Analogically, for $\lambda_2^d = 1$, we would get $\partial E(U_2^m)/\partial\lambda_2^d = 0$. The same results can be derived for Case 12 by analogy. □

Intuition of the sincereness in decentralization is straightforward. Without loss of generality, consider Case 21 (corresponding to standard price elastic case). In state X, Region 1 has better access, but cannot control the weakest link. In state Y, Region 1 has worse access, but controls the weakest link. The Region 1's median voter is therefore perfectly satisfied with sincerely delegated citizen-candidate in state Y, because he delivers her optimal result, $G|_Y = \hat{G}_1^d$; the voter delegates sincerely, $\hat{G}_1^d = \hat{G}_1^m$. However, in state X, this sincerely delegated citizen candidate has a too big optimum to control the weakest link. Median voter however has no incentive to control the weakest link; in the first place, he wants to increase public good level, not decrease it, and secondly, non-sincere delegation would only distort delegate's decision in state Y, which would be costly.

Comparing to the fixed construction asymmetry, randomness induces voters to sincere voting. This has a constitutional conclusion: sincereness or true representation of preferences may be enhanced by allowing random asymmetry instead of fixed asymmetry. Other things being equal, it may be superior to agree on random asymmetry, for instance in the form of less frequent elections.

Finally, in Proposition 6, we explore centralization of randomly asymmetric access. In this case, we have to make assumptions about joint distribution of the two probabilities of the delegates, the probability of becoming a ruler, and the probability of getting better access. We examine formally only the case when the two probabilities are perfectly correlated, i.e. the region selected by Nature has both better access and is in power, so $G|_X = \hat{G}_1^d$ and $G|_Y = \hat{G}_2^d$.

Proposition 6: In centralization with the random asymmetry of access and the control of power perfectly correlated with better access, both regions delegate sincerely, $\lambda_1^d = \lambda_2^d = 1$.

Proof In centralization, $E(U_1^m) = \frac{1}{2}b(G_1^d|_X) - \frac{1}{2}G_1^d|_X + \frac{1}{2}b(\psi_2 G_2^d|_Y) - \frac{1}{2}G_2^d|_Y$. The first-order condition yields $\partial E(U_1^m)/\partial\lambda_1^d = \frac{1}{2}[b'(G_1^d|_X - c)]dG_1^d|_X d\lambda_1^d$. For $\lambda_1^d = 1$, we have $b'(G_1^d|_Y) = c$, implying $\partial E(U_1^m)/\partial\lambda_1^d = 0$. By analogy, for $\lambda_2^d = 1$, we get $b'(G_2^d|_Y) = c$, implying $\partial E(U_2^m)/\partial\lambda_2^d = 0$. □

In any case, we may immediately recognize that the shape of the joint distribution function is irrelevant. The reason is that for centralization, delegate of either of regions is either fully powerful (with half probability), or entirely powerless (with half probability). When the delegate controls power, she enforces her first-best optimum regardless of the rate of access. When not controlling power, the delegate cannot change anything. Therefore, median voter wants delegate to behave like him when in power (no matter what the rate of access is), which implies sincere delegation. More generally: non-cooperative centralization, as modeled here, cannot bring anything but sincere delegation.

3.6 Asymmetric cost

Asymmetry in costs, $c_x \neq c_y$, yields identical results we had for asymmetry in access and *elastic* demand for real public good. In this case, as there is no difference between nominal and real public good demand, and no difference between costs of the nominal and real public good, citizens always tend to decrease nominal public consumption when facing higher cost.

Definition 4 (Asymmetric cost): In the case of the fixed asymmetric cost, $G = \min\{x, y\}$, $\psi_1 = \psi_2 = 1$ and $c_x < c_y$. The randomly asymmetric cost is defined by $G = \min\{x, y\}$, $\psi_1 = \psi_2 = 1$ and two equiprobable states of the world: $c_x < c_y$ (state X) and $c_x > c_y$ (state Y). The state of the world is determined by Nature between Stages 1 and 2.

3.6.1 Fixed asymmetry

When asymmetry is fixed, we get results similar to elastic case of asymmetric access in Propositions 3 and 4.

Proposition 7: In decentralization for the fixed asymmetry in costs, we have conservative delegation in Region 1 and sincere delegation in Region 2, $\lambda_1^d < \lambda_2^d = 1$ and $G = \hat{G}_2^m < \hat{G}_1^m$.

Proof Identical to the proof of the first part of Proposition 3, which refers to case of price-elastic real public good demand. □

Proposition 8: In centralization with the fixed asymmetry of access, we have sincere delegation, $\lambda_1^d = \lambda_2^d = 1$, and the total amount of public good is $E(G) = \frac{1}{2}\hat{G}_1^m + \frac{1}{2}\hat{G}_2^m$, where $\hat{G}_1^m > \hat{G}_2^m$.

Proof From (3.5), (3.6), $b'(\cdot) > 0$, and $b''(\cdot) < 0$, we receive $\hat{G}_1^m > \hat{G}_2^m$. The rest is identical with the proof of Proposition 4. □

3.6.2 Random asymmetry

For random asymmetry, we again replicate results of elastic case of asymmetric access, given in Propositions 5 and 6.

Proposition 9: Decentralization with the random asymmetry of costs implies sincere delegation of both regions, $\lambda_1^d = \lambda_2^d = 1$.

Proof We use that real demand for public good is declining in real price, so for symmetrical sincere delegation we have by Lemma 8 Case 21. Consider benefits of deviations from this strategy profile. In state X of Case 21, Region 1 has lower cost, but cannot control the weakest link. In state Y, Region 1 has higher cost, and controls the weakest link. The Region 1's median voter is therefore perfectly satisfied with sincerely delegated citizen-candidate in state Y, because the delegate delivers her optimal result, $G|_Y = \hat{G}_1^d|_Y = \hat{G}_1^m|_Y$. In state X, the sincerely delegated citizen-candidate doesn't control the weakest link, but the median voter has no incentive to vote conservatively, because lower λ_1^d would decrease $G|_Y$ below optimal $\hat{G}_1^m|_Y$ and, if conservative delegation were excessive and Case 21 switched to Case 11, also $G|_X$ would further fall. The best response is $\lambda_1^d(1) = 1$, which holds also symmetrically, $\lambda_2^d(1) = 1$, and therefore Nash equilibrium writes $\lambda_1^d = \lambda_2^d = 1$. □

Proposition 10: In centralization with the random asymmetry of costs and the control of power perfectly correlated with lower cost, both regions delegate sincerely, $\lambda_1^d = \lambda_2^d = 1$.

Proof We have $E(U_1^m) = \frac{1}{2}b(G_1^d|_X) - \frac{1}{2}c_{x|X}G_1^d|_X + \frac{1}{2}b(G_2^d|_Y) - \frac{1}{2}c_{x|Y}G_2^d|_Y$. The F.O.C. yields $\partial E(U_1^m)/\partial\lambda_1^d = \frac{1}{2}[b'(G_1^d|_X) - c_{x|X}]dG_1^d|_X d\lambda_1^d$. For $\lambda_1^d = 1$, we have $b'(G_1^d|_X) = c_{x|X}$, implying $\partial E(U_1^m)/\partial\lambda_1^d = 0$. By analogy, for $\lambda_2^d = 1$, we get $b'(G_2^d|_Y) = c_{y|Y}$ and consequently $\partial E(U_2^m)/\partial\lambda_2^d = 0$. □

3.7 Penalty for the strictly weakest link

Consider a case when access to public good is not exogenous, but endogenous. Specifically, the region with the strictly smallest contribution (controlling thus the weakest link) may have lower access $\psi < 1$ than other regions. This is a situation of protection against floods, when region with the lowest dam is the one which suffers "the first strike". Even if other regions are also affected, this region is disadvantaged by lower amount of time to prepare and higher exposure to water.

Definition 5 (Weakness penalty): When penalty for the strictly weakest links is in place, we have $G = \min\{x, y\}$, $c := c_x = c_y$, $0 < \psi < 1$ and

$$(\psi_1, \psi_2) = \begin{cases} (\psi, 1) : x < y \\ (1, 1) : x = y \\ (1, \psi) : x > y. \end{cases}$$

Whether to match contributions of the other region or rather control the weakest link is now influenced by two extra considerations. First, controlling the weakest link is costlier. Second, when the weakest link is under control, optimum amount of public good may decrease (for price-elastic demand for real public good), but also increase (for price-inelastic demand).

First of all, we find best responses of delegates in Stage 2. Utility of delegate i writes $U_i^d(\psi_i, G)$, and since $\psi_i \in \{\psi, 1\}$, we have two utility functions $U_i^d(\psi, G)$ and $U_i^d(1, G)$, with maxima in $\hat{G}_i^d(\psi)$ and $\hat{G}_i^d(1)$. By Lemma 7, we have $\hat{G}_i^d(\psi) < \hat{G}_i^d(1)$ for price-elastic demand for real public good, and $\hat{G}_i^d(\psi) > \hat{G}_i^d(1)$ for price-inelastic demand for public good. Further, as $b'(\cdot) > 0$, we have that $U_i^d(\psi, G) < U_i^d(1, G)$ holds for all G.

Solving for best responses, recognize firstly that the delegate in Region 1 has basically three options: match the other contribution ($x = y$), control the weakest link ($x < y$), or overspend ($x > y$). Overspending is clearly dominated by matching the contribution, since for $x \geq y$, public good is constant ($G = y$), and costs in terms of forgone private good fall. Thus, we have only two effective options.

To control the weakest link (the 1st option) implies to optimize $U_1^d(\psi, G)$ on $x \in \langle 0, y)$. As $U_1^d(\psi, G)$ is quasiconcave in G, this yields that for $y < \hat{G}_1^d(\psi)$, utility grows in G, so $x = y$ in optimum. For $y \geq \hat{G}_1^d(\psi)$, one obviously promotes the optimum by setting $x = \hat{G}_1^d(\psi)$.

The 2nd option is to match consistently, $x = y$. In order to derive the best response, we compare utilities of both options. For $y < \hat{G}_1^d(\psi)$, both options yields identical responses, so we have $\tilde{x}(y) = y$. For $y > \hat{G}_1^d(\psi)$, we use that 1st option gives constant utility, $U_1^d(\psi, \hat{G}_1^d(\psi)) < U_1^d(1, \hat{G}_i^d(\psi))$. Therefore, the optimum of $U_1^d(\psi, G)$, namely $\hat{G}_1^d(\psi)$, is not always desirable. If y is too close to it, it pays off to match, i.e. overspend, and suffer worse access (penalty).

The 2nd option, by assumption of $b''(\cdot) < 0$, gives decreasing utility $U_1^d(1, y)$. Therefore, this willingness to overspend must end for sufficiently high y. Unless $b'''(\cdot)$ is too large (extremely close to zero), we get that there exists intersection of these functions; more formally, there exists a critical threshold $\bar{G}_1$, for which $U_1^d(\psi, \hat{G}_1^d(\psi)) = U_1^d(1, \bar{G}_1)$. Notice immediately that $\bar{G}_1 > \hat{G}_1^d(1)$. We summarize

$$\tilde{x}(y) = \begin{cases} y : y \leq \bar{G}_1 \\ \hat{G}_1^d(\psi) : y > \bar{G}_1. \end{cases} \tag{3.16}$$

Obviously, the best response of delegate in Region 2 is symmetric. If we define $\bar{G}_2$ by $U_2^d(\psi, \hat{G}_2^d(\psi)) = U_2^d(1, \bar{G}_2)$, we have

$$\tilde{y}(x) = \begin{cases} x : x \leq \bar{G}_2 \\ \hat{G}_2^d(\psi) : x > \bar{G}_2. \end{cases} \tag{3.17}$$

Where is Nash equilibrium of the game? As usual for games where matching is a frequent best response, we have infinite amount of Nash equilibria, namely the interval $G \in \langle 0, \min\{\bar{G}_1, \bar{G}_2\}\rangle$, where $x = y = G$. As usual, we can use the pre-play communication as a refinement ruling out Pareto-inefficient equilibria. We can also use other refinements as below.

3.7.1 Pareto-efficient equilibria

If pre-play communication enables delegates to coordinate on Pareto-efficient equilibria only, the set restricts to $G \in \langle \min\{\hat{G}_1^d(1), \hat{G}_2^d(1)\}, \max\{\hat{G}_1^d(1), \hat{G}_2^d(1)\}\rangle$.

We will see that for Nash equilibrium in the whole game, it is irrelevant which of the Pareto-efficient equilibria is chosen.

Proposition 11: In decentralization with the weakest-link penalty, where delegates coordinate on some of Pareto-efficient equilibria, sincere delegation is a subgame-perfect Nash equilibrium, $\lambda_1^d = \lambda_2^d = 1$.

Proof Best responses of the delegates are derived in (3.16) and (3.17). Since all Nash equilibria consist of symmetric strategy profiles, no weakest-link penalty takes place in equilibrium. It is sufficient to examine only stability of any symmetric strategy profile, $\lambda_1^d = \lambda_2^d$. By symmetry, $\hat{G}_1^d(1) = \hat{G}_2^d(1)$. Due to non-effectiveness of weakness penalty and symmetry, the set of equilibria reduces to singleton $G \in \{\hat{G}_1^d(1)\}$. Obviously, this yields optimum only for sincere voting, so we have that $(\lambda_1^d, \lambda_2^d) = (1, 1)$ is (at least weak) Nash equilibrium.

3.7.2 Overshooting equilibrium

Pre-play communication may not lead to Pareto-efficient equilibrium, if extra players are in charge of the pre-play announcements, e.g. bureaucracy. For many reasons, bureaucrats may have the prerogative to convey prior messages to the other regions. If the bureaucracy is interested in maximizing output, they promote the Nash equilibrium which features maximal output (overshooting equilibrium), $G = \min\{\bar{G}_1, \bar{G}_2\}$, which affects the game.

Proposition 12: In decentralization with the weakest-link penalty, where delegates coordinate on the overshooting equilibrium, conservative delegation is a subgame-perfect Nash equilibrium, $\lambda_1^d = \lambda_2^d < 1$.

Proof Best responses of the delegates are derived in (3.16) and (3.17). The overshooting equilibrium is Nash equilibrium with the maximal output, i.e. $G = \min\{\bar{G}_1, \bar{G}_2\}$. The equilibrium is symmetric, hence without weakness penalty. Consider best responses of median voter in Region 1. She knows that $\bar{G}_1 > \hat{G}_1^d(1)$. To get $G = \hat{G}_1^m$, she has to delegate in a way that brings him first-best optimum, $\bar{G}_1 = \hat{G}_1^m$, so $\hat{G}_1^d(1) < \hat{G}_1^m$, which by Lemma 6 implies conservative delegation $\lambda_1^d < 1$. Of course, we must at the same time have $\bar{G}_2 \geq \bar{G}_1 = \hat{G}_1^m$. Using $\hat{G}_2^d(1) < \bar{G}_2$ and monotonicity in Lemma 6, we realize that it holds for $\lambda_2^d \geq \lambda_1^d$. This reveals that λ_1^d expressed above is a feasible best response to $\lambda_2^d = \lambda_1^d$. As best replies of median voter in Region 2 are symmetric, we get that by delegating $\lambda_1^d = \lambda_2^d < 1$, they get first-best optima $\hat{G}_1^m$ and $\hat{G}_2^m$.

3.7.3 Non-cooperative centralization

Centralization again makes one region powerful and the other powerless, with exogenous probability. The ruling delegates can enforce their first best allocations, which leads median voters to delegate sincerely.

Proposition 13: In centralization with the penalty for the strictly weakest link, both regions delegate sincerely, $\lambda_1^d = \lambda_2^d = 1$.

Proof Since $U_i^d(\psi, G) < U_i^d(1, G)$, the ruling delegate doesn't impose the penalty on her region, so $x = y = G$. With full control over the allocation, she sets $G = \hat{G}_i^d$. The first-order condition yields $\partial E(U_1^m)/\partial \lambda_1^d = \frac{1}{2}[b'(G_1^d|_X - c)]dG_1^d|_X d\lambda_1^d$. For $\lambda_1^d = 1$, we have $b'(G_1^d|_Y) = c$, implying $\partial E(U_1^m)/\partial \lambda_1^d = 0$. By analogy, for $\lambda_2^d = 1$, we get $b'(G_2^d|_Y) = c$, implying $\partial E(U_2^m)/\partial \lambda_2^d = 0$. □

3.8 Conclusion

Outcomes for all asymmetries are listed in Table 3.2. All types of asymmetries in decentralization result in strategically conservative delegation, $\lambda < 1$. For the fixed asymmetric access and normal (elastic) real public good demand, the region with better access votes conservatively to match contribution of the region with worse access. If the demand is inelastic, it is the region with worse access which delegates conservatively. For the fixed asymmetric cost, we again receive that the region with lower costs votes conservatively in order to match lower contributions of the less fortunate region.

The interesting result occurs for the weakness penalty, especially if delegates for some reason (e.g., bureaucratic control of pre-play communication) coordinate on an overshooting equilibrium. Then, both regions vote conservatively to overcome this bias. Strategic delegation is welfare-enhancing, unlike the previous cases when it favored one region at the expense of the other region.

Non-cooperative centralization is shown to reduce strategic delegation in all asymmetries: asymmetric access, asymmetric costs, and the existence of penalty for the strictly weakest contribution. In the simple form, it always induces sincere delegation, since delegates are either powerful or powerless.

Tab. 3.2: Delegation and the public good provision in decentralization: the summary

Case	Def.	Delegation	Provision
Full symmetry	2	$\lambda_1 = \lambda_2 = 1$	$G = \hat{G}_1^m = \hat{G}_2^m$
Fixed access, elastic	3	$\lambda_1 < \lambda_2 = 1$	$G = \hat{G}_2^m < \hat{G}_1^m$
Fixed access, inelastic	3	$\lambda_2 < \lambda_1 = 1$	$G = \hat{G}_1^m < \hat{G}_2^m$
Random access, elastic	3	$\lambda_1 = \lambda_2 = 1$	$G = \frac{1}{2}\hat{G}_2^m\|_X + \frac{1}{2}\hat{G}_1^m\|_Y$
Random access, inelastic	3	$\lambda_1 = \lambda_2 = 1$	$G = \frac{1}{2}\hat{G}_1^m\|_X + \frac{1}{2}\hat{G}_2^m\|_Y$
Fixed cost	4	$\lambda_1 < \lambda_2 = 1$	$G = \hat{G}_2^m < \hat{G}_1^m$
Random cost	4	$\lambda_1 = \lambda_2 = 1$	$G = \frac{1}{2}\hat{G}_2^m\|_X + \frac{1}{2}\hat{G}_1^m\|_Y$
Penalty, Pareto	5	$\lambda_1 = \lambda_2 = 1$	$G = \hat{G}_1^m = \hat{G}_2^m$
Penalty, overshooting	5	$\lambda_1 = \lambda_2 < 1$	$G = \hat{G}_1^m = \hat{G}_2^m$

This scope of the chapter is limited in three respects. First, attention is paid only to symmetric weakest-link technologies, hence no complementarity between public bads that move costly is examined. Second, centralization is purely non-cooperative, and simplistic. Third, there are no other financial flows connecting the regions. These are the natural avenues for further research.

References

[1] Arrow, K. J. & A. C. Enthoven (1961). "Quasiconcave Programming," *Econometrica*, 29 (October), 779–800.

[2] Bardhan, P., M. Ghatak & A. Karaivanov (2002). "Inequality, Market Imperfections, and the Voluntary Provision of Collective Goods", in P. Bardhan, S. Bowles & J. M. Baland (ed.): *Economic Inequality, Collective Action, and Environmental Sustainability*. Princeton University Press, Princeton.

[3] Barro, R. J. & X. Sala-i-Martin (1999). *Economic Growth*, Cambridge MA, MIT Press.

[4] Besley, T. & S. Coate (2003). "Centralized versus decentralized provision of local public goods: a political economy approach", *Journal of Public Economics*, 87, 2611–2637.

[5] Cornes, R. (1993). "Dyke Maintenance and Other Stories: Some Neglected Types of Public Goods," *The Quarterly Journal of Economics*, 108 (1), 259–271.

[6] Dur, R. & Roelfsema, H. (2005). "Why does centralisation fail to internalise policy externalities?", *Public Choice*, 122. 395–416.

[7] Grossman, V. (2003). "Income inequality, voting over the size of public consumption, and growth", *European Journal of Political Economy*, 19, 265-287.

[8] Heal, G. & H. Kunreuther (2005). "IDS Models of Airline Security," *Journal of Conflict Resolution*, 49 (2), 201–217.

[9] Hirshleifer, J. (1983). "From weakest link to best shot: The voluntary provision of public goods," *Public Choice*, 41 (3), 371–386.

[10] Leontief, W. (1941). *The Structure of the American Economy: 1919–1929*, Cambridge MA, Harvard University Press.

[11] Sandler, T. (2006). "Regional public goods and international organizations," *Review of International Organizations*, 1, 5–25.

[12] Sandler, T. & S. Vicary (2001). "Weakest-link public goods: giving in-kind or transferring money in a sequential game", *Economics Letters*, 74, 71-75.

[13] Vicary, S. & T. Sandler (2002). "Weakest-link public goods: giving in-kind or transferring money," *European Economic Review*, 46, 8, 1501–1520.

[14] Weingast, R. B., Shepsle K. A. & C. Johnsen (1981). "The political economy of benefits and costs: A neoclassical approach to distributive politics," *Journal of Political Economy*, 89(4), 642-664.

4. TOLERABLE INTOLERANCE

4.1 Introduction

Man needs to confess in order to make life meaningful. The subjects of confession can be various religious, ideological, or personal cults, especially in contemporary western societies with unregulated supply of confessions. The absence of regulations in the market with confessions is nevertheless subject to discussions and often even a reason for red alert, given the religious hatred spread especially by Muslim clerics in Europe, or proliferation of books promoting the Auschwitz Lie. Contrarily, in the United States, Conservatives intensifies the "cultural war", arguing about (in)correctness of liberal norms separating religion and society. In Europe, we observe retreat of tolerant and defensive monotheistic religions that have been little interested in evangelization in last decades; in their place come often virulent sects, idiosyncratic philosophies and personal cults.

The Continental-European approach to intolerant expression is to regulate entry and prevent conflicts by banning cults historically involved in civilization disasters. Swastikas cannot be worn even by members of British royal family, Muslim girls in French schools cannot cover their faces by traditional habits, and Holocaust refuters are prosecuted. The Anglo-Saxon approach is distinctly more liberal; it believes that decent citizens manage to isolate individuals with faulty and immoral opinions, and prosecutes only ideas directly involved in criminal and terrorist acts, such as London preachments by Abu Hamza Masri.

This chapter aims to contribute to this discussion in a rigorous manner and clarify the working properties of the liberal (Anglo-Saxon) approach. We model a non-regulated society where people with confessions of different content and different level of tolerance meet in random, bilateral interactions. Confession, as in Durkheim ([1912] 1995), is defined as the set of taboos.[1] Individuals adjust confessions on the basis of payoffs in interaction; confessions thus unequally spread in social interactions. Tolerance and

[1] Durkheim ([1912] 1995) maintained that "a religion is a unified set of beliefs and practices relative to sacred things, that is to say, things set apart and forbidden".

intolerance is the property of confession which regulates much of social behavior. In specific, tolerance is understood as indifference; intolerance is propensity to reward resemblance and punish difference, even at a material cost.[2] We seek for evolutionarily stable states defined such that the profile of confessions held in population cannot be invaded by any competing profile. We identify when outcomes of evolution are socially efficient. This involves the issue whether confessions detrimental to human welfare are sustainable in a society with free entry.

After discussing methodology in Section 4.2, we create setup of the model. In Section 4.3, we solve it in pure strategies and extend to mixed strategies interpreted as a polymorphic state. Both in pure and mixed strategies, we obtain equivalent results with regard to stability of intolerant confessions. We derive policy implications and discuss extensions in Section 4.4, while Section 4.5 concludes.

4.2 Methodology

4.2.1 Rational choice approach

The economics of religion is typically ascribed to the rational choice approach, rooted in Chicago economics (Iannacconne 1998). The founding axioms of the rational choice approach constitute the famous trinity (Iannacconne 1996):

A1. Individuals weigh the costs and benefits of potential actions, and choose those actions that maximize their net benefits.

A2. The ultimate preferences that individuals use to assess costs and benefits tend not to vary much from person to person or time to time.

A3. Social outcomes constitute the equilibria that emerge from the aggregation and interaction of individual actions.

At least A1 and A2 resist consistent application in the market with confessions. Among confessions, perfectly rational choice is a much more heroic assumption than in areas where calculation is the daily bread. The traditional short-cut is to replace perfect rationality by evolution: “Evolutionary forces will favor maximizing behaviors even if religious firms do not consciously strive for ‘success’...” (Iannnacconne 1996, p. 4). However,

[2] Among others, this allows to enter confession into endogenous growth models, as in Blum and Dudley (2001).

there is no reason to presume that evolutions bring outcomes identical to hypothetical perfect rationality equilibria.

Inclusion of internal norms into the rational choice framework is more than difficult. In one extreme, the norms can be interpreted as component of ultimate preferences in A2. Considering that people in the world show significant variation of internal norms is however inconsistent with A2, where ultimate preferences are defined to have little variation. A quasi-solution is to subsume inner norms into human capital. Norms are specified such that the human capital follows the optimal time-consistent path. Any change of constraint (e.g. leveraging taboo on premarital sex after leaving high school, starting to smoke in the age of 70) is thus a pre-planned response to a pre-planned change in incentives. But, in order to get to the plan, one requires either some explicitly evolutionary interpretation (not justified nor modeled in A1-A3), or an assumption that at certain point of time, individual is able to consider one's own ultimate preferences; any later decision would involve accumulated capital, i.e. sunk costs, which need not to be ex ante optimal. To overstretch, the only individual that could make a perfect plan would be either a kid or a culture-free barbarian.

Applying the human-capital approach to confessions, we have it difficult to interpret why people commit and admit errors. Conversions would also have to be radically reinterpreted—the model of conversion would probably be St. Augustine who converted in his 30s just after he had enjoyed all beauties of vicious life.

We rather suggest that norms are nothing but continuously evolving instruments for good life, subject to current situations and/or past experience, not to a given plan. Norms are indeed internalized, and abandoned only under evolutionary pressures. To say that cultural norms are accepted as binding, because they express the (whatever temporary) opinion on the good life, sounds trivial; simplicity however avoids philosophically delicate issues of time consistency in the evolution of culture, and arguments that babies or barbarians invest optimally into human capital.

4.2.2 Evolutionary approach

Full shelves in Departments of Theology illustrate that assessment of available information on confessions is extremely difficult for an individual. Promotion of confessions and ideologies arguably doesn't resemble information-processing. Therefore, we do not hasten to put excessive requirements on the ability of individuals, and follow the bounded rationality complemented by evolutionary processes, as spelt out in the following preliminaries:

B1. Individuals act boundedly rationally, adjusting confessions on the basis of experience in social interactions.

B2. The ultimate preferences on confessions (metaphysical truth) tend not to vary much from person to person or time to time, and are revealed in social interactions.

B3. Social outcomes constitute the distribution of confessions that cannot be invaded by little perturbations in the distribution of confessions.

In B1, we relax perfect rationality in the realm of choice of confession.[3] In B2 (very much like in A2), we assume that universal, the most efficient confession exists. Unlike human-capital approach, we understand constraints as elements of preferences, so good norms are goals per se, not instruments to any other goal (e.g. non-smoking can be instrumental to health). Thereby we omit paradoxical implications discussed above.

By B2, individuals decide only on the basis of social interactions, including observations of other people. We justify this simplification in three ways. First, complex confessions are apparently unverifiable by reasoning of a common man, whereas imitation of (real or fictitious) individuals who have a higher payoff (demonstrate satisfaction in life personally or in media) is available and very simple indeed. Second, we can comfortably study social functions of confessions. Third, the relative importance of social effects and adjustment towards the metaphysical truth can be properly addressed by a well-defined payoff function.

Interactions can be interpreted as day-to-day cultural encounters of adult people in a fixed population, which involve two basic subsets: interactions related to social activities and mass media communication. The former involves numerous reciprocal acts, which explains benefits and costs especially when individuals occur in social dilemmas. Social role of elements of confessions is merely binary (tolerance and intolerance), therefore deviations from socially beneficial (intolerant) confession is only for the reason of "protecting one's own consciousness". This allow to clearly study the conflict between individual consciousness and group cooperation (or group survival, cf. Wilson 2002). Regarding media, I presume that competitive mass media used by entire population conform to the assumption of random communication; as the media have interest to virtually mirror the society proportionally, they approximate the random interaction of cultural norms. This concept is of course subject to several qualifications, discussed in Section 4.4.

[3] Notice that to date, no-one has proved what the meaning of life actually is. Novel hypotheses suggest even purely abstract purposes, such as Answer 42 in Adams (1979).

B3 points at states that are evolutionarily stable, where — if a small amount of mutants (differently confessing individuals) appears in the population — the mutants have a lower utility in this post-entry population, and therefore are driven out of the game.

4.3 The model

4.3.1 Assumptions

We will model social interactions as a repeated play of a symmetric bilateral stage-game, where players are randomly drawn from a large population. Denote the drawn players as A, B. Player A enters the game with a *confession* in the form of $m \times 2$ matrix $\mathbf{C}_A = (\mathbf{a}, \mathbf{u})$ which describes constraints and intolerance in m elements of life. The confession prescribes each element $i \in \{1, 2, \ldots, m\}$ to be either regulated, i.e. *taboo* ($a_i = 1$), or not regulated, i.e. *freedom* or no-taboo ($a_i = 0$). Also, it describes whether to be *tolerant* in the element ($u_i = 0$) or *intolerant* ($u_i = 1$). Assume that intolerance exists only for taboos ($u_i \leq a_i$). And, by analogy, let Player B enter the game with a confession $\mathbf{C}_B = (\mathbf{b}, \mathbf{v})$ with identical properties ($b_i \in \{0, 1\}, v_i \in \{0, 1\}, v_i \leq b_i$).

Definition 6 (Confessions): Let $M = \{1, 2, \ldots, m\}$ be the set of elements. Define vectors of taboos $\mathbf{a} = (a_i)_{i \in M}$, $\mathbf{b} = (b_i)_{i \in M}$ and vectors of intolerance $\mathbf{u} = (u_i)_{i \in M}$, $\mathbf{v} = (v_i)_{i \in M}$. By $\mathbf{C}_A = (\mathbf{a}, \mathbf{u})$ and $\mathbf{C}_B = (\mathbf{b}, \mathbf{v})$, denote confessions as the pure strategies of Players A and B. Define the identical pure strategy set for Players A and B as $S_p = \{(\mathbf{x}, \mathbf{y}) \in \{0, 1\}^m \times \{0, 1\}^m : y_i \leq x_i\}$. The set of pure strategy profiles writes as $\Theta_p \equiv S_p \times S_p$ and $\mathbf{C} = (\mathbf{C}_A, \mathbf{C}_B) \in \Theta$ is the $m \times 4$ matrix denoting the pure-strategy profile of Players A and B.

Notice that the restrictions on vectors of intolerance $\mathbf{u} \leq \mathbf{a}$ and $\mathbf{v} \leq \mathbf{b}$ diminish cardinality of S_p in comparison with a hypothetical situation without restrictions:

$$\text{card}(S_p) = 2^m \sum_{i \in M} \frac{2^i m!}{i!(m-i)!} < 2^{2m}$$

To interpret and identify the elements of life over which taboos are imposed is not complicated–obvious examples are Sunday work, blood transfusion, premarital sex, annual pilgrimage and the like. Some confessions are not very restrictive, and impose restrictions on Sunday mornings, marriages and funerals at most, whilst others take nearly all wealth, free time and children; the most extreme sects demand control over all details of daily life.

In our understanding, tolerance and intolerance are nothing but proxy for social function of confessions. An intolerant person enjoys two functions of the confession–access to metaphysical *truth*, and access to like-minded *community*. Unlike him, a tolerant person can enjoy only the latter. Put in other words, tolerance represents the case when people are oriented toward private salvation, and intolerance represents orientation toward public salvation. The relative pros and cons of tolerance and intolerance are discussed and formally introduced in Assumption 1.

We have further assumed that intolerance takes place only in taboos; otherwise, the individual is tolerant. Consider an example of two taboos, stealing and no baptism. A Christian who obeys both taboos may be intolerant in one taboo and tolerant in another one. In this example, he may be relentless on stealing (which may affect pocket, anyway), but may feel no discomfort if others don't baptize their children. By postulating that intolerance doesn't appear in the absence of taboo, we imply in this example that thieves (individuals without taboo on stealing) cannot be socially harmed by meeting decent citizens, and atheists (individuals without no taboo regarding baptism) cannot be shocked when observing baptism.

Definition 7 (Pure-strategy payoffs): For any $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \in \Theta_p$, let the combined expected payoffs for Player A and Player B be a function $\pi : \Theta_p \to \mathbb{R}^2$, where $\pi(\mathbf{C}) = (\pi_A(\mathbf{C}), \pi_B(\mathbf{C}))$. Given some vector $\mathbf{x} = (x_i)_{i \in M}$, denote by (y, x_{-i}) a vector whose i-th coordinate is y and all j-th coordinates, $j \in M, j \neq i$, are identical with vector $\mathbf{x}$. For any $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \in \Theta_p$, let $\pi_A^i(a_i, u_i, b_i, v_i, \mathbf{C}) \equiv \pi_A((a_i, a_{-i}), (u_i, u_{-i}), (b_i, b_{-i}), (v_i, v_{-i}))$. Define the vectors of differences $\mathbf{\Delta}$, $\nu^{\mathbf{0}}$, $\nu^{\mathbf{1}}$, $\mathbf{\Phi^0}$, $\mathbf{\Phi^1}$, $\rho^{\mathbf{0}}$, and $\rho^{\mathbf{1}}$ with coordinates $i \in M$ in the following way:

$$\begin{aligned}
\Delta_i(\mathbf{C}) &\equiv \pi_A(0,0,0,0,\mathbf{C}) - \pi_A(1,0,1,0,\mathbf{C}) \\
\nu_i^0(\mathbf{C}) &\equiv \pi_A(0,0,0,0,\mathbf{C}) - \pi_A(0,0,1,0,\mathbf{C}) \\
\nu_i^1(\mathbf{C}) &\equiv \pi_A(1,0,1,0,\mathbf{C}) - \pi_A(1,0,0,0,\mathbf{C}) \\
\Phi_i^0(\mathbf{C}) &\equiv \pi_A(0,0,1,0,\mathbf{C}) - \pi_A(0,0,1,1,\mathbf{C}) \\
\Phi_i^1(\mathbf{C}) &\equiv \pi_A(1,0,0,0,\mathbf{C}) - \pi_A(1,1,0,0,\mathbf{C}) \\
\rho_i^0(\mathbf{C}) &\equiv \pi_A(1,1,1,0,\mathbf{C}) - \pi_A(1,0,1,0,\mathbf{C}) \\
\rho_i^1(\mathbf{C}) &\equiv \pi_A(1,1,1,1,\mathbf{C}) - \pi_A(1,1,1,0,\mathbf{C}) \\
\rho_i^2(\mathbf{C}) &\equiv \pi_A(1,1,1,1,\mathbf{C}) - \pi_A(1,0,1,1,\mathbf{C})
\end{aligned}$$

Besides $\boldsymbol{\Delta}$, which indicates truthfulness (efficiency) of a taboo, all other variables relate to social function of confessions. Variables $\nu^{\mathbf{0}}$ and $\nu^{\mathbf{1}}$ capture interactions of tolerant players; the former denotes how much a non-believer loses when meeting a believer instead of non-believer, and the latter denotes the respective loss of a believer.

Losses associated with social behavior of intolerant players (facing heterogeneity) are described by $\boldsymbol{\Phi}^{\mathbf{0}}$ and $\boldsymbol{\Phi}^{\mathbf{1}}$. Variable $\boldsymbol{\Phi}^{\mathbf{0}}$ is for ostracism, refusal to cooperate and other types of social punishment which a non-believer suffers when meeting an intolerant believer; $\boldsymbol{\Phi}^{\mathbf{1}}$ is the reverse side of the identical phenomenon, namely the loss (disgust, refusal to reciprocate) of an intolerant believer who meets a non-believer.

The three remaining variables represent extra gains from social behavior of intolerant players, and apply only for interactions between believers. Variable $\rho^{\mathbf{0}}$ is for the temptation to become intolerant, when interacting with a tolerant believer. Variable $\rho^{\mathbf{2}}$ is a similar gain, in the case of interaction with an intolerant believer. Finally, $\rho^{\mathbf{1}}$ is the gain of an intolerant believer who witnesses switch from tolerance to intolerance. Table 4.1 summarizes payoffs of Player A for differences in i-th element, where, without loss of generality, $\pi_A(1,0,1,0,\mathbf{C}) \equiv 0$.

Tab. 4.1: Payoffs of Player A, π_A, when $\pi_A(1,0,1,0,\mathbf{C}) \equiv 0$

$a_i, u_i / b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i	$\Delta_i - \nu_i^0$	$\Delta_i - \nu_i^0 - \Phi_i^0$
$1,0$	$-\nu_i^1$	0	$\rho_i^0 + \rho_i^1 - \rho_i^2$
$1,1$	$-\nu_i^1 - \Phi_i^1$	ρ_i^0	$\rho_i^0 + \rho_i^1$

In the following assumptions, we impose restrictions on these differences; the restrictions express the idea that tolerance is associated with indifference and that intolerance involves numerous social functions. This makes intolerance strong in homogenous society, but fragile in heterogenous society.

Assumption 1: The payoff function π satisfies:

1. Symmetry: $\forall \mathbf{C}_A, \mathbf{C}_B \in S_p : \pi_B(\mathbf{C}_A, \mathbf{C}_B) = \pi_A(\mathbf{C}_B, \mathbf{C}_A)$
2. Additivity: $\forall i \in M, \forall \mathbf{C}^1, \mathbf{C}^2 \in \Theta_p, (a^1_{-i}, u^1_{-i}, b^1_{-i}, v^1_{-i}) = (a^2_{-i}, u^2_{-i}, b^2_{-i}, v^2_{-i})$:

$$\Delta_i(\mathbf{C}^1) = \Delta_i(\mathbf{C}^2), \Phi_i^0(\mathbf{C}^1) = \Phi_i^0(\mathbf{C}^2), \Phi_i^1(\mathbf{C}^1) = \Phi_i^1(\mathbf{C}^2),$$
$$\rho_i^0(\mathbf{C}^1) = \rho_i^0(\mathbf{C}^2), \rho_i^1(\mathbf{C}^1) = \rho_i^1(\mathbf{C}^2), \rho_i^2(\mathbf{C}^1) = \rho_i^2(\mathbf{C}^2)$$

3. Zero social impact of tolerant players: $\nu^0(\mathbf{C}) = \nu^1(\mathbf{C}) = \mathbf{0}$

4. In heterogeneous interaction, intolerance has a negative social impact and the intolerant player pays a cost of intolerance,

$$\forall i \in M : \Phi_i^0(\mathbf{C}) > 0, \Phi_i^1(\mathbf{C}) > 0.$$

5. In homogenous interactions, intolerance has a positive social impact and the intolerant player gets a reward from intolerance,

$$\forall i \in M : \rho_i^1(\mathbf{C}) > 0, \rho_i^0(\mathbf{C}) > 0, \rho_i^2(\mathbf{C}) > 0.$$

Symmetry anticipates that no additional factors beyond the confession itself determine the relative interest in confessions, so the identity of players in the game is irrelevant. Thereby, we abstract from real-world instances of asymmetries, such as the private interests of the church establishment, discussed in Section 4.4. Additivity, which is the crucial property of Independent taboo model, helps to solve the game by isolating individual taboos. Thereby, we can infer optimal decision within each taboo without focusing on the composition of confessions in remaining taboos.

Second, tolerance is indifference, so we assume that tolerance brings neither benefits nor costs for either the tolerant player or his tolerant opponent. A tolerant player simply doesn't care; he doesn't feel any benefit himself, and has no social need to reward or harm the opponent. The assumption describes interactions of two tolerant players, who obtain payoff strictly on the basis of truth.

For intolerance, the picture is much more colorful; in heterogenous interaction, intolerance is beneficial to both the intolerant player and his opponent. An intolerant player enjoys similarity, and rewards the other player for being similar. In homogenous society, the complete reverse holds. The intolerant player is deprived, and is hostile to the opponent. Both suffer a cost, though not necessarily of equal size.

We can see that intolerance dominates for homogeneity and tolerance dominates for heterogeneity. Meeting a person with the same confession, one would like to enjoy benefits of the social function. Meeting a person with a different confession is rather discomfort to both players; in such a case, tolerance is better, because tolerance is nothing but indifference.

Finally, since we perfectly separated access to truth and access to community in the payoff function, we can very easily define a vector of truthful taboos. This is a vector $\mathbf{a}^*$ corresponding to an individually efficient confession $(\mathbf{a}^*, \mathbf{0})$. The individually efficient confession can be identified in a

Pareto-optimum from all homogeneous, fully-tolerant populations. "Truthful" taboos thus yield the highest possible payoff in homogeneous, fully-tolerant societies.

Definition 8 (Truth): We call $\mathbf{a}^* \in \{0,1\}^m$ truthful taboos, if

$$\forall \mathbf{a} \in \{0,1\}^m, \mathbf{a} \neq \mathbf{a}^* : \pi_A(\mathbf{a}^*, \mathbf{0}, \mathbf{a}^*, \mathbf{0}) > \pi_A(\mathbf{a}, \mathbf{0}, \mathbf{a}, \mathbf{0}). \tag{4.1}$$

4.3.2 Evolutionary game

Let us have a large population where each individual has a confession $(\mathbf{x}, \mathbf{y}) \in \Theta_p$. In an infinite number of repetitions, two individuals are randomly drawn from the population to play a pure-strategy stage-game with the confessions being the (inherited, not chosen) strategies, where payoff function writes π. Denote the individuals A, B and the respective confessions $\mathbf{C}_A = (\mathbf{a}, \mathbf{u})$ and $\mathbf{C}_B = (\mathbf{b}, \mathbf{v})$. After the stage game, they observe the payoffs $\pi(\mathbf{C}_A, \mathbf{C}_B)$, and adjust confessions in the following way:

$$\begin{aligned} \pi_A(\mathbf{C}_A, \mathbf{C}_B) > \pi_B(\mathbf{C}_A, \mathbf{C}_B) &\Rightarrow \mathbf{C}_A = \mathbf{C}_B = (\mathbf{a}, \mathbf{u}) \\ \pi_A(\mathbf{C}_A, \mathbf{C}_B) < \pi_B(\mathbf{C}_A, \mathbf{C}_B) &\Rightarrow \mathbf{C}_B = \mathbf{C}_A = (\mathbf{b}, \mathbf{v}) \\ \pi_A(\mathbf{C}_A, \mathbf{C}_B) = \pi_B(\mathbf{C}_A, \mathbf{C}_B) &\Rightarrow \mathbf{C}_A, \mathbf{C}_B \in \{(\mathbf{a}, \mathbf{u}), (\mathbf{b}, \mathbf{v})\} \end{aligned} \tag{4.2}$$

How can we interpret this evolution of confessions? A priori, people cannot know what level of satisfaction (payoff) the alternative confessions bring, unless they are tried. An incentive to try a new confession arrives in the interaction with a randomly assigned partner, as long as it is possible to observe and compare satisfaction (payoff) of the partner. The interaction represents the random opportunity to compare confession and adjust to it, which, as we have discussed, is best approximated by large non-specialized media. Following assumptions above, we use here that people are social animals, so beside the truth defined in (4.1), they tend to praise homogeneity (i.e. engage in intolerance) when the others are similar and praise for heterogeneity (i.e. are tolerant) when the others are different. The brings interesting features into the evolutionary game.

This adjustment dynamics in (4.2) reveals that a confession survives only if it is a best-reply to itself, i.e. Nash equilibrium.

Definition 9 (Best replies): A pure-strategy best reply for player A to a strategy profile $\mathbf{C} = (\mathbf{C}_A, \mathbf{C}_B) \in \Theta_p$ is the correspondence $\tilde{\beta}_A : \Theta_p \to S_p$ such that:

$$\tilde{\beta}_A(\mathbf{C}) = \{\mathbf{C}_A \in S_p : \pi_A(\mathbf{C}_A, \mathbf{C}_B) \geq \pi_A(\tilde{\mathbf{C}}_A, \mathbf{C}_B), \forall \tilde{\mathbf{C}}_A \in S_p\} \tag{4.3}$$

A strategy profile $\mathbf{C} \in \Theta_p$ belongs into the set of *pure strategy Nash equilibria* $\Theta_p^{NE} \subseteq \Theta_p$, if and only if it consists of mutual best replies:

$$\mathbf{C} = (\mathbf{C}_A, \mathbf{C}_B) \in \Theta_p^{NE} \Leftrightarrow \mathbf{C}_A \in \tilde{\beta}_A(\mathbf{C}_B) \wedge \mathbf{C}_B \in \tilde{\beta}_B(\mathbf{C}_A) \tag{4.4}$$

For a strict Nash equilibrium, $\{\mathbf{C}_A\} = \tilde{\beta}_A(\mathbf{C}_B)$ and $\{\mathbf{C}_B\} = \tilde{\beta}_B(\mathbf{C}_A)$.

This is however insufficient. Suppose a mutant confession $\mathbf{C}_2 \in S_p$ survives in a normal population of $\mathbf{C}_1$, because it is a best reply to the normal confession: $\mathbf{C}_2 \in \tilde{\beta}_A(\mathbf{C}_2, \mathbf{C}_1)$. Then, the symmetric strategy profile of normal confessions constitutes only a weak Nash equilibrium, and we have to consider also how fast $\mathbf{C}_1$ adjusts in a population of mutants $\mathbf{C}_2$. To eliminate mutants, we need the adjustment to be faster, so $\pi_A(\mathbf{C}_1, \mathbf{C}_2) > \pi_A(\mathbf{C}_2, \mathbf{C}_2)$. With these preliminaries, we can define the evolutionary stability concept formally (for alternative concepts, see Weibull 1995).

Definition 10 (Evolutionary stability): Define the set of evolutionarily stable states Θ_p^{ESS} as all symmetric strategy profiles $(\mathbf{C}_1, \mathbf{C}_1)$ that satisfy the first-order and the second-order best reply conditions for any $\mathbf{C}_2 \in S_p, \mathbf{C}_2 \neq \mathbf{C}_1$:

$$\pi_A(\mathbf{C}_2, \mathbf{C}_1) \leq \pi_A(\mathbf{C}_1, \mathbf{C}_1) \tag{4.5}$$

$$\pi_A(\mathbf{C}_2, \mathbf{C}_1) = \pi_A(\mathbf{C}_1, \mathbf{C}_1) \Rightarrow \pi_A(\mathbf{C}_1, \mathbf{C}_2) > \pi_A(\mathbf{C}_2, \mathbf{C}_2) \tag{4.6}$$

4.3.3 Solutions

Obviously, we have $\Theta_p^{ESS} \subseteq \Theta_p^{NE}$. In hunt for evolutionary stability, we thus firstly examine Nash equilibria of the game, for which we have to identify best replies.

Lemma 9: Denote the set of i-th coordinates of the best replies of Player A as $\tilde{\beta}_A^i(\mathbf{C})$, $\tilde{\beta}_A^i(\mathbf{C}) \equiv \{(a_i, u_i) : (\mathbf{a}, \mathbf{u}) \in \tilde{\beta}_A(\mathbf{C})\}$. Then, for any $i \in M$

$$\forall \mathbf{C}, \bar{\mathbf{C}} \in \Theta_p : a_i = \bar{a}_i, u_i = \bar{u}_i, b_i = \bar{b}_i, v_i = \bar{v}_i \Rightarrow \tilde{\beta}_A^i(\mathbf{C}) = \tilde{\beta}_A^i(\bar{\mathbf{C}}). \tag{4.7}$$

Proof For $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v})$, let $(\mathbf{a}, \mathbf{u}) \in \tilde{\beta}_A(\mathbf{C})$. By contradiction, for some $i \in M$ suppose $(a_i, u_i) \notin \tilde{\beta}_A^i(c)$. By best-reply definition, we have $\forall (\bar{\mathbf{a}}, \bar{\mathbf{u}}) \in S_p$,

$$\pi_A(\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \geq \pi_A(\bar{\mathbf{a}}, \bar{\mathbf{u}}, \mathbf{b}, \mathbf{v}) \Rightarrow \pi_A(\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \geq \pi_A((\bar{a}_i, a_{-i}), (\bar{u}_i, u_{-i}), \mathbf{b}, \mathbf{v}).$$

By additivity in Ass. 1, this inequality holds also when any of j-th elements $(j \neq i)$ changes:

$$\pi_A((a_i, \bar{a}_{-i}), (u_i, \bar{u}_{-i}), (b_i, b_{-i}), (v_i, v_{-i})) \geq \pi_A(\bar{\mathbf{a}}, \bar{\mathbf{u}}, (b_i, b_{-i}), (v_i, v_{-i})). \quad \square$$

We call the sets of i-th coordinates of the best replies as *element-specific best replies*. Lemma 9 explains that the element-specific best replies are invariant to changes out of the i-th element. Not only that; by definition of additivity, the vector differences change not only sign, but remain even constant. Moreover, we exploit the fact that the if taboo is not truthful in i-th element ($a_i^* = 0$), the difference Δ_i must be positive, and that truthful taboo ($a_u^* = 1$) implies negative difference.

Lemma 10: For any $i \in M$, $a_i^* = 0 \Rightarrow \Delta_i > 0$ and $a_i^* = 1 \Rightarrow \Delta_i < 0$.

Proof By definition of vector difference $\Delta_i(\cdot)$, we have $\Delta_i := \pi_A(0,0,0,0,\mathbf{C}) - \pi_A(1,0,1,0,\mathbf{C})$. By additivity, we generalize that $\Delta_i = const.$ for any $\mathbf{C}$. By definition of truth, $a_i^* = 0 \Rightarrow \pi_A(0,0,0,0,\mathbf{C}) > \pi_A(1,0,1,0,\mathbf{C})$, which is equivalent to $a_i^* = 0 \Rightarrow \Delta_i > 0$. By definition of truth, $a_i^* = 1$ implies $\pi_A(0,0,0,0,\mathbf{C}) < \pi_A(1,0,1,0,\mathbf{C})$, equivalent to $a_i^* = 1 \Rightarrow \Delta_i < 0$. □

Because of Lemma 9, we can split the search for best replies into m elements, and solve for best replies when strategy sets are reduced to the i-th element. This is the main property of Independent Taboo model. In finding pure-strategy Nash equilibria, we use that $\Phi_i^0, \Phi_i^1, \rho_i^0, \rho_i^1, \rho_i^2 > 0$ holds by Ass. 1, which we input into Table 4.1.

Proposition 14 (Nash equilibria in pure strategies): Any Nash equilibrium pure-strategy profile $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \in \Theta_p^{NE}$ satisfies for any $i \in M$:

$$\begin{aligned} a_i^* = 0 \wedge \Delta_i \leq \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Longrightarrow a_i = b_i = u_i = v_i \in \{0,1\} \\ a_i^* = 0 \wedge \Delta_i > \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Longrightarrow a_i = b_i = u_i = v_i = 0 \\ a_i^* = 1 &\Longrightarrow a_i = b_i = u_i = v_i = 1 \end{aligned}$$

Proof The existence of a Nash equilibrium in pure strategies is illustrated by the tables based on Table 4.1, assuming without loss of generality that $\pi_A(1,0,1,0,\mathbf{C}) = \pi_B(1,0,1,0,\mathbf{C}) = 0$. For convenience, we depict only best responses, and the other responses are blank. To recognize which of the two payoffs in a pair (π_A, π_B) represents the best reply, I use for the irrelevant payoff symbol $\times$. Consider firstly elements, where $a_i^* = 1$.

Tab. 4.2: Payoffs (π_A, π_B) $(a_i^* = 1)$

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$		$\times$, 0	
$1,0$	$0,\times$		$\times$, ρ_i^0
$1,1$		$\rho_i^0,\times$	$\rho_i^0 + \rho_i^1$, $\rho_i^0 + \rho_i^1$

As a result, the only strict Nash equilibrium is $a_i = b_i = u_i = v_i = 1$, which proves the third part of the proposition. Consider now $a_i^* = 0$. We distinguish among five cases, depending on $\Delta_i <> \rho_i^0$ and $\Delta_i <> \rho_i^0 + \rho_i^1 + \Phi_i^0$.

Tab. 4.3: Payoffs (π_A, π_B) $(a_i^* = 0,\ \Delta_i < \rho_i^0)$

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i,Δ_i		
$1,0$			$\times$, ρ_i^0
$1,1$		ρ_i^0, $\times$	$\rho_i^0 + \rho_i^1$, $\rho_i^0 + \rho_i^1$

Tab. 4.4: Payoffs (π_A, π_B) $(a_i^* = 0,\ \Delta_i = \rho_i^0)$

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i, Δ_i	Δ_i, $\times$	
$1,0$	$\times$, Δ_i		$\times$, ρ_i^0
$1,1$		ρ_i^0, $\times$	$\rho_i^0 + \rho_i^1$, $\rho_i^0 + \rho_i^1$

Tab. 4.5: Payoffs (π_A, π_B) $(a_i^* = 0,\ \rho_i^0 < \Delta_i < \rho_i^0 + \rho_i^1 + \Phi_i^0)$

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i, Δ_i	Δ_i, $\times$	
$1,0$	$\times$, Δ_i		
$1,1$			$\rho_i^0 + \rho_i^1$, $\rho_i^0 + \rho_i^1$

In all the three cases, we have two strict equilibria, to be denoted as $(a_i, u_i, b_i, v_i) = (0,0,0,0)$ and $(a_i, u_i, b_i, v_i) = (1,1,1,1)$. Interestingly, condition $\Delta_i <> \rho_i^0$ is irrelevant as it relates to out-of-equilibrium best replies.

Tab. 4.6: Payoffs (π_A, π_B) ($a_i^* = 0$, $\rho_i^0 + \rho_i^1 + \Phi_i^0 = \Delta_i$)

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i, Δ_i	Δ_i, $\times$	$\Delta_i - \Phi_i^0$, $\times$
$1,0$	$\times$, Δ_i		
$1,1$	$\times$, $\Delta_i - \Phi_i^0$		$\rho_i^0 + \rho_i^1$, $\rho_i^0 + \rho_i^1$

Still, there are two equilibria. Yet, the intolerant equilibrium $(1,1,1,1)$ is not strict anymore. This proves the first part of the proposition.

Tab. 4.7: Payoffs (π_A, π_B) ($a_i^* = 0$, $\rho_i^0 + \rho_i^1 + \Phi_i^0 < \Delta_i$)

$a_i, u_i/b_i, v_i$	$0,0$	$1,0$	$1,1$
$0,0$	Δ_i, Δ_i	Δ_i, $\times$	$\Delta_i - \Phi_i^0$, $\times$
$1,0$	$\times$, Δ_i		
$1,1$	$\times$, $\Delta_i - \Phi_i^0$		

The equilibrium $(1,1,1,1)$ vanishes and we end up in a unique strict Nash equilibrium $(0,0,0,0)$. □

In order to obtain Θ_p^{ESS}, we generally eliminate asymmetric strategy profiles and unstable weak Nash equilibria from the set Θ_p^{NE}.

Proposition 15 (Evolutionary stability in pure strategies): All evolutionarily stable pure-strategy profiles $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \in \Theta_p^{ESS}$ satisfy for any $i \in M$:

$$a_i^* = 0 \wedge \Delta_i < \Phi_i^0 + \rho_i^0 + \rho_i^1 \Longrightarrow a_i = b_i = u_i = v_i \in \{0,1\}$$
$$a_i^* = 0 \wedge \Delta_i > \Phi_i^0 + \rho_i^0 + \rho_i^1 \Longrightarrow a_i = b_i = u_i = v_i = 0$$
$$a_i^* = 1 \Longrightarrow a_i = b_i = u_i = v_i = 1$$

Proof We have to eliminate asymmetric equilibria and non-stable weak equilibria from the set Θ_p^{NE}. The former ones are not in the game, but we have one weak Nash equilibrium $(a_i, u_i, b_i, v_i) = (1,1,1,1)$ for $a_i^* = 0$ and $\Delta_i = \rho_i^0 + \rho_i^1 + \Phi_i^0$). Here, both $(1,1)$ and $(0,0)$ are the best replies of Player A. For $(1,1,1,1)$ to be evolutionarily stable, we need that $(1,1)$ reproduces faster in a population of "mutants" $(0,0)$, $\pi_A(1,1,0,0,\mathbf{C}) > \pi_A(0,0,0,0,\mathbf{C})$. This is not satisfied, because $\pi_A(1,1,0,0,\mathbf{C}) = -\Phi_i^1 < \Delta_i = \pi_A(0,0,0,0,\mathbf{C})$. □

Example Consider $m = 5$, $\Delta_2 < \Phi_2^0 + \rho_2^0 + \rho_2^1$, $\Delta_4 = \Phi_4^0 + \rho_4^0 + \rho_4^1$, and $\Delta_5 > \Phi_5^0 + \rho_5^0 + \rho_5^1$. If $a^{*\prime} = (1,0,1,0,0)$, then $\Theta_p^{NE} = \{\mathbf{C^1}, \mathbf{C^2}, \mathbf{C^3}, \mathbf{C^4}\}$ and $\Theta_p^{ESS} = \{\mathbf{C^1}, \mathbf{C^2}\}$, where:

$$\mathbf{C^1} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \mathbf{C^2} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\mathbf{C^3} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \mathbf{C^4} = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

4.3.4 Interpretation

The solutions allow for straightforward interpretation. If a taboo is truthful, it is always reinforced by intolerance and is evolutionarily stable. Therefore, all freedoms remaining in a stable state are necessarily efficient. The stability of taboo is however not a sufficient indicator of efficiency; if the social function of taboo is strong enough, it overrides attractiveness of tolerant, truthful freedom. The critical condition for each i-th element is $\Delta_i < \Phi_i^0 + \rho_i^0 + \rho_i^1$. On the left side, we have benefit from truth, and on the right side, gains of intolerance; recall Φ^0 denotes absence of discomfort (penalty) from meeting an intolerant person, ρ^0 is the intrinsic reward from sharing taboo, and ρ^1 is the external reward for like-minded players, paid by intolerant persons.

When freedom is truthful, we thus get a coordination game, where stability depends on the initial distribution of population. If the majority of population initially believes in all freedoms and mutants are rare, the population always necessarily arrives at a socially efficient state, while the population starting with comprehensive list of taboos maintains all true taboos, but also those false taboos that satisfy the condition above. This is in line with divergent evolutions of confessions in different societies.

Efficiency implications of the model in pure strategies are twofold. First, since efficient taboos reinforce themselves again any mutations, there is *no policy reason to allow any involvement in enforcement of taboos.* An efficient taboo never disappears in a society tolerant in a sense of absence of legal restrictions. Contrary to common wisdom, the rule of free entry is not detrimental to any taboo; in fact it may weaken stability of freedoms.

Second, a socially efficient state sustains small invasions of (either too liberal or too restrictive) "mutants". There may nevertheless be a large amount of stable states with inferior social efficiency, as long as temptation from intolerance overrides benefits of truth. Hence, there is a case for support of taboo breakers, because they could shift the population from an inefficient state toward a more efficient state.

The central problem of this intervention is that by assumption, payoffs are realized only in bilateral interactions, players are boundedly rational, and a policy maker is presumably no different from ordinary players, so there is no guide for intervening policy maker. Although some subsidies of no-taboos (e.g., cultural subsidies) may swing the socially inefficient taboo into socially efficient freedom, in other cases, the subsidy turns into pure waste. Unless the policy maker has additional information on the distribution of truth across taboos and is able to calculate expected benefits of the intervention, he cannot conclude that certain freedoms deserve policy support. The only unambiguous constitutional rule is liberal: protect free entry and forbid protection of any particular confession.

4.3.5 Mixed strategies

Pure strategies are only special cases of mixed strategies when confessions don't randomize. One could successfully argue against this reduction, and introduce mixed strategies. The first (*monomorphic*) interpretation of mixing is nonetheless largely implausible. It admits that a confession is binding, perceived as expression of truth, but at the same time probabilistic (sometimes obeyed, sometimes not). This is uneasy to capture in philosophical terms. The other interpretation of mixed strategies, *polymorphic* interpretation, is rather uncontroversial—a mixed strategy is a heterogenous population where groups of individuals playing different pure strategies have various shares.

Generally, one only needs to re-define all arguments of pure-strategies in terms of probability measures to introduce mixed strategies. For example, a_i shall be re-written as a vector of coordinates interpreted as probability measures of regulation and non-regulation, belonging to a unit simplex. Since the projection of this unit simplex is into the interval $[0, 1]$, one can represent such a vector by a single value, denoted as $\alpha_i \in [0, 1]$. This method maintains reasonable notation. Pure strategies thereby become only special cases when some of coordinates take value of one, and the others zero.

Definition 11 (Polymorphic populations): For all $i \in M$, interpret α_i and β_i as the players' probabilities of choosing regulation of element i, and let μ_i and υ_i be the probabilities of their intolerance in the element i. By $\mathbf{C}^m_A = (\alpha, \mu)$ and $\mathbf{C}^m_B = (\beta, \upsilon)$, denote the mixed strategies of Players A and B. Define the identical mixed strategy set for Players A and B as $S_m = \{(\mathbf{x}, \mathbf{y}) \in [0,1]^{2m} \times [0,1]^{2m} : x_i \in [0,1] \wedge y_i \leq x_i\}$. Let $\mathbf{C}^m = (\mathbf{C}^m_A, \mathbf{C}^m_B) \in \Theta_m$ be the mixed-strategy profile of Players A and B, where the set of strategy profiles writes as $\Theta_m = S_m \times S_m$.

In order to get the expected payoffs in mixed strategies, we have to find conditional probabilities of realizations of all strategy profiles $\tilde{\mathbf{C}}$ in the definition set of the payoff function.

Definition 12 (Mixed-strategy payoff): For each $\mathbf{C} \in \Theta_p$, denote $\mathbf{a}(\mathbf{C})$ as the vector corresponding to that profile, namely $\mathbf{C}((\mathbf{a}(\mathbf{C}), \mathbf{u}) \quad \mathbf{C}_B) = \mathbf{C}$. By analogy, define $\mathbf{b}(\mathbf{C})$, $\mathbf{u}(\mathbf{C})$, and $\mathbf{v}(\mathbf{C})$. The expected value of Player A associated with any mixed strategy profile $\mathbf{C}^m \in \Theta_m$ is:

$$U_A(\mathbf{C}) = \sum_{\tilde{\mathbf{C}} \in \Theta_p} \pi_A(\tilde{\mathbf{C}}) \alpha_{\mathbf{a}(\tilde{\mathbf{C}})}(\mathbf{C}) \mu_{\mathbf{u}(\tilde{\mathbf{C}})}(\mathbf{C}) \beta_{\mathbf{b}(\tilde{\mathbf{C}})}(\mathbf{C}) \upsilon_{\mathbf{v}(\tilde{\mathbf{C}})}(\mathbf{C}) \tag{4.8}$$

The combined mixed strategy payoff function is $U : \Theta_m \to \mathbb{R}^2$, where $U(\mathbf{C}) = (U_A(\mathbf{C}), U_B(\mathbf{C}))$.

In order to solve the game, we redefine the best replies, Nash equilibrium, and evolutionary stability.

Definition 13 (Mixed-strategy best replies): Define a mixed best reply for player A to a strategy profile $\mathbf{C} \in \Theta_m$ as the correspondence $\tilde{\beta}^m_A : \Theta_m \to S_m$.

$$\tilde{\beta}^m_A(\mathbf{C}) = \{\mathbf{C}_A \in S_m : U_A(\mathbf{C}_A, \mathbf{C}_B) \geq U_A(\tilde{\mathbf{C}}_A, \mathbf{C}_B) \forall \tilde{\mathbf{C}}_A \in S_m\} \tag{4.9}$$

The mixed best reply for player B, $\tilde{\beta}^m_B(\mathbf{C})$, is to be defined symmetrically. Let the set of mixed-strategy Nash equilibria, Θ^{NE}_m, write as follows:

$$\mathbf{C} = (\mathbf{C}_A, \mathbf{C}_B) \in \Theta^{NE}_m \Leftrightarrow \mathbf{C}_A \in \tilde{\beta}^m_A(\mathbf{C}) \wedge \mathbf{C}_B \in \tilde{\beta}^m_B(\mathbf{C}) \tag{4.10}$$

Let the evolutionary stability be analogical to the one in Definition 10, where S_m replaces S_p and we write Θ^{ESS}_m instead of Θ^{ESS}_p. With this notation, we can solve the game.

Proposition 16 (Nash equilibria in mixed strategies): Any Nash equilibrium pure-strategy profile $\mathbf{C} = (\alpha, \nu, \beta, \upsilon) \in \Theta_m^{NE}$ satisfies for any $i \in M$:

$$\begin{aligned} a_i^* = 0 \wedge \Delta_i \leq \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i \in \left\{0, \frac{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \Delta_i}{\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1}, 1\right\} \\ a_i^* = 0 \wedge \Delta_i > \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i = 0 \\ a_i^* = 1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i = 1 \end{aligned} \tag{4.11}$$

Proof First, we construct for each $\mathbf{C} \in \Theta_p^{NE}$ a mirror strategy profile C^m, where $\forall \in M : \alpha_i = a_i, \beta_i = b_i, \nu_i = u_i, \upsilon_i = v_i$. From Definitions 12 and 4.4, we conjecture that $C^m \in \Theta_m^{NE}$, because no combination of best replies in pure-strategies can give a mixed strategy with a higher payoff than the payoff of the best-replies. Therefore, we look for (symmetric) mixed strategies that are not pure. Moreover, we consider only the case of truthful freedom, because otherwise (for truthful taboo $a_i^* = 1$), all strategies beyond $(a_i, u_i) = (1, 1)$ are strictly dominated strategies, and of course cannot be part of a Nash equilibrium profile.

In each $i \in M$, we have $(a_i, u_i) \in \{(0, 0), (1, 0), (1, 1)\}$. Denote for convenience (x_i, y_i, z_i) as probabilities imposed on each feasible pure strategy in the i-th element (a_i, u_i), i.e. probabilities of $(0, 0)$,$(1, 0)$, and $(1, 1)$, where the measure is equal one, $x_i + y_i + z_i = 1$. We get that $\alpha_i = y_i + z_i$, $1 - \alpha_i = x_i$, $\nu_i = z_i$, and $1 - \nu_i = x_i + y_i$.

We can apply Lemma 9 and write the expected utility of a mixed strategy (x_i, y_i, z_i) playing against some strategy $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$, both in i-th element, as $U_A^i(\mathbf{C}) = (x_i, y_i, z_i; x_i, y_i, z_i)$. We can write utility that simply because all elements beyond the i-th element are irrelevant in the sense that they only linearly transform the expected utility, therefore don't change best responses. Now, we use the polymorphic interpretation by which in Nash equilibrium (x_i^*, y_i^*, z_i^*), each strategy that is present has to have the same expected utility (otherwise those with higher utility grow, and the others diminish). Consider first that all strategies are in the support, so $x_i > 0$, $y_i > 0$, and $z_i > 0$:

$$\begin{aligned} &U_A^i(1, 0, 0; x_i^*, y_i^*, z_i^*) = \Delta_i - \Phi_i^0 z_i^* \\ &U_A^i(0, 1, 0; x_i^*, y_i^*, z_i^*) = (\rho_i^0 + \rho_i^1 - \rho_i^2) z_i^* \\ &U_A^i(0, 0, 1; x_i^*, y_i^*, z_i^*) = \Phi_i^1 x_i^* + \rho_i^0 y_i^* + (\rho_i^0 + \rho_i^1) z_i^* \\ &U_A^i(1, 0, 0; x_i^*, y_i^*, z_i^*) = U_A^i(0, 1, 0; x_i^*, y_i^*, z_i^*) = U_A^i(0, 0, 1; x_i^*, y_i^*, z_i^*) \end{aligned}$$

A unique solution of this linear system is as follows:

$$x_i^* = \frac{\rho_i^0(\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2 - \Delta_i) + \Delta_i\rho_i^2}{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2}$$
$$y_i^* = \frac{\Phi_i^1(\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2 - \Delta_i) - \Delta_i\rho_i^2}{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2}$$
$$z_i^* = \frac{\Delta_i}{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2}$$

This solution satisfies $0 < x_i^*, y_i^*, z_i^* < 1$ if and only if:

$$\rho_i^2 > \rho_i^0 \wedge \Delta_i < \min\left\{\frac{\rho_i^0(\rho_i^0 + \rho_i^1 - \rho_i^2)}{\rho_i^2 - \rho_i^0}, \frac{\Phi_i^1(\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2)}{\rho_i^2 - \rho_i^0}, \frac{\Phi_i^1(\Phi_i^0 + \rho_i^0 + \rho_i^1 - 1) - \rho_i^2}{\Phi_i^1 + \rho_i^2}; \Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2\right\}$$
$$\vee\ \rho_i^2 \leq \rho_i^0 \wedge \Delta_i < \min\left\{\frac{\Phi_i^1(\Phi_i^0 + \rho_i^0 + \rho_i^1 - 1) - \rho_i^2}{\Phi_i^1 + \rho_i^2}; \Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2\right\}$$

Now, is this strategy a best response to itself? To check it, write the expected utility of any alternative strategy $U_A^i(x_i, y_i, z_i; x_i^*, y_i^*, z_i^*)$:

$$U_A^i(x_i, y_i, z_i; x_i^*, y_i^*, z_i^*) = \frac{\Delta_i(\rho_i^0 + \rho_i^1 - \rho_i^2)}{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \rho_i^2} - \Delta_i y_i$$

Obviously, $y_i = 0$ is the best response. Therefore, we can consider a special case when $z_i = 1 - x_i$ and $x_i, y_i > 0$. In other words, we can consider cases when only two strategies are in the support:

$$U_A^i(1, 0, 0; x_i^*, 0, 1 - x_i^*) = \Delta_i - \Phi_i^0(1 - x_i^*)$$
$$U_A^i(0, 0, 1; x_i^*, 0, 1 - x_i^*) = \Phi_i^1 x_i^* + (\rho_i^0 + \rho_i^1)(1 - x_i^*)$$
$$U_A^i(1, 0, 0; x_i^*, 0, 1 - x_i^*) = U_A^i(0, 0, 1; x_i^*, 0, 1 - x_i^*)$$

The solution is:

$$x_i^* = \frac{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \Delta_i}{\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1} \qquad 1 - x_i^* = \frac{\Delta_i + \Phi_i^1}{\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1}$$

To satisfy $0 < x_i^* < 1$, we only need $\Delta_i < \Phi_i^0 + \rho_i^0 + \rho_i^1$. Under this condition, we get that $(x_i^*, 0, 1 - x_i^*; x_i^*, 0, 1 - x_i^*)$ is a weak Nash equilibrium, because for any $x_i \in [0,1]$, $(x_i, 0, 1-x_i)$ is the best response to $(x_i^*, 0, 1-x_i^*)$:

$$U_A^i(x_i, 0, 1-x_i; x_i^*, 0, 1-x_i^*) = \frac{\Delta_i \rho_i^0 + \Delta_i \rho_i^1 + \Delta_i \Phi_i^1 - \Phi_i^0 \Phi_i^1}{\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1}$$
$$\frac{dU_A^i(x_i, 0, 1-x_i; x_i^*, 0, 1-x_i^*)}{dx_i} = 0$$

For completeness, we can check that there are no other corner solutions. Consider $z_i = 0$, i.e. $y_i = 1 - x_i$. We get that $U_A^i(1,0,0; x_i, 1-x_i, 0) = U_A^i(0,1,0; x_i, 1-x_i, 0) \Rightarrow \Delta_i = 0$, which contradicts Lemma 10, where $a_i^* = 0 \Rightarrow \Delta_i > 0$. The remaining non-pure case is for $x_i = 0$, i.e. $z_i = 1 - y_i$:

$$U_A^i(0,1,0;,0,y_i,1-y_i) = U_A^i(0,0,1;0,y_i,1-y_i) \Rightarrow y_i^* = \frac{\rho_i^2}{\rho_i^2 - \rho_i^0}$$

For $\rho_i^2 \geq \rho_i^0$, we have $y_i^* > 1$, which contradicts that $y_i < 1$. For $\rho_i^2 < \rho_i^0$, we have that $y_i > 0 \Rightarrow \rho_i^2 < 0$, which contradicts that $\rho_i^2 > 0$. □

We obtain Θ_m^{ESS} again by elimination of unstable weak Nash equilibria.

Proposition 17 (Evolutionary stability in mixed strategies): All evolutionarily stable pure-strategy profiles $\mathbf{C} = (\mathbf{a}, \mathbf{u}, \mathbf{b}, \mathbf{v}) \in \Theta_m^{ESS}$ satisfy for any $i \in M$:

$$\begin{aligned} a_i^* = 0 \wedge \Delta_i < \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i \in \{0,1\} \\ a_i^* = 0 \wedge \Delta_i > \Phi_i^0 + \rho_i^0 + \rho_i^1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i = 0 \\ a_i^* = 1 &\Rightarrow \alpha_i = \beta_i = \nu_i = \upsilon_i = 1 \end{aligned} \tag{4.12}$$

Proof By Definition 10, a weak Nash equilibrium is evolutionarily stable as long as strategies involved in it reproduce faster against alternative best-replies than the best replies reproduce to themselves:

$$\forall x_i \in [0,1], x_i \neq x_i^* :$$
$$U_A^i(x_i^*, 0, 1-x_i^*; x_i, 0, 1-x_i) - U_A^i(x_i, 0, 1-x_i; x_i, 0, 1-x_i) > 0$$

Let $\Delta U_A(x_i) \equiv U_A^i(x_i^*, 0, 1-x_i^*; x_i, 0, 1-x_i) - U_A^i(x_i, 0, 1-x_i; x_i, 0, 1-x_i)$.

We examine behavior of this polynomial in the surrounding of x_i^*:

$$\frac{d\Delta U_A(x_i)}{dx_i} = 2(\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1)^2 \left[-x_i + \frac{\Phi_i^0 + \rho_i^0 + \rho_i^1 - \Delta_i}{\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1}\right]$$

$$\frac{d\Delta U_A(x_i)}{dx_i}(x_i = x_i^*) = 0$$

$$\frac{d^2\Delta U_A(x_i)}{d{x_i}^2}(x_i = x_i^*) = -2(\Phi_i^0 + \Phi_i^1 + \rho_i^0 + \rho_i^1)^2 < 0$$

As a result, for any $x + \varepsilon$ or $x_i - \varepsilon$, we have $\Delta U_A(x_i) < 0$, so the strategy profile $(x_i^*, 0, 1 - x_i^*; x_i^*, 0, 1 - x_i^*)$ is not evolutionarily stable. □

The new mixed-strategy profiles that were absent in pure-strategy specifications are not evolutionarily stable. Although they are Nash equilibria, they are weak, and can be destabilized by even very small amounts of invaders. To recognize the shares in this weak polymorphic equilibrium is however useful since the shares indicate critical thresholds around which small amounts of invaders completely change the composition of population in the stable state. Of course, this is only for elements in which freedom is truthful and attractiveness of in truth Δ_i is not prohibitively large.

4.4 Extensions

Our framework is applicable to evolution of any norms, which exhibit both intrinsic and social effects. This is closely related to cultural economics, where qualities of the product are continuously tested, depend on the consumption of the others, and the ultimate payoff function is unknown. Interactions of competing ideas that are not easily verifiable (academia, philosophy, or aesthetics) could be the best subsumed into this framework, although interpretation of elements and restrictions may be less intuitive, and the role of reason arguably more pronounced than in the case of metaphysical confessions.

The crucial assumptions of this framework are repeated random interactions and symmetric stage-game. Both are nearly satisfied for competitive media, which tend to reflect population norms proportionally, and are proactive in cultural communication. Nonetheless, there are potentially large barriers to proportionality in the real world, deserving separate discussion.

Communication and interactions may be non-random, especially in undeveloped or fragmented societies. Language barriers and cultural path dependencies prohibit free flow of competing norms. Benefits of free entry thus may depend on having large markets, i.e. full-blown *globalization*. Inefficient taboos in fragmented, or clustered populations, imply larger obstacles to the truth than in an open society. An additional policy implication of our framework is thus to support openness of societies as the second-best instrument of tolerance.

In our framework, we don't study consequences of political involvement of confessions; we only see how socially successful are confessions that are socially hostile to non-believers. If confessions can protect themselves by putting enforcement on the others, *the political market is endogenous*, and the final outcome may be different. To solve this situation, one would need to examine when an intolerant confession changes the "rules of evolutionary game" by suppressing tolerant confessions with political power.

Private interest clearly modifies the stability of confessions. Those who have stake at running some confession hardly change their opinion. Fortunately, this may not be a too big a part of population. This *asymmetry* is more pronounced in the evolution of ideologies, where almost anyone has a private interest related to certain policy.

Media market can be uncompetitive. There may also be *communication bias* against certain values. If acceptance of a norm requires not only "carrot", but often also a "stick", media which cannot effectively invest into long-term concentration cannot transmit certain values as effectively as other types of interactions, like peers or family.

Media target groups attractive for advertising, especially those with high *purchasing power*. If certain confessions are correlated with purchasing power, it should also have a disproportionately large share in media. For instance, the classic trade-off "Protestants eat well, but Catholics sleep well" might transform into "Protestants run newspapers, but Catholics run for newspapers". Lipford and Tollison (2003) provide evidence on the relationship between income and religious participation (in our definition, a form of confession). Another example is exclusive (income-based) membership, like Scientology. Exclusivity is understood as additional social function, involving decreasing amount of interactions with non-members.

Explanations of heterogeneities present in contemporary societies are thus manifold. First and foremost, we are still out of the stable state, in the course of the long-run evolution (no end of history). Second, frequency and extent of interactions is endogenous to norms, so certain confessions preserve due to this induced localization effect. Third, some societies are so separate that they may have locally stable states. Fourth, media are likely not the best transmitter of deep norms because of detraction potential involved in marketing, accessibility, and time management of consumers.

4.5 Conclusion

Adjustment of confessions can be studied on the basis of bounded rationality. This chapter imposed several plausible restrictions on such dynamics in the market with confessions and studied whether the evolution of confessions produces the socially most efficient outcome. Individual efficiency of a confession (truth) has been complemented by a social function. This is consistent with recent evolutionary approaches to religion by which individual interests may subserve the interests of groups (Wilson 2002).

When intolerance prospers in homogenous societies, but decays in heterogeneous societies, and when tolerance allows to identify the efficient confession, we find that the most efficient confession is always evolutionarily stable. However, it is not necessarily unique.

On one hand, we may have stable societies with too many taboos and too much intolerance, like in Boyer (2001), where religion is adaptation to past environments, maladaptive in modern environments. On the other hand, intolerance cannot prohibit efficient confessions from existence, as long as there is a critical mass of tolerant believers switching from one stable state to another and if interactions are sufficiently random (global). From this point of view, plus considering policy maker's impossibility to determine optimal (metaphysically truthful) confession a priori, we can conclude that intolerance is tolerable, but the state shall also preserve free entry of no-taboos into market with confessions.

References

[1] Adams, D. (1979). *The Hitchhiker's Guide to the Galaxy.* London: Pen Books.

[2] Blum, U. & L. Dudley (2001). "Religion and economic growth: was Weber right?" *Journal of Evolutionary Economics*, 11, 207–230.

[3] Boyer, P. (2001). *Religion Explained.* New York: Basic Books.

[4] Durkheim, E. [1912] (1995). *The Elementary Forms of Religious Life.* New York: The Free Press.

[5] Friedman, D. (1998). "On economic applications of evolutionary game theory", *Journal of Evolutionary Economics*, 8, 15–43.

[6] Iannaccone, L. R. (1996). "Rational Choice: Framework for the Scientific Study of Religion", in L. A. Young, ed., *Rational Choice Theory & Religion*, Routledge, London, pp. 25–46.

[7] Iannaccone, L. R. (1996). "Introduction to the economics of religion", *Journal of Economic Literature*, 36, 1465–1496.

[8] Lipford, J. D. & R. D. Tollison (2003). "Religious participation and income", *Journal of Economic Behavior and Organization*, 51, 249–260.

[9] Weibull, J. (1995). *Evolutionary Game Theory.* Cambridge, Mass.: MIT Press.

[10] Wilson, D. S. (2002). *Darwin's Cathedral: Evolution, Religion, and the Nature of Society.* Chicago: University of Chicago Press.

5. THE PROS AND CONS OF BANKING SOCIALISM

5.1 Introduction

The literature on privatization in transition often abstracts from differences in the financial and non-financial sectors (see Biais and Perotti 2002). Nonetheless, to privatize banks may be a largely different task than to privatize state-owned enterprises. Bank privatization implies loss of the credit channel as a policy instrument, which might be costly for an employment-maximizing government, facing strong unions and private banks with large market power.

Gordon (2003) explains incomplete privatization by attempt to overcome distortions of corporate taxation; the government privatizes primarily capital-intensive firms, for distortions therein are relatively small. State-owned banks then provide cheap credit to privatized firms to overcome distortions. We complement this view by another plausible aim; the government may prefer credit control considering sticky wages and deadweight loss associated with market power of privatized banks.

In this chapter, we seek an explanation why the benevolent government has incentive not only to privatize enterprises, but, at the same time, to keep control of banks. The intuition is driven by experience of those transition economies which were reluctant to privatize banks but not enterprises; in the Czech Republic, known for the very fast voucher privatization, four major banks went private as late as in the end of 1990s. The situation when the government handed out enterprises, but kept significant control over credit, and consequently also employment, won a label of "banking socialism".

Three topics in transition economics are closely related to banking socialism: optimal sequencing of reform, modes of privatization, and the soft-budget constraint. Regarding sequencing (see, e.g., Lian and Wei 1998), we side-step the problem by studying a one-period economy. We thus focus on the short run, when voters evaluate the policy-maker retrospectively. It would be possible to extend the article into more periods, and allow growth in the economy affect herein exogenous parameters (e.g. efficiency gap), which would change the government's optimum in next stages in favor of bank privatization, and nicely justify banking socialism as a transient phenomenon. However, no additional explanation beyond a simple change in parameters would be embedded in that plain extension.

The second stream of transition literature concerns modes of privatization. On one side, there are influential papers emphasizing strategic ("Machiavellian") aspect of privatization, namely how the governing right-wing politicians structure the value of shareholdings to affect political preferences of the middle class (Biais and Perotti 2002). Dispersed ownership of capital is then a safeguard against possible renationalization, which explains why right-wing governments resort to underpricing and quick speed. Ahrend (2002) suggests a rather standard mechanism: the efficiency gains from private ownership can be redistributed into the hands of the policy-maker's constituency, namely via underpriced assets.

These explanations however omit the importance of employment for the transition government, and the widespread use of privatization to further extra political and economic policy goals (see the comprehensive empirical study by Jones *et al.* 1999). Moreover, the strategic motive to prevent from re-nationalization (ensure time-consistent privatization policy), at least in the case of banks, loses appeal in a global economy where transition countries struggle for FDIs. A related issue is whether extra effort into maximization of employment is necessarily detrimental to efficiency; Börner (2003) argues that the commitment not to influence the profit-maximizing employment may be socially suboptimal, which invites a normative issue of desirable privatization, addressed by this chapter as well.

The literature on bank privatization is not as far-reaching as on general privatization, and the main findings are nearly stylized facts: bank privatization improves bank efficiency (Clarke, Cull and Shirley 2005), the quality of the nation's banking sector is significant determinant of bank privatization, but the political variables are not (Boehmer, Nash and Netter 2005). The latter invites an immediate inference that it is the median voter who is decisive. Therefore, we stick to median voter model in the assumptions instead of modeling special interest politics.

As to the explanation of delays in bank privatization, Ambrus-Lakatos and Hege (1998) suggest that the bank sale may lead to excessive liquidation or even credit crunch. The reason is that the private owner has to signal hard-budget constraint to minimize moral hazard, but the signalling is costly. Here, we have complete information without moral hazard, thus the model gives a complementary explanation of delay.

The third extensive stream of literature, concerned with the soft-budget constraint, is the most helpful for the purpose of our chapter. Privatization is not only driven by efficiency gains and the ensuing transfers, but also by employment motive when labor markets feature some significant barrier to adjustment. Desai and Olofsgard (2006) attribute the employment subsidies to asymmetric information on the incumbent government's inability to promote job creation. This explains why first-best policy, i.e. a comprehensive reform, is not adopted; the less competent politicians mimic the more competent ones by using hidden subsidies. We argue that information asymmetry is plausible yet not necessary; subsidies occur even when interests of the government is aligned with interests of the median voter as in early (normative) approach to soft-budget constraint, emphasized by Kornai, Maskin and Roland (2003).

In this chapter, we consider private ownership genuinely private. Thereby, we abstract from institutional maze involved, for instance, in the Czech banking sector, where the state-owned banks in fact exerted ownership control over the privatized enterprises. In that particular case, the lack of alternative domestic private capital (shallow pocket) empowered banks to launch major investment privatization funds, purchasing shares in voucher privatization, and the non-privatized banks became ultimate owners of "privatized" enterprises.

In Section 5.2, we construct a transition economy with nominal wage rigidity and market power of banks. We find equilibria in four institutional modes (socialism, industrial socialism, banking socialism, and capitalism) in Section 5.3. In the next Section 5.4, we enrich the analysis of state-owned enterprises by the possibility to use profits for wage or credit subsidies. The last Section 5.5 concludes.

5.2 The economy

5.2.1 Enterprises

The economy consists of the brownfield (B) and the greenfield (G) sector. The brown companies are owned by the government, and can be privatized. The green companies are owned by the private sector and cannot be nationalized. Production requires labor L, capital K, and technology A, aggregated by a Cobb-Douglas function with parameters 1/2 for labor and

1/4 for capital,[1]

$$Y_B = A_B L_B^{1/2} K_B^{1/4}, \qquad Y_G = A_G L_G^{1/2} K_G^{1/4}. \tag{5.1}$$

The brown sector starts with an old technology, $A_B := a$; if restructured, the sector can imitate a new technology used by green companies, $A_G := 1$, where $a < 1$. Thus, prior restructuralization, the better technology is situated in greenfields, while after restructuralization, both sectors are identical.

The output is sold on international market for fixed price P. The unit labor cost is w, and capital is purchased for r_G and r_B (in general case, we allow banks to discriminate between green and brown sectors).

The private owners always maximize profits, that is they seek $(L_G, K_G) = \arg\max \pi_G = \arg\max PY_G - wL_G - r_G K_G$. Denote $\Delta := P^4/64w^3$. We get

$$(L_G, K_G) = (2\Delta r_G^{-1}, w\Delta r_G^{-2}). \tag{5.2}$$

The government maximizes employment. In the elementary case (section 3), it simply maximizes labor demand of brownfields. In a more sophisticated case (section 4), the government can use profits of brownfields and redistribute them as pure transfers or wage subsidies. In the elementary case, the government is only restrained by the necessity to generate non-negative profits, $\pi_B \geq 0$. For any Cobb-Douglas production function, even zero amount of capital contributes to the production, so the maximization of labor yields $K_B = 0$ and we obtain

$$(L_B, K_B) = \begin{cases} (P^2 a^2 w^{-2}, 0) & \text{if state-owned,} \\ (2\Delta r_B^{-1}, w\Delta r_B^{-2}) & \text{if privatized.} \end{cases}$$

5.2.2 Labor market

Let n be the measure of individuals in the economy. An individual i, $i = 1..n$, is working l units of time and enjoys $1 - l$ of free time, where $l \in \{0, 1/2\}$. This means that only full-time contracts are available. Moreover, throughout the text, assume a constant wage rate w. The assumptions of discrete choice over extent of job and of extreme wage rigidity can be justified by the legacy of the socialist labor code, where the unions had a

[1] The specific parameters are algebraically convenient. A more important limitation is that we use negative returns to scale, which allows us to avoid corner solutions for extremely rigid labor market. With exactly zero returns to scale, a more complex model would emerge.

prominent position.[2] Notice that we forgo any heterogeneity in skills to capture aggregate effects.[3]

We define a quasi-linear utility function, where $f(1-l)$ is for utility of leisure time, $g(wl)$ denotes utility of income, and $\omega_i \in [0,1]$ denotes the individual preference for work,

$$U_i(w,l) = f(1-l) + \omega_i g(wl). \tag{5.3}$$

For convenience, we use $f(x) = x^{1/2}$, $g(x) = 2x$, which satisfies unique optimum due to $f_x < 0$ (marginal utility decreasing in leisure time) and $g_x = 2$ (marginal utility constant in working time). Suppose ω is uniformly distributed on the interval $[0,1]$, $\omega \sim U[0,1]$.

Individual labor supply l_i^* is given by a binary choice, as the individual compares the utilities of accepting and declining a full-time job offer, $U_i(w,0)$ and $U_i(w,1/2)$. In other words, the individual labor supply l_i^* is derived from the individual reservation wage, for which the utilities are equal,

$$l_i^* = \begin{cases} 0 & \text{if } w < \frac{2-\sqrt{2}}{2\omega_i}, \\ \frac{1}{2} & \text{if } w \geq \frac{2-\sqrt{2}}{2\omega_i}. \end{cases}$$

Alternatively, for each wage, we can identify the marginal individual, who is indifferent between job and unemployment, as $\hat{\omega}(w) = \frac{2-\sqrt{2}}{2w}$. Total labor supply is $L_S = \sum_i l_i^* = \frac{n}{2}(1-\hat{\omega})$ (recall uniform distribution of ω).

In market with hypothetically excess demand ($L_S \leq L_D$), each working individual is satisfied. In market with excess supply (standard case of sticky-wage market), unemployment appears at amount $1 - \frac{L_D}{L_S}$; we assume random rationing, i.e. individuals with non-zero individual labor supply are randomly matched with available jobs. Hence, utility function depends not only on individual labor supply, but also on the status of the non-clearing job market:

$$\begin{aligned} U_i|_{\omega_i<\hat{\omega}} &= U_i(0) \\ U_i|_{\omega_i\geq\hat{\omega};L_S\leq L_D} &= U_i(\tfrac{1}{2}) \\ E(U_i|_{\omega_i\geq\hat{\omega};L_S>L_D}) &= \frac{L_D}{L_S}U_i(\tfrac{1}{2}) + \left(1 - \frac{L_D}{L_S}\right)U_i(0) \end{aligned}$$

[2] In fact we need only partial rigidity to derive our main results, but allowing for adjustment in wages would bring only excessive notation. We could also do without constraints on contracts, with the effect of a more complicated total labor supply.

[3] Heterogeneity of workers in a spirit of Balla *et. al* (2005) would not change the model; if labor demand exceeded the size of high-skilled workers, they all would be employed, and the unemployment would affect only low-skilled workers. If not all high-skilled workers could be employed, the low-skilled workers would be beyond reach of a policy maker.

We get the first comparative properties. An increase in wage and labor demand increases utility ($\frac{dU_i}{dw} \geq 0$, $\frac{dU_i}{dL_D} \geq 0$), while an increase in population enlarges the probability of unemployment when seeking a job offer, thus diminishes utility, $\frac{dU_i}{dn} \leq 0$.

5.2.3 Credit market

The banks set interest rate for creditors at r_C, and receive savings $S(r_C) = \sigma r_C$. Then, they distribute all credit to brown companies for r_B and to green companies for r_G, $K_B(r_B) + K_G(r_G) = S(r_C)$. Suppose that savings come from abroad, so the level of savings has no direct influence on utility of individuals considered in the model.

Efficiency of banking system is denoted by the parameter σ, where $\sigma \in \{\overline{\sigma}, \underline{\sigma}\}$. The more efficient banking system, the more savings are attracted for the identical level of interest rate. Although this definition looks like marketing efficiency, it can represent operational efficiency, risk management, and the like. Transition economy obviously starts with non-efficient banking, $\sigma = \underline{\sigma}$, while privatized banking system features $\sigma = \overline{\sigma} > \underline{\sigma}$, which reflects an assumption that the private owners dispose with contract and profit-sharing schemes which motivate managers to restructure.

5.2.4 Private banks

Besides improved operational efficiency, assume that the private banks have market power over creditors, given by stringent entry requirements and other banking sector regulation. Specifically, they can discriminate in price between marginal and inframarginal creditors. The marginal creditor is paid r_G for his credit bid, while the inframarginal creditor receives some $r < r_C$, which still exceeds opportunity costs of capital. As a result, the banks don't pay $r_C S$ for savings at amount $S(r_C)$; suppose they pay exactly $\frac{3}{4} r_C S$. [4]

Private owners maximize profits, π. We use that a profit-maximizing banker sets $r_G = r_B$, which is derived in Section 5.3.3. Thus, we can write $K(r_G) := K_G(r_G) + K_B(r_G)$, and private banks set $(r_C, r_G) = \arg\max \pi = \arg\max r_G K(r_G) - \frac{3}{4}\overline{\sigma} r_C^2$, under feasibility constraint $K(r_G) \leq S(r_C)$,

$$(r_C, r_G) = K(r_G) \left(\frac{1}{\overline{\sigma}}, \frac{3}{2\overline{\sigma}} - \frac{1}{\frac{dK(r_G)}{dr_G}} \right). \tag{5.4}$$

[4] Perfect discrimination would be for costs $\frac{1}{2} r_C S$, while non-discrimination is for $r_C S$. Our case is therefore exactly in the middle of the interval.

5.2.5 Public banks

The government cannot employ an efficient contract scheme, thus the state-controlled banking sector is relatively inefficient. More to that, the state-owned bank cannot use market power over creditors, thus pays exactly $r_C S$ for credit at amount S.[5] This difference can be attributed also to pronounced propensity for bailouts, incorporated into ex ante costs.

In the elementary case, suppose that the government is using banks as a source of cheap credit for enterprises, regardless of private or public. The low level of additional output can be interpreted also as increased risk, hence higher share of qualified loans. (In the advanced case, the government can use bank profits as pure transfers or wage compensations.) Maximization of labor demand obviously involves minimization of interest rate for lenders, hence $r_C = r_G$. We denote this unified (socialist) interest rate r_S, $r_S := r_C = r_G$. The interest rate is constrained only by necessity to generate zero profits, $\pi = r_S K(r_S) - r_S S(r_S) = r_S[K(r_S) - \underline{\sigma} r_S] = 0$, which yields

$$r_S = K(r_S)\frac{1}{\underline{\sigma}}. \tag{5.5}$$

5.3 Ownership

We study the optimal decision-making of a government which is driven purely by interest in reelection. The sequence of activities is as follows. We begin in socialist status quo; the government owns brownfields and banks. Then, the government decides whether to privatize banks, brownfields, both, or none. In each institutional configuration, banks and enterprises maximize objective functions as explained above. On the basis of that, individuals get utilities and re-elect the government. We abstract from political rents.

Since the utilities are linear in ω_i, the utilities for any of the four institutional options are quasiconcave in ω. Therefore, it is the individual with median preference $\omega_i = 1/2$ whose preference is decisive for the institutional configuration; if he prefers one institutional option to another, his preference must be backed by a majority of voters. Throughout the chapter, we reasonably set that the median voter is job-seeking, in other words $\hat{\omega} < 1/2$ (otherwise, we would have a Monte-Carlo-like renters economy).

Anticipating preferences of a median voter, the reelection-seeking government in targets median voter's optimum, hence maximizes the utility of the median voter from the set of available institutional choices.

[5] This assumption could be relaxed without affecting the main result; the state-controlled bank could discriminate in price.

The crucial question is what happens with profits and proceedings from privatization. As to profits of private owners, we suppose that they don't reflect in the utility of the median voter. One interpretation is that they flow abroad; another is that the share of investors in the population is negligible to be worth incorporating into the utilities. Profits of the government operating brownfields and banks can be used for direct transfers, or wage subsidies, which is studied in section 4.

We recognize the following four institutional configurations:

Socialism (SO) Brownfields and banks are not privatized. Banks are inefficient ($\sigma = \underline{\sigma}$), and set such r_S that profits are zero. Brownfields are not restructuralized ($A_B = a$) and maximize labor demand up to the point of zero profits.

Industrial socialism (IS) Banks are private and efficient ($\sigma = \overline{\sigma}$), and maximize profits, setting (r_C, r_G). State-owned brownfields are not restructuralized ($A_B = a$) and the state therein maximizes labor demand.

Banking socialism (BS) Banks are not privatized, thus inefficient ($\sigma = \underline{\sigma}$), setting r_S to have zero profits. Brownfields are privatized, restructuralized ($A_B = 1$), and maximize profits.

Capitalism (CA) Both banks and brownfields are privatized and restructured, ($\sigma = \overline{\sigma}$ and $A_B = 1$), and both maximize profits.

We are particularly interested in cases when banking socialism dominates not only socialism and industrial socialism, but also capitalism.

5.3.1 Socialism

The costs of socialism are obvious, namely large inefficiencies in all sectors. The low ability of a socialist bank to allocate credit efficiently is reflected by $\underline{\sigma} < \overline{\sigma}$, and the low performance in state-owned enterprises is reflected by $a < 1$.

What are the benefits of socialism? In banking sector, the government doesn't collect profits, so it may set a lower interest rate than the private banks, unless relative inefficiency is too large. In brownfields, the government is targeting the total amount of labor, not the marginal productivity of labor like the private owner.

We already know that for the socialist bank, $K_B = 0$, and it funds only the private (greenfield) sector. Since the private sector contributes to the total labor demand by $L_G^{SO}(r_S)$, it sets r_S so as to maximize the green labor demand. From (5.2) and (5.5), we have:

$$r_S^{SO} = \frac{1}{\underline{\sigma}} K_G(r_S) = \left(\frac{\Delta w}{\underline{\sigma}}\right)^{\frac{1}{3}} \tag{5.6}$$

Accordingly, the labor demand is as follows:

$$L_D^{SO} = L_B^{SO} + L_G^{SO} = \left(\frac{aP}{w}\right)^2 + 2\Delta^{\frac{2}{3}}\left(\frac{\underline{\sigma}}{w}\right)^{\frac{1}{3}} \tag{5.7}$$

5.3.2 Industrial socialism

Like in socialism, the brown sector requires no bank funding, so brownfield labor demand is identical to the case of socialism. However, labor demand of greenfield sector may diff. The private banks make credit more expensive because of market-power exploitation, but are more efficient, which lowers the interest rates. What is the optimal interest rate of the private bank for lenders, r_G? By (5.4):

$$r_G^{IS} = \left(\frac{3\Delta w}{\overline{\sigma}}\right)^{\frac{1}{3}} \tag{5.8}$$

The greenfields set labor demand in accordance with (5.2) and (5.8):

$$L_D^{IS} = L_B^{SO} + L_G^{IS} = \left(\frac{aP}{w}\right)^2 + 2\Delta^{\frac{2}{3}}\left(\frac{\overline{\sigma}}{3w}\right)^{\frac{1}{3}} \tag{5.9}$$

Proposition 18: Median voter prefers industrial socialism to socialism, if and only if $3\underline{\sigma} < \overline{\sigma}$.

Proof In elementary model, only L_D is variable of institutional configurations. For median voter's utility, $\frac{dU_i}{dL_D} \geq 0$. By (5.2), $L_G^{IS} > L_G^{SO}$ if and only if $r_G^{IS} < r_S^{SO}$. As $L_B^{IS} = L_B^{SO}$, $L_D^{IS} > L_D^{SO}$ is equivalent to $L_G^{IS} > L_G^{SO}$. With (5.6) and (5.8), we easily rewrite $r_G^{IS} < r_S^{SO}$ into $3\underline{\sigma} < \overline{\sigma}$. Putting all equivalences together complements the proof. □

To lose credit channel but not privatize brownfields is thus suboptimal to socialist status quo, unless differences between banking sector performances are overwhelming. This drives our intuition on the importance of state control of banks when labor demand stimulation is politically desirable.

5.3.3 Banking socialism

In this case, the state controls the banks, but none of the productive sectors. Labor demand is maximized by provision of cheap credit.

We firstly consider possibility to lower interest rate to one sector at the expense of another sector. The pair (r_B, r_G) would have to satisfy $K_B(r_B) + K_G(r_G) = K_B(r_S) + K_G(r_S) = \underline{\sigma} r_S^2$. Imposing capital demands of private owners from (5.2), the constraint re-writes into

$$\frac{1}{r_B^2} + \frac{1}{r_G^2} = \frac{\underline{\sigma} r_S^2}{w\Delta}.$$

We maximize labor demand subject to $r_B = \arg\max L_G(r_G) + L_B(r_B)$ and the constraint above, which yields

$$\frac{1}{r_B^2} = \frac{\underline{\sigma} r_S^2}{2w\Delta}.$$

This implies $r_G = r_B$. Selective credit manipulation is not a feasible way for the government aiming to boost employment. Moreover, as the state-owned bank minimizes credit, we have $r_S = r_G = r_B$. From that, we derive the optimal (lowest feasible) interest:

$$r_S^{BS} = \left(\frac{32\Delta w}{\underline{\sigma}}\right)^{\frac{1}{3}} \tag{5.10}$$

Entering into labor demands of private firms (5.2) and using $L_B^{BS} = L_G^{BS}$:

$$L_D^{BS} = \left(\frac{2\Delta\underline{\sigma}}{4w}\right)^{\frac{1}{3}} \tag{5.11}$$

5.3.4 Capitalism

In a full-fledged market, banks set r_G by (5.4):

$$r_G^{CA} = \left(\frac{96\Delta w}{\overline{\sigma}}\right)^{\frac{1}{3}} \tag{5.12}$$

The enterprises set factor demands as in (5.2), given the interest, r_G^{CA}:

$$L_D^{CA} = \left(\frac{2\Delta\overline{\sigma}}{12w}\right)^{\frac{1}{3}} \tag{5.13}$$

Proposition 19: Median voter prefers capitalism to banking socialism, if and only if $3\underline{\sigma} < \overline{\sigma}$.

Proof Analogically to proof of Proposition 18, we use that $r_G^{CA} < r_G^{BS}$ if and only if $3\underline{\sigma} < \overline{\sigma}$. □

5.3.5 Pros and cons

Propositions 18 and 19 give, under all circumstances, a threshold condition for bank privatization. Privatization of banks is contingent only upon the relatively inefficiency of banks under private and public ownership; unless the difference is large, banks are better to rest under state control.

We look for all conditions under which banking socialism dominates other alternatives. By the threshold condition, we get that if banking socialism dominates capitalism, then socialism dominates industrial socialism. Therefore, we only rest to find the missing condition under which banking socialism dominates socialism.

Proposition 20: Median voter prefers banking socialism to all alternatives, if and only if

$$\underline{\sigma} > \max\left\{\frac{\overline{\sigma}}{3}, \frac{16^3 w a^6}{[2^{\frac{5}{3}} - 1]^3 P^2}\right\}.$$

Proof Part one is implication of Propositions 18 and 19. Analogically to proofs of previous propositions, we have that utility of the median voter is maximized for maximum L_D. Re-writing $L_D^{BS} > L_D^{SO}$ yields $\underline{\sigma} > 16^3 w a^6 [2^{\frac{5}{3}} - 1]^{-3} P^{-2}$. □

With Proposition 20, we can derive comparative properties leading to the dominance of banking socialism:

1. Narrow bank efficiency gap. The smaller is the efficiency gap of public banks ($\overline{\sigma}/\underline{\sigma}$), the less pronounced is incentive to privatize banks.

2. Development of financial market. The more efficient public banks ($\underline{\sigma}$ grows), the higher incentive to privatize brownfields. This may also be the case when total amount of savings increase, for instance due to external price liberalization.

3. Large technology gap. The higher is the technology gap between non-restructured and restructured brownfields ($1/a$), the more important it is to privatize brownfields.

4. Low wage rigidity. The lower wage rigidity (the clearer labor market), the better it is to privatize brownfields.

5. High product price. The higher price of the product P, the more beneficial is privatization of brownfields.

5.4 The use of profits

So far, we supposed that the government creates zero profits when owning brownfields and banks. This is optimal when profits have no other use and the only goal is labor-demand maximization. However, profits can be used in a variety of other ways, which contribute to the utility function. In the case of brownfields, a state-owned firm could create profits and redistribute them to all individuals. The second option is to withdraw profits from brownfields and use them as wage subsidies to greenfields and/or brownfields. These options modify optimums in socialist and industrial socialist cases.

5.4.1 Pure transfers

The government owning brownfields now decides to decrease production of brownfields from the scope where marginal labor productivity is very low, and use the profits for direct transfers. Suppose the profits Π_B are distributed by a uniform per capita transfer Π_B/n. The government aims to maximize utility subject to L_B and Π_B, where profits are generated by $\Pi_B = aPL_B^{1/2} - wL_B$.

We put the constraint into the utility function of the median voter ($\omega_i = \frac{1}{2}$), under assumption of positive unemployment ($L_D < L_S$), and the optimum yields the optimal (Π_B, L_B), where

$$L_B = \left(\frac{2w - 2 + 2^{\frac{1}{2}}}{2(2 - 2^{\frac{1}{2}})} \right)^2 L_B^{SO}. \tag{5.14}$$

Hence, the extent of transfers and labor stimulation depend only on the level of rigidity, where $\hat{w} = \frac{3}{2}(2 - \frac{1}{2})$ is the critical level:

$$\begin{aligned} w \leq \hat{w} &\Longrightarrow L_B \leq L_B^{SO} \Longrightarrow \Pi_B \geq 0 \\ w > \hat{w} &\Longrightarrow L_B > L_B^{SO} \Longrightarrow \Pi_B < 0 \end{aligned} \tag{5.15}$$

This reveals that the government optimizes the magnitude of production in two ways. For sufficiently small rigidity, it generates positive profits that are redistributed among citizens. In a more relevant case (given positive unemployment), it creates deficit, paid by citizens, that is used for stimulation of production. Of course, the latter option requires that costless and non-distortional taxes can be imposed. If the government cannot tax, it is bound by liquidity constraint, and $L_B = L_B^{SO}$. Then, transfers are used as an instrument only for cases of low nominal wage rigidity; otherwise, the government fully stimulates production of brownfields, eliminating all profits.

We have an indication that the relative importance of direct transfers from brownfields declines in wage, i.e. with higher nominal wage rigidity. We can establish this finding formally. Since we have an additively separable utility function, where marginal utility of additional transfer and additional labor demand are constant, we can directly compare utility increases given by increase in transfers and decrease in the labor demand in brownfields (respectively, increase in taxes and increase in labor demand). For any pair (Π_B, L_B), we have:

$$U(\Pi, L_B) - U(\Pi, 0) = U(0, L_B) - U(0, 0)$$
$$U(\Pi, L_B) - U(0, L_B) = U(\Pi, 0) - U(0, 0)$$

Thus, the relative importance of transfers to job-creation is given by

$$\frac{U(\Pi_B, L_B) - U(0, L_B)}{U(\Pi_B, L_B) - U(\Pi_B, 0)} = \frac{2(6 - 3\sqrt{2} - 2w)}{2\omega_i w + \sqrt{2} - 2}. \tag{5.16}$$

This reinforces intuition that each individual has different relative interest in maintaining transfers vs. creating jobs. The ratio clearly declines in ω_i, so those who value income highly welcome transfers relatively less than those who emphasize leisure time.

We can further maintain that an increase in wage (the more pronounced problem of nominal rigidity) leads to the higher role of job-creation and lower role of transfers. It is sufficient to examine how the relative importance of transfers declines in w for the median voter,

$$\frac{d\frac{U(\Pi_B, L_B) - U(0, L_B)}{U(\Pi_B, L_B) - U(\Pi_B, 0)}}{dw} = \frac{\sqrt{2} - 2}{(w + \sqrt{2} - 2)^2} < 0. \tag{5.17}$$

The median voter favors less transfers when the wage rate increases; a similar outcome could be derived for any i, where $l_i^* = \frac{1}{2}$. The only group which opposes that is the group of those who refuse to work ($l_i^* = 0$); for them, there is no trade-off between transfers and additional jobs, because $U(\Pi_B, L_B) - U(\Pi_B, 0) = 0$. In other words, when transfers are available, those out of the labor market are no more indifferent about the level of labor demand, but prefer the demand that maximizes profits of brownfields. Paradoxically, it is the group of non-workers who push the government to behave like a private owner. Nonetheless, in our simple median-voter specification, they are not in majority unless $\hat{\omega} \geq 1/2$ (renters' economy).

Can banking socialism be chosen even under the possibility of direct transfers from brownfield profits? First, since the transfer option is identical in socialism and industrial socialism, the two differ only in the level of employment in greenfield sector, which is described in condition $3\underline{\sigma} > \overline{\sigma}$. Therefore, this condition is unchanged. What surely changes is the other condition; since the case with positive transfers necessarily weakly dominates the case without transfers (which is a special case of zero transfers), we can conclude that $\underline{\sigma}$ has to pass a higher threshold to sustain victory of banking socialism. In particular, the modified condition writes as follows:

$$\underline{\sigma} > \frac{a^6 w(w-1)}{(2^{5/3}-1)^3 P^2} \frac{2w + 2^{1/2} - 2}{w + 2^{1/2} - 2} \frac{w + 2^{3/2} - 3}{2^{3/2} - 3} \tag{5.18}$$

5.4.2 Wage subsidies within brownfields

Excessive wage is the key source of mismatch between labor demand and labor supply. If the government can use a wage subsidy, it contributes to market clearing, even at an inefficiently high level employment. The labor supply can't adjust, but jobs that pay an effectively lower wage than w can be filled if the government provides a wage subsidy to the enterprises (e.g., in the form of social insurance deductions).

To pay a wage subsidy from state-owned brownfield profits is unfortunately nothing but a transfer from one pocket to another. Suppose that the government selects an optimal pair of transfer and labor in brownfields (Π_B, L_B). This pair must be feasible, that is it corresponds to the production function, $\Pi_B = aPL_B^{1/2} - wL_B$.

Suppose the enterprises increase production with subsidized work, and generate additional profits Π_a. The profits have to be re-paid by wage subsidies at amount Π_a/L_B, therefore the effective wage paid by brownfields is $w - \Pi_a/L_B$. The new allocations have to satisfy

$$aPL_B^{1/2} - \left(w - \frac{\Pi_a}{L_B}\right) L_B = \Pi_B + \Pi_a. \tag{5.19}$$

By multiplication, we of course only replicate the standard constraint, so the optimal pair (Π_B, L_B) is identical. In other words, the government has no incentive to change behavior, and will not use wage subsidies for brownfields that are withdrawn from their profits.

5.4.3 Wage subsidies to greenfields

Government may target the wage subsidies into the greenfield sector, for greenfields have better technology than non-privatized brownfields, thus a higher potential of job-creation. How would that work? For the moment, we neglect the simultaneous possibility of wage subsidies and pure transfers and assume profits to be used exclusively as wage subsidies to greenfields. The government thus determines the size of the profits extracted, Π_a, and commands brownfields to produce as much as to cover costs plus the profit extracted, and distributed the profits to greenfields in a form of a wage subsidy per unit of labor Π_a/L_G. Therefore, greenfields operate with lower labor costs, namely $w - \Pi_a/L_G$. Labor demand in brownfields and greenfields then satisfy:

$$P\underline{\kappa}L_B^{1/2} - wL_B - \Pi = 0$$
$$L_G = \frac{P^4}{32r_G}\left(w - \frac{\Pi}{L_G}\right)^{-3} \tag{5.20}$$

From maximization of $L_D = L_B + L_G$, we might get an (albeit extremely long) polynomial that would lead to the explicit solution. Here, a sufficient finding to derive is the relationship between L_B and L_G,

$$L_G^{1/3} = L_B^{1/2}\frac{8P^{8/9}}{3(6wL_B^{1/2} + aP)}. \tag{5.21}$$

We want to check if the wage subsidies to greenfields are positive, or negative. In the latter case, the government would tax greenfields and use the revenues for the production of brownfields (of course only if such a corporate tax is feasible). The sign of wage subsidy is derived from the sign of profits; positive profits entail a positive subsidy and vice versa. Denote the labor demands when the wage subsidy is zero (known from preceding section) as $L_B(0)$ and $L_G(0)$.

First, for strictly positive profits, we have $\Pi_B > 0$, which implies $L_B < L_B(0)$ and $L_G > L_G(0)$. On the other hand, for strictly negative profits, we have that $L_B > L_B(0)$ and $L_G < L_G(0)$. To sum up:

$$\Pi_B > 0 \Longrightarrow \frac{L_B(0)}{L_G(0)} > \frac{L_B}{L_G}$$
$$\Pi_B < 0 \Longrightarrow \frac{L_B(0)}{L_G(0)} < \frac{L_B}{L_G}$$

As we know the function $L_G(L_B)$ from (5.21), we can compute $L_G(L_B(0))$, and compare:

$$L_G(L_B(0)) > L_G(0) \Longrightarrow \frac{L_B(0)}{L_G(L_B(0))} < \frac{L_B(0)}{L_G(0)} \Longrightarrow \Pi > 0$$
$$L_G(L_B(0)) < L_G(0) \Longrightarrow \frac{L_B(0)}{L_G(L_B(0))} > \frac{L_B(0)}{L_G(0)} \Longrightarrow \Pi < 0 \quad (5.22)$$

We simply used the fact that relative share of employment between brownfields and greenfields has to be higher after a positive subsidy for greenfields, and lower after a negative subsidy (tax) for greenfields. In our example, we rewrite into conditions:

$$w\underline{\sigma}^{1/3} < \tfrac{2^{21}}{7^9} \Longrightarrow \Pi > 0$$
$$w\underline{\sigma}^{1/3} > \tfrac{2^{21}}{7^9} \Longrightarrow \Pi < 0 \quad (5.23)$$

The incentive to favor greenfields instead of brownfields thus falls in wage rate, and efficacy (extent) of financial system. We also found that cheap capital is a substitute of government intervention; here, higher $\underline{\sigma}$ makes government less interested in assistance to capital-using sector, that is to greenfields.

5.5 Conclusion

This chapter developed a model where the government seeking reelection by the retrospectively voting median voter in the optimum doesn't privatize banks, which results in status quo with lower bank efficiency. We identified two conditions which describe the willingness of the government not to privatize banks, but at the same time privatize state-owned enterprises. The two conditions also identify when any of the other three institutional configurations (socialism, industrial socialism, or capitalism) is optimal. The conditions can be used for comparative empirical studies on the relative benefits of delays in privatization.

In the simple model, the government couldn't generate profits. We extended that into a case when brownfields profits could be used as pure transfers, or wage subsidies for greenfields. In the case of pure transfers, we identified critical conditions for the non-negative level of transfers, and discussed political involvement of pure beneficiaries who are out of labor market. In the case of wage subsidies, we derived situations when subsidies are actually negative, namely represent additional taxes imposed upon greenfields.

The model can be enlarged by considering credit subsidies to greenfields paid from profits of brownfields. Another important option is to insert proceedings from the privatization sales into the cost-benefit analysis of the government.

References

[1] Ahrend, R. (2002). "Beyond Machiavellism: The Political Economy of Privatisation," DELTA Working Paper 2002-12.

[2] Ambrus-Lakatos, L. and U. Hege (1998). "Delegation and Delay in Bank Privatisation," WDI Working Paper 181, William Davidson Institute.

[3] Balla, K., Kertesi, G., Köllo, J. and A. Simonovits (2005). "Privatization, Unemployment and Subsidy for Low-skilled Labor," 2005 Meeting Papers 75, Society for Economic Dynamics.

[4] Biais, B. and E. Perotti (2002). "Machiavellian Privatization," *American Economic Review*, 92 (1), 240–258.

[5] Boehmer, E., Nash, R. C. and J. M. Netter (2005). "Bank Privatization in Developing and Developed Countries: Cross-sectional Evidence on the Impact of Economic and Political Factors," *Journal of Banking & Finance*, 29, 1981–2013.

[6] Börner, K. (2004). "The Political Economy of Privatization: Why Do Governments Want Reforms?" Munich Economics Discussion Paper 2004-1.

[7] Bortolotti, B., Fantini, M. and D. Siniscalco (2001). "Privatisation: Politics, Institutions, and Financial Markets," *Emerging Markets Review*, 2, 109–136.

[8] Clarke, G. R. G., Cull, R. and M. M. Shirley (2005). "Bank Privatization in Developing Countries: A Summary of Lessons and Findings," *Journal of Banking & Finance*, 29, 1905–1930.

[9] Desai, R. M. and A. Olofsgard (2006). "The Political Advantage of Soft Budget Constraints," *European Journal of Political Economy*, 22, 2, 370–387.

[10] Gordon, R. H. (2003). "Taxes and Privatization," in Cnoseen, S. & H.-W. Sinn, eds., *Public Finance and Public Policy in the New Century*, MIT Press, pp. 185–211.

[11] Jones, S. L., Megginson, W. L., Nash, R. C., and J. M. Netter (1999). "Share Issue Privatizations As Financial Means to Political and Economic Ends," *Journal of Financial Economics*, 53, 217–253.

[12] Kornai, J., Maskin, E. & G. Roland (2003). "Understanding the Soft-Budget Constraint," *Journal of Economic Literature*, 41, 1095–1136.

[13] Lian, P. and S. Wei (1998). "To Shock or Not to Shock? Economics and Political Economy of Large-scale Reforms," *Economics and Politics*, 10 (2), 161–183.

6. PUBLIC SECTOR EFFICIENCY IN THE NEW EU MEMBER STATES

6.1 Introduction

Countries with large public sectors are typically associated with high average costs. The standard explanation is that due to increasing marginal costs, total output requires higher average costs. This is nonetheless an incomplete story since cost functions may differ across countries. Part of the difference stems from the fact that services for the government are non-tradable, and the price for factors of production may not clear on the international market. Even if these effects are neglected by assuming average cost functions proportional with country income (reflecting, among others, correlation between wages in public sector and private sector, thus correlation between wages in public sector and country's income), we can still expect differences between average costs of different public sectors. These differences shall be attributed to different efficiency of public sectors, representing most likely organizational slack and institutional backwardness.

Effectiveness of public spending is a topic of growing interest (for the latest research, see Afonso and St. Aubyn 2004, Hjerppe *et al.* 2006), but little have been devoted to Central and Eastern European countries. An exception is Gupta *et al.* (2002), who analyzed whether public spending on education and health care matter for education attainment and health status in 50 developing and transition countries. They found strongest evidence for effectiveness of spending for education, and weaker evidence for health.

In this study, we focus on public sector efficiency in 10 new EU member states. We derive several indices of public sector efficiency, following methodology by Afonso *et al.* (2005). Using indicators of the World Bank (2005), we start with indices of public sector performance, and then compute efficiency of public finances both in the aggregate (composite) index and disaggregated into sector indices. Finally, as the chapter aims at Czech audience, we discuss the position of the Czech public sector among the new EU member states. We also point to shortcomings of the methodology.

6.2 Public sector performance (PSP)

6.2.1 Indicators

Public sector produces an enormous variety of outcomes-goods, social insurance, services and institutions. Some are final consumption goods, while some are inputs in production of private sector. The problem is that their importance can be hardly estimated if shadow prices of these outcomes are not available. To overcome the problem, Afonso et al. (2005) identify 7 important public sector *categories* and for each they select up to 4 *variables*. The categories include administrative/institutional outcomes, education, health, public infrastructure, economic performance, economic stability, and distribution. They use World Bank data (World Development Indicators 2001), OECD (Main Indicators, Social Expenditures, PISA Report 2000), European Commission (Ameco), World Economic Forum (The World Competitiveness Yearbook 2001), and IRIS. For each category, they construct a *sub-indicator*. The *total indicator* is then the unweighted average of sub-indicators.

We replicate their methodology as closely as possible to get results allowing for comparisons. We use the dataset of the World Development Indicators 2005 because it comprises most of variables mentioned in Afonso et al. (2005). For variables not represented in WDI, we find close substitutes, as reported in Table 6.1. The list of all variables and years (or periods) of observations is given in Table 6.2.

Tab. 6.1: Disjunct variables and sources in Afonso *et al.* (2005) and Gregor (2006)

Author	Variable	Source
Afonso et al.	Confidence in administration of justice/ Justice	WEF 1990, 2001
Gregor	Courts (% of managers surveyed lacking confidence in courts to uphold property rights)	WDI 2002
	Courts (% of managers surveyed ranking this as a major constraint)	WDI 2002
Afonso et al.	Regulatory environment/ Bureaucracy	WEF 1990, 2001
	Size of shadow economy	Schneider 2002
Gregor	n. a.	
Afonso et al.	Education achievement	OECD 2001
Gregor	School enrollment, tertiary	WDI 1990, 2002
Afonso et al.	Communications & transport quality	IRIS
Gregor	Roads, paved (% of total roads)	WDI 1990, 2002
	Roads (km) per population and square root of size in sq km	WDI 1990, 2002

6.3 Data

We want to compare starting conditions and current development. Therefore, we work with data from 1990 and 2002 (or, the closest year available). For 1990, the reason is that most of WDI variables are available as early in 1990; the older statistics from communist countries can be disregarded as biased or distorted (esp. in GDP growth, inflation). For new data, 2002 is the most recent year where the set of observations is complete.

For stocks, we use annual observations. For flows (growth, inflation), we consider years in period 1990-2004 when data is available for the majority of countries. Then, we split the time span into half, and get typically six- or seven-year averages. The first average is assigned into 1990 set, and the

	WDI	Period 1 (~1990)	Period 2 (~2002)	Index	Comment
ADMINISTRATION					Only one observation in 1990-2004
Corruption (% of managers surveyed ranking this as a major constraint)	#76	2002 (PL 2003)	2002 (PL 2003)	reversed	CY and MT inputted as average (1)
Courts (% of managers surveyed lacking confidence in courts to uphold property rights)	#77	2002 (PL 2003)	2002 (PL 2003)	reversed	CY and MT inputted as average (1)
Courts (% of managers surveyed ranking this as a major constraint)	#78	2002 (PL 2003)	2002 (PL 2003)	reversed	CY and MT inputted as average (1)
EDUCATION					SK 1990 inputted from CZ 1990
School enrollment, secondary (% gross)	#537	1990	2002		
School enrollment, tertiary (% gross)	#542	1990	2002		
HEALTH					
Mortality rate, infant (per 1,000 live births)	#390	1990	2003	reversed	
Life expectancy at birth, total (years)	#347	1990	2003		
INFRASTRUCTURE					
Roads, paved (% of total roads)	#521	1990 (MT 1994)	2002 (PL 2001)		
Roads (km) per population and square root of size in sq km	#522 * SQRT (#452) / #456				
* Roads, total network (km)	#522	1990 (MT 1994)	2002 (PL 2001)		
* Population density (people per sq km)	#452	1990-1994	2002		
* Population, total	#456	1990	2002		
DISTRIBUTION					
Income share held by lowest 40%	#286 + #287				Both quintiles come from the same year.
* Income share held by lowest 20%	#286	1996-2002	1996-2002		CY and MT inputted as average (1)
* Income share held by second 20%	#287	1996-2002	1996-2002		CY and MT inputted as average (1)
STABILITY					
GDP growth (annual %)	#173	s. d. of 1991-1997	s. d. of 1998-2004	reversed	s. d. = standard deviation
Inflation, GDP deflator (annual %)	#297	average of 1991-1997	average of 1998-2004	reversed	For CPI inflation, we have only 1994-2002.
PERFORMANCE					
GDP growth (annual %)	#173	Average of 1991-1997	average of 1998-2004		
Unemployment, total (% of total labor force)	#613	Average of 1991-1996	average of 1997-2002	reversed	LT 1991 inputted from LT 1992; SO 1991 and 92 inputted from SO 1993

Tab. 6.2: Construction of sub-indices: data, time span, and method

other to 2002 set. For growth variability, we compute standard deviations of growth over the respective periods. The only exception to stock vs. flow division rule is unemployment. Though a stock, we compute the average value of unemployment exactly as Afonso et al. (2005) did.

6.3.1 Indices

For each sub-indicator of performance in each period, we compute a *simple* and an *advanced* index. This is always an unweighted average of simple (advanced) indices calculated for each variable. Denote $E(X)$ the mean, $M(X)$ the minimum and $D(X)$ the standard deviation of sample X. Then, a simple index s_i and an advanced index a_i, attributed to an observation x_i, are defined as follows:

$$s_i = \frac{x_i}{E(X)} \qquad a_i = \frac{x_i - E(X)}{2D(X)}$$

Both indices are linear and satisfy $s_i = a_i = 1$ for $x_i = E(X)$. They differ such that $s_i = 0$ for $x_i = 0$, while $a_i = 0$ for $xi = E(X) - 2D(X)$. Note that the indices assume linearity in variables, so we could say that "exchange ratio" across variables is almost fixed. (Hypothetically, if a country increased variable A smaller by a small, it would always pay the same unit price in terms of a decrease in variable B to keep the sub-indicator constant.)

The simple index assumes zero fixed costs, while the advanced index allows for positive/negative fixed costs of public output production. The negative fixed costs can be interpreted as a production function achieving a strictly positive output for zero costs, which is usually the case when external factors contribute to the level of provision, such as in education.

The advanced index is justified whenever we suspect that, regardless of zero public effort, we get a strictly positive value of x_i. On the other hand, since the relative importance of a sub-indicator is determined only by variation in the sample, the relative position of a country is reflected only by this variation. This is not necessarily good, since for tiny differences in absolute terms, the index becomes very sensitive (for more, see below).

The advanced index solves a problem emerging when zero public effort causes a negative value, like for economic growth. To avoid this problem for simple index, we decided to re-compute a simple index on the explicit assumption that $s_i = 0$ for $x_i = M(X)$, which we do whenever negative values are present in the sample:

$$s_i = \frac{x_i - M(X)}{E(X) - M(X)}$$

When a high value of observation is defined as a diminished performance, we use reversed values, defined as $\hat{s}_i = 2 - s_i$ and $\hat{a}_i = 2 - a_i$. Afonso et al. (2005), in contrast, recalculated sample values by setting $1/x_i$. By hyperbolic transformation, they however violated the assumption of linearity in data. The aggregated index hence loses part of rationale, since it arbitrarily mixes linearity and non-linearity.

When a sub-indicator is composed from several variables, we firstly compute indices from each variable and then construct the sub-indicator as the average of indicators. This aggregation is also maintained when constructing a total indicator.

Differences between the indices

Figure 6.1 demonstrates differences between the indices for life expectancy in 2003. The simple index shows very low variation, while the adjusted index has large variation. For comparison, growth volatility in 1998-2004 features almost identical simple and advanced indices.

6.3.2 Evolution of public sector performance

The 1990 and 2002 indices of public sector performance (PSP 1990 and PSP 2002) indicate whether there has been some relative progress in their peer groups, comparing starting, post-communist conditions, to the situation in the end of millennium. The initial performances are reported in Tables 6.3 and 6.4, and the more recent performance in Tables 6.5 and 6.6. Recall that performance measures still don't indicate efficiency—they disregard costs necessary to achieve the output.

Starting conditions (PSP 1990) Simple index reveals clearly superior performance in Malta and Cyprus to former Communist states in 1990. However, Czech Republic and Slovenia were not that behind at that time. The laggards are former parts of the Soviet Union, namely Latvia and Lithuania. This is however mainly due to strong economic instability in early 90s in the successors of the Soviet Union. Note that the difference between maximum and minimum index of stability is 2.01; in health, it amounts only to 0.34.

By advanced index, we again get the quarter of Malta, Slovenia, Czech Republic and Cyprus on the top. The lowest performance is found in Latvia and Poland. Stability doesn't play such a dominant role because of normalization; the difference between maximum and minimum is 1.33, while for health it is 1.2.

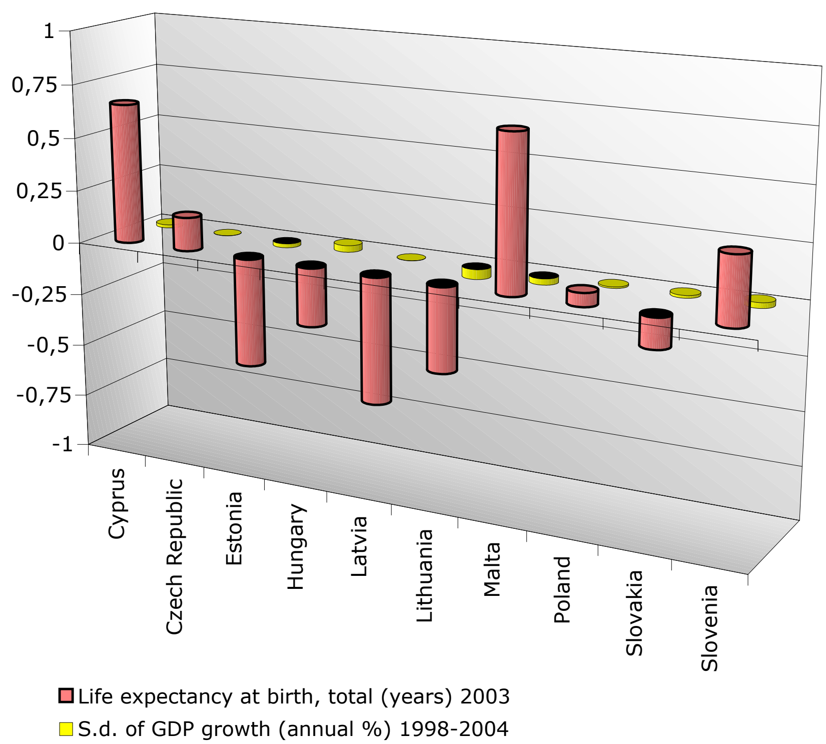

Fig. 6.1: Differences between the simple and advanced indices, $s_i - a_i$

Current performance (PSP 2002) and comparison With simple index, we get that Slovenia maintained the top position, unlike Mediterranean countries. Surprisingly, Latvia got to the very top, which signals that Latvia is perhaps the country with the biggest progress. Poland remains among the worst, and is joined by Slovakia. Here we get the first sign of inferior Slovak performance in 90s, possibly caused by the regime of authoritative Vladimir Meciar.

Slovenia features the best also by advanced index, and the Czech Republic is the second best. Again, Poland and Slovakia is the couple of countries with the lowest performance.

What about relative improvement across the periods, measured in basis points of simple/advanced index? The most dramatic improvement is found in Latvia (+20/+34), in consistence with the previous findings. Lithuania (+10/+26), Estonia (+2/+17), Hungary (0/12) and Slovenia (+5/+1) also show clear improvements. Poland is equivocal (-9/+5). A relative downfall in found in Mediterranean countries, Malta (-27/-35) and Cyprus (-5/-13), plus the successors of Czechoslovakia, i.e. the Czech Republic (-8/-12) and Slovakia (-12/-8).

Tab. 6.3: Public sector performance (PSP 1990), simple index

	ADM	EDU	HEA	INF	DIS	STA	PER	PSP 90[1]
Cyprus	1	0.72	1.13	1.01	1	1.74	1.6	1.17
Czech Rep.	1.06	0.93	1.05	1.04	1.14	1.48	1.27	1.14
Estonia	1.54	1.21	1	0.66	0.84	0.07	0.88	0.89
Hungary	1.38	0.8	0.88	0.91	1.08	1.44	0.81	1.04
Latvia	1.29	1.14	0.91	0.88	0.91	-0.19	0.43	0.77
Lithuania	0.88	1.35	1.01	1.10	0.95	-0.37	0.39	0.76
Malta	1	0.79	1.08	1.63	1	1.91	1.76	1.31
Poland	0.28	1	0.84	0.98	0.9	1.42	1.03	0.92
Slovakia	0.27	0.93	0.92	0.9	1.09	1.3	0.77	0.88
Slovenia	1.31	1.12	1.18	0.89	1.08	1.2	1.07	1.12

[1] ADM = Administration, EDU = Education, HEA = Health, INF = Infrastructure, DIS = Distribution, STA = Stability, PER = Performance

Tab. 6.4: Public sector performance (PSP 1990), advanced index

	ADM	EDU	HEA	INF	DIS	STA	PER	PSP 90
Cyprus	1	0.25	1.7	0.99	1	1.44	1.57	1.14
Czech Rep.	1.04	0.99	1.1	1.09	1.72	1.25	1.3	1.21
Estonia	1.7	1.58	0.8	0.61	0.19	0.48	0.94	0.9
Hungary	1.4	0.51	0.5	0.88	1.42	1.24	0.8	0.96
Latvia	1.18	1.3	0.58	0.8	0.53	0.26	0.54	0.74
Lithuania	0.7	1.62	0.97	1.13	0.74	0.28	0.5	0.85
Malta	1	0.6	1.5	1.71	1	1.59	1.66	1.29
Poland	0.41	0.88	0.53	0.97	0.52	1.25	0.9	0.78
Slovakia	0.29	0.99	0.74	0.93	1.46	1.09	0.74	0.89
Slovenia	1.27	1.27	1.58	0.88	1.4	1.13	1.04	1.23

Tab. 6.5: Public sector performance (PSP 2002), simple index

	ADM	EDU	HEA	INF	DIS	STA	PER	PSP 02
Cyprus	1	0.8	1.21	0.87	1	1.27	1.09	1.04
Czech Rep.	1.06	0.83	1.2	1.14	1.14	1.06	0.7	1.02
Estonia	1.54	1.12	0.85	0.94	0.84	0.8	1.3	1.06
Hungary	1.38	1.02	0.88	0.84	1.08	0.87	1.19	1.04
Latvia	1.29	1.17	0.69	1.15	0.91	1.06	1.45	1.1
Lithuania	0.88	1.2	0.86	1.14	0.95	0.93	1.17	1.02
Malta	1	0.77	1.14	1.25	1	0.86	0.71	0.96
Poland	0.28	1.1	1.03	0.81	0.9	1.05	0.65	0.83
Slovakia	0.27	0.78	0.94	0.85	1.09	1.04	0.64	0.8
Slovenia	1.31	1.2	1.2	1	1.08	1.05	1.09	1.13

Tab. 6.6: Public sector performance (PSP 2002), advanced

	ADM	EDU	HEA	INF	DIS	STA	PER	PSP 02
Cyprus	1	0.67	1.63	0.8	1.0	1.3	1.2	1.09
Czech Rep.	1.04	0.65	1.38	1.2	1.72	1.07	0.86	1.13
Estonia	1.7	1.06	0.53	0.96	0.19	0.79	1.18	0.92
Hungary	1.4	1.26	0.69	0.77	1.42	0.81	1.2	1.08
Latvia	1.18	1.1	0.25	1.23	0.53	1.06	1.23	0.94
Lithuania	0.7	1.4	0.59	1.22	0.74	0.97	1.02	0.95
Malta	1	0.48	1.55	1.4	1	0.87	0.9	1.03
Poland	0.41	1.33	1.08	0.69	0.52	1.06	0.66	0.82
Slovakia	0.29	0.39	0.85	0.75	1.46	1.03	0.62	0.77
Slovenia	1.27	1.66	1.45	0.97	1.40	1.03	1.14	1.28

6.4 Public sector efficiency (PSE)

We make two distinctively different attempts to measure public sector efficiency. In the first, non-parametric approach (Free Disposable Hull analysis, FDH), we estimate a production possibility frontier on the basis of the best total performance in the sample. This is done by grouping countries by total public expenditures. In the other approach, we simply measure average costs in each sub-indicator by assigning relevant expenditures to each sub-indicator, and then weight the sub-indicators of efficiency by several weights. The first approach allows us to distinguish between input and output efficiency and needs not the assumption of constant marginal costs. On the other, if we applied it for particular aspects of efficiency like the second one, we would magnify the problems involved in definition of relevant expenditures and the small size of the sample.

6.4.1 Total input and output efficiency

The Free Disposable Hull analysis is a simple non-parametric method constructing the hypothetical production possibility frontier, proposed by Deprins, Simar and Tulkens (1984). The frontier, located in space of PSP and total costs, is defined simply: it incorporates all countries whose PSP and costs are not dominated, i.e. there is no country which would achieve the same or higher PSP with the same or lower costs. To get the frontier, we take a country with the smallest costs; denote it A for the moment. The

PSP of A must be on the frontier, since there is no other country that would produce higher PSP with the same or lower costs (there is no country with lower costs). Then we look for another country on the frontier. In specific, we look for a country with the second lowest costs, to be denoted B. If PSP of B is identical to or lower than the PSP of A, we get that efficiency of B is dominated by efficiency of A; country A achieved higher PSP with lower costs. But, if PSP of B is higher, B is located on the frontier. Then, we make additional step in the process, as we again search another country with higher costs that would have a higher PSP, now PSP of B.

By this method, we can distinguish between input and output efficiency. The input efficiency indicates how much less input the country could use to achieve the same level of output. The output efficiency gives how much more output could a country make using the same inputs. These indices take necessarily a value between zero and one, and are not identical because average costs are not independent in amount of production.

Here, we work with PSP 2002 and the cost measure is the total public expenditures (WDI 123), the lagged value of 1997 (or the closest year available). We chose the lag as the beginning year of the second period for flow variables; moreover, 1997 data were the most complete.

FDH is not suitable for samples of small sizes. Most of observations may happen to be found undominated because of large dispersion of costs. In our case, we have 4 out of 10 countries on the production possibility frontier for both simple and advanced index, which is acceptable.

Tab. 6.7: Input/output efficiency (value/rank)

	Simple index				Advanced index			
Efficiency	Input		Output		Input		Output	
Cyprus	0.82	5.	0.94	5.	1	1.-4.	1	1.-4.
Czech Rep.	0.73	8.	0.93	6.	1	1.-4.	1	1.-4.
Estonia	1	1.-4.	1	1.-4.	0.97	5.	0.97	6.
Hungary	0.69	10.	0.92	7.	0.84	8.	0.85	8.
Latvia	1	1.-4.	1	1.-4.	0.96	6.	0.99	5.
Lithuania	1	1.-4.	1	1.-4.	1	1.-4.	1	1.-4.
Malta	0.73	9.	0.87	8.	0.92	7.	0.95	7.
Poland	0.76	6.	0.75	9.	0.76	9.	0.75	9.
Slovakia	0.73	7.	0.73	10.	0.73	10.	0.71	10.
Slovenia	1	1.-4.	1	1.-4.	1	1.-4.	1	1.-4.

The Table 6.7 identifies two clear leaders of rankings, Slovenia and Lithuania. Cyprus is on the frontier only in the advanced index, and Estonia and Latvia only in the simple index. The case of the Czech Republic is unclear - although situated on the frontier in the advanced index, it is nonetheless 8th/6th in the input/output efficiency scores for simple index. Hungary, Poland, and Malta are consistently below average, and Slovakia clearly occupies the worst position.

This indicates several stories. Slovenia is a shining example—it has consistently high performance, which has been slightly improved and, moreover, the sector is very efficient. Lithuania made a huge improvement in performance, largely due to an increase in efficiency. Also the remaining two Baltic countries feature very high efficiency. Of Mediterranean countries, Cyprus fell down the ranking, but still keeps a relatively efficient public sector, which is no more the case of Malta. Of the four Central European countries, Czech Republic seems to be the best, and Slovakia consistently the worst.

6.4.2 Total efficiency based on sector-specific performances

Another approach is to decompose total expenditures into relevant public expenditures, and assign each part to each sub-indicator. Then we can define efficiency in each sub-indicator as the average costs in the category (i.e. benefit-cost ratios). How is the comparison of average costs justified? If objective costs (costs with hypothetical full efficiency) are proportional to country's income and marginal costs are nearly constant, then average costs can be compared across countries with different incomes and different size of public sectors. Differences in average costs adjusted for country's income are then indicative of differences in sectoral efficiency.

Like in the previous example, we examine 2002 performance variables and expenditure data from 1997. For definition of relevant public expenditure per each category, see Table 6.8.

Tab. 6.8: Relevant expenditures for each sector/category

	Relevant expenditure	No.	Year
ADM	Expense (% of GDP)	123	1997
EDU	Public spending on education, total (% of GDP)	493	1998; SK99; LT01; SO02
HEA	Health expenditure, public (% GDP)	242	1998
INF	Gross fixed capital formation (% of GDP)	219	1997
DIS	Social contributions (% of GDP)	557	
	GDP (current LCU)	170	1991–1997
	Social contributions (current LCU)	557	1997; CZ, HU, LT 00; PL01; SK03
STA	General government final consumption expenditure (% of GDP)	189	1997
PER	General government final consumption expenditure (% of GDP)	189	1997

To get a more interesting picture, we aggregate the sectoral efficiency indices into total efficiency indices by five ways. The first one is a Baseline indicator with equal weights over all 7 categories. Then, we impose special preference either for administration, distribution, stability, or economic performance, getting Administration, Distribution, Stability, or Economic performance total efficiency scores. Figure 6.2 presents the weights.

More formally, suppose Y_j is the sample of expenditures in category j, and $E(Y_j)$ is mean. For expenditure y_{ij} of country i, when s_{ij} and a_{ij} are her sectoral indices, we get the following sectoral efficiency indices:

$$e_{ij}^s = \frac{s_{ij} y_{ij}}{E(Y_j)} \qquad e_{ij}^a = \frac{a_{ij} y_{ij}}{E(Y_j)}$$

The total efficiency indices of country i are derived as $t_i^s = \mathbf{w}\mathbf{e}_\mathbf{i}^\mathbf{s}$ and $t_i^a = we_i^a$, where $\mathbf{w}$ is a vector of weights over seven sectoral efficiency indices, $\mathbf{w} = (w_1, , w_7)$. Tables 6.9 and 6.10 show the resulting scores as well as ranks within the group.

Slovenia happens to be no more a distinct show-case. The role is now taken by Cyprus and, quite surprisingly, by Hungary. Slovenia and Poland remain to be the robust laggards. Comparing to FDH analysis, there are differences in the upper and middle part of ranking. They stem from the

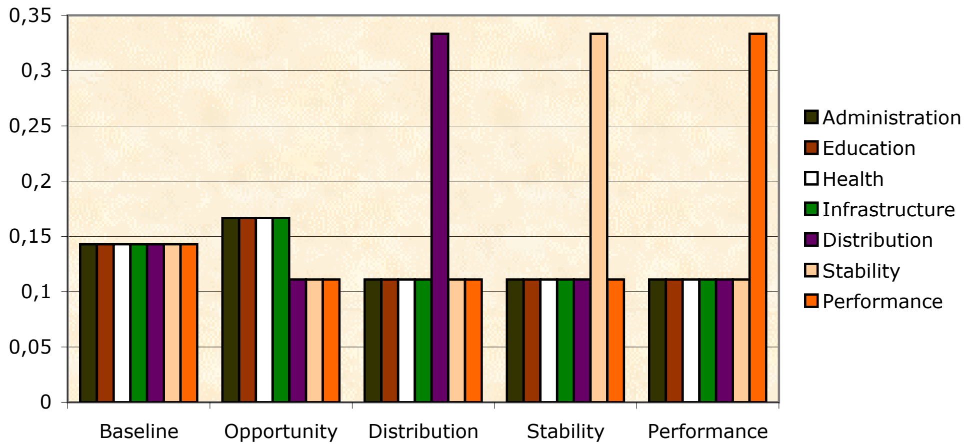

Fig. 6.2: Weights assigned to each total efficiency indicator

different assumptions; the FDH analysis rejects the assumption of constant marginal costs, since the production possibility frontier is not linear. With this finding, one has several potential solutions. Either we enlarge the sample to improve comparative efficiency measures, or we redefine relevant public expenditures. Another solution is to use FDH for total efficiency scores, while simple average costs approach can be applied for sectoral sub-indices.

Tab. 6.9: Efficiency for 5 weights (value/rank), simple index

	Baseline		Admin.		Distribution		Stability		Perform.	
Cyprus	1.43	1.	1.41	1.	1.61	1.	1.41	1.	1.37	2.
Czech Rep.	0.91	8.	0.93	8.	0.9	8.	0.92	8.	0.85	8.
Estonia	1.03	5.	1.05	4.	1	7.	0.96	7.	1.06	5.
Hungary	1.26	2.	1.2	3.	1.18	2.	1.34	2.	1.47	1.
Latvia	1.2	3.	1.22	2.	1.16	3.	1.16	3.	1.24	3.
Lithuania	1.03	6.	1.04	6.	1.04	5.	0.98	5.	1.03	6.
Malta	1.01	7.	1.01	7.	1.16	4.	0.97	6.	0.94	7.
Poland	0.88	9.	0.88	9.	0.86	9.	0.92	9.	0.83	9.
Slovakia	0.72	10.	0.71	10.	0.75	10.	0.77	10.	0.69	10.
Slovenia	1.05	4.	1.05	5.	1.01	6.	1.05	4.	1.06	4.

Tab. 6.10: Efficiency for 5 weights (value/rank), advanced

	Baseline		Admin.		Distribution		Stability		Perform.	
Cyprus	1.54	1.	1.53	1.	1.69	1.	1.5	1.	1.48	2.
Czech Rep.	0.99	6.	0.98	6.	1.05	5.	0.98	6.	0.94	6.
Estonia	0.9	8.	0.94	8.	0.75	10.	0.86	9.	0.93	8.
Hungary	1.29	2.	1.23	2.	1.27	2.	1.35	2.	1.51	1.
Latvia	1.03	5.	1.04	5.	0.93	6.	1.02	4.	1.06	4.
Lithuania	0.95	7.	0.96	7.	0.92	7.	0.92	7.	0.93	7.
Malta	1.07	4.	1.06	4.	1.2	3.	1.02	5.	1.02	5.
Poland	0.87	9.	0.89	9.	0.78	8.	0.92	8.	0.83	9.
Slovakia	0.67	10.	0.63	10.	0.77	9.	0.73	10.	0.64	10.
Slovenia	1.17	3.	1.17	3.	1.16	4.	1.14	3.	1.17	3.

6.5 A puzzle of the Czech Republic

The Czech Republic in 1990 scored twofold 3rd place in performance rankings. In other words, the Czech part of Czechoslovakia enjoyed the most stable economic environment in the early 90s. The Czech Republic in 2002 shows a rather different, unclear picture (see Table 6.11). By advanced index, it remains affront, overall on the 2nd spot. With simple index, the sectoral ranks paradoxically improve, but the overall score fells down to 6th. This reminds of sensitivity to computation of indices, demonstrated by Figure 6.1.

Tab. 6.11: Ranks of PSP sub-indicators (CZE, 2002)

Index	Simple	Advanced
Administration	5.	5.
Education	7.	8.
Health	2.	4.
Infrastructure	4.	4.
Distribution	1.	1.
Stability	2.	2.
Performance	8.	8.
PSP	**6**.	**2**.

The differences between simple and advanced index are maintained also when considering efficiency by FDH analysis. For simple index, the Czech Republic is only 8th and 6th, while for advanced index, it is situated on the production possibility frontier, which defines the maximum efficiency from the sample. To investigate effect of segregation of indices, examine further ranks of total efficiency for all five weights of sectoral indices in Table 6.12.

Tab. 6.12: Total efficiency ranks of the Czech Republic, 2002

Index	Simple	Advanced
Baseline	8.	6.
Administration	8.	6.
Distribution	8.	5.
Stability	8.	6.
Performance	8.	6.

These results indicate that the Czech Republic may have a slightly above-average public sector performance, but the performance is achieved with high average cost, i.e. low benefit-cost ratio. However, we get this only on the assumption of constant average costs.

To sum up, the Czech Republic in 2002 scores well using the advanced (normalizing) indexing of performance, and efficiency is relatively high using the non-parametric FDH analysis. The outcome is unfortunately not robust with the simple assumptions of constant marginal costs and zero fixed costs. And, of course, we discuss only static efficiency caused by 1997 expenditures, not efficiency of current expenditures, long-term public investments, not to speak about the sustainability of public finances.

6.6 Conclusion

We have analyzed the performance and efficiency of the public sector in the new EU member states. We constructed a composite indicator and seven sub-indicators of performance, and measured input and output efficiency.

The highest public sector performance in observed in Slovenia, and robustly superior performance in Cyprus, Czech Republic and Hungary. Contrariwise, Poland and Slovakia are at the bottom of the ranking. Considering efficiency, Free Disposable Hull method reveals that Slovenia keeps the best position along with Lithuania, while Slovakia is again the worst. Using aggregation of sector-specific efficiency measures, Cyprus and Hungary move up, while Slovakia and Poland turn to be the worst. The cases of Cyprus and Slovenia show that the high performance can be achieved even with the relatively low average costs. Slovakia and Poland, on the other side, show that even a large public sector may produce a low public sector output.

We took the Czech Republic as an example of a country whose position was significantly affected by the construction of index. Using advanced methods, the Czech Republic scores well both in performance and efficiency, while by simpler methods, the output and efficiency are inferior. This echoes methodological literature on weaknesses of composite indices (Jacobs *et al.* 2006). We also pointed to the fact that one must distinguish between static efficiency achieved in 2002 because of expenditures in 1997 or 1998, efficiency of current public spending, and dynamic efficiency of current public finances.

References

[1] Afonso, A., Schuknecht, L., & V. Tanzi (2005). "Public sector efficiency: An international comparison", *Public Choice*, 123, 321-347.

[2] Afonso, A., & M. St.Aubyn (2004). "Non-parametric Approaches to Education and Health Expenditure Efficiency in OECD countries", ISEG/UTL, Department of Economics, Working Paper 1/2004.

[3] Deprins, D., Simar, L., & H. Tulkens (1984). "Measuring labor-efficiency in post office". In M. Marchand, P. Pestieau, & H. Tulkens (eds.), *The Performance of Public Enterprises: Concepts and Measurements*. Amsterdam, North-Holland.

[4] Gupta, S., Verhoeven, M., & E. R. Tiongson (2002). "The effectiveness of government spending on education and health care in developing and transition economies", *European Journal of Political Economy*, 18, 717–737.

[5] Hjerppe, R., Kiander, J., & M. Viren (2006). "Are government expenditures productive? Measuring the effect on private sector production". VATT, Government Institute for Economic Research, Helsinki.

[6] Jacobs, R., Goddard, M., & P. C. Smith (2006). "Public Services: Are Composite Measures a Robust Reflection of Performance in the Public Sector", CHE Research Paper 16, Centre for Health Economics, University of York.

[7] World Bank (2005). World Development Indicators 2005.