ESSAYS ON

THE DECENTRALISATION OF DEMOCRATIC INSTITUTIONS

Antonios Haniotis

Submitted for the Degree of Ph.D.

London School of Economics and Political Science

University of London

1995

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

This is a thesis on the political economy of decentralisation. Part one analyses the effects of fiscal and electoral decentralisation on the accountability of public officials. Chapter 2 introduces a model of repeated elections to public office. The interaction between voters and officials is seen as a principal-agent relationship. Within this framework, two potential justifications for political decentralisation are provided. It is shown that voters may strictly prefer to limit the funds they allocate to a public office by setting up more than one office (fiscal decentralisation). A smaller budget implies a stronger incentive for the officeholder to try to be re-elected. The second argument for political decentralisation is that splitting up the electorate into homogeneous jurisdictions (electoral decentralisation) implies that the voters will tend to have more control over the use of public funds by the officials.

Chapter 3 introduces asymmetric information into the model. Officeholders are assumed to have better knowledge of the technological conditions under which public funds are provided. Asymmetric information gives rise to an information rent for the officials. Total information rents can be limited by having more than one bureau and the voters using relative performance evaluation for their re-election decisions (=electoral yardstick competition). It is also shown that fiscal decentralisation can be beneficial to the voters, if they are better informed about the local rather than the federal conditions of public goods provision.

Part two studies the political economy of the transition from a collectivist economy to a market economy. Chapter 4 develops a general equilibrium model of income distribution and risk-taking. The formation of entrepreneurial and working classes as well as the allocation of talent are derived endogenously. In a market economy, no insurance can be provided due to the non-observability of output (and employment) in private firms. The more able and the less risk averse individuals tend to become entrepreneurs. Inefficiencies arise due to insufficient risk-taking. A different (second-best) balance is reached in a collectivist economy where output is shared equally. In this case, non-observability implies that workers will not reveal their talent and managers will not reveal the productivity of their plant. It is shown that either of the two systems may overall dominate the other in terms of the majority criterion or even the Pareto criterion.

Different forms of transition from collectivist to market-based production are considered. A total (or "big bang") reform is viewed as the collectivist sector being closed down completely and all individuals having to move to the competitive sector. Gradual reform is viewed as the free migration of individuals between the collectivist sector and the competitive sector. This leads, in general, to a dual economy. A majority of individuals may want to leave the traditional sector to try their luck in the modern sector and, therefore, vote for reform. Despite this, under both gradual and total reforms, a majority of individuals may *ex post* regret the changes and vote for a return to the collectivist system. It is also shown that the majority choice between total reform, gradual reform and no reform is always transitive. Moreover, gradual reform is never the most preferred option for the majority. These results may be useful for interpreting some of the developments in the economies in transition.

TABLE OF CONTENTS

NOTATION

Part One:

δ time discount factor.

g_t output in period t.

a_t fraction of funds used for the production of public services in period t.

θ_t realisation of the technological shock in period t.

s public funds.

w private benefits of office.

$\hat{g}_{it}$ cutoff level of output for re-election set by voter i in period t.

v_t value to the incumbent of being in office in period t.

N number of voters or jurisdictions.

M number of votes required for re-election.

$\mathcal{E}$ set of voters.

β_{nk} benefit spillover from region k to region n.

G_n cutoff level of output for region n staying in the federation.

γ measure of technological inferiority of decentralisation.

$\hat{s}$ critical level of funds, beyond which the bureau's accountability is diminishing under symmetric information.

$\hat{s}^{\alpha}$ critical level of funds, beyond which the bureau's accountability is diminishing under asymmetric information.

π_t probability of the incumbent being re-elected, in equilibrium.

σ mean-preserving spread on θ_t.

Part Two:

v certainty equivalent.

y income.

π_H probability of success for entrepreneurs of high ability.

π_L probability of success for entrepreneurs of low ability.

R index of risk aversion.

$P(R)$ cumulative distribution of risk attitudes in the population.

g firm output.

λ variable labour input.

γ measure of decreasing returns to scale.

α proportion of talented individuals in the population.

ε proportion of managers in the population.

G aggregate output.

C collectivist economy (no reform).

CE competitive economy (total reform).

DE dual economy (partial reform).

Δ set of workers in the competitive sector.

Γ set of entrepreneurs in the competitive sector.

Σ set of workers in the collectivist sector.

σ proportion of collectivist sector workers in the population.

r_{CE} entrepreneurial income, in case of success.

ACKNOWLEDGEMENTS

This thesis would not exist if I had not been lucky enough to be one of Martin Hellwig's students at the University of Basle. He made modern economic theory accessible to us and showed us that there are still open questions in economics. He always manages to turn criticism into constructive thinking and disappointment into new ideas. I am extremely grateful for all his time and encouragement.

This thesis was written under the auspices of the *European Doctoral Programme in Quantitative Economics*. I have benefited from its stimulating environment in Bonn and London. I would like to thank its organisers and in particular Urs Schweizer at the University of Bonn who dedicates much of his valuable time and effort to the Programme.

At various stages of the writing of this theses, I benefited from discussions with and comments from numerous teachers, colleagues and friends such as Philippe Aghion, Patrick Bolton, Tilman Börgers, Gilles Chelma, Kathy Hayes, Stephen Howes, Alison Hole, Ashok Kotwal, Mikko Leppamaki, Georg Nöldeke, Thomas Piketty, Sven Rady, Klaus Schmidt, Reinhard Selten, Nick Stern, Max Steuer, Chenggang Xu, and in particular Gilles Duranton and Denis Gromb. All remaining limitations are, of course, my responsibility.

Financial support from the *Swiss National Science Foundation*, the *Max Geldner-Stiftung* and the *Nikolaus und Bertha Burckhardt-Bürgin-Stiftung* is gratefully acknowledged.

Chapter 1

INTRODUCTION

1.1 The traditional theory of political decentralisation

Political science and economics are two closely related social sciences. There exists a long standing history of thought on political economy and on economic policy. Yet, it seems puzzling that the issue of decentralisation has so strongly dominated economic thinking while, until recently, it has received the attention of very few political scientists. A related fact is that there have not been many fruitful attempts by economists to understand bureaucracy. Until recently, neither economists nor political scientists have directed sufficient efforts towards understanding competitive forces and incentives within political and bureaucratic institutions.[1]

This thesis deals with the political economy of decentralisation. The main underlying theme concerns the costs and benefits of decentralisation in democratic institutions. The arguments I am developing below apply to democratic nation states, but most of them are

1. A notable exception is, of course, the famous Downs-Hotelling model of political competition.

equally valid in the context of other democratic institutions such as corporate firms, cooperatives, sports clubs etc.

The motivation for my analysis is based on the following well-known puzzle[2]: *Take any organisation where formal authority is decentralised. Such an organisation will, in general, be inefficient because of different types of external effects and coordination problems among its units. These inefficiencies can be overcome through an agreement signed by all units inducing them to behave efficiently. Alternatively, formal authority can be concentrated on one benevolent centre which directly implements an efficient outcome. This implies that, as long as complete contracts can be signed and enforced, an efficient arrangement will be found and the distinction between centralisation (where authority is concentrated on a central decision-maker) and decentralisation (where authority is concentrated on a comprehensive contract and the court which enforces it) is purely semantic.*

In a seemingly very different context, a large literature on federalism has been developed. This literature, which can be traced back at least to Tocqueville (1835) and the Federalist Papers (Madison, Hamilton and Jay, 1787), has tried to prescribe how the functions of the state should be allocated between different levels of government. More recent standard references are Buchanan (1950), Musgrave (1959, 1961, 1971), Oates (1972), Riker (1964) and King (1984).

Several *advantages of centralisation* (=allocating functions to the federal government) have been identified in this literature. They range from increasing returns to scale in the production of public goods and services, a need for the coordination of policies and the internalisation of inter-jurisdictional externalities, to the view that the costs of redistribution would be lower under centralisation. More arguments for the superiority of centralisation have been developed more recently.[3] It has been argued that central government will in general strictly dominate decentralisation since it does not have to respect local autonomy: As Klibanoff and Morduch (1993) show, efficient policies in a federation may

2. See Coase (1939) and Williamson (1975). A more sophisticated version of this argument is known as the "revelation principle". See Dasgupta, Hammond and Maskin (1979).

3. For an excellent optimal taxation approach identifying several sources of inter-jurisdictional externalities and showing how they can be internalised by the central government see Gordon (1980). Inter-jurisdictional externalities have also been the subject of the literature on fiscal competition. For surveys on the theory of fiscal competition, see Wildasin (1988 and 1992). The advantages of centralisation are also reviewed in Prud'homme (1995).

violate local autonomy, i.e. some regions will prefer to leave the federation rather than being part of it under an efficient arrangement. This means that, if centralised government is in the position to violate local autonomy while decentralisation (by definition) has to respect it, the former will strictly dominate the latter.[4]

On the other hand, several *advantages of decentralisation* (=allocating functions to lower levels of government) have been identified, such as better matching of services to local tastes and incomes or reduced bureaucracy arising from the fact that local government is somehow "closer to the people" and therefore knows their preferences more precisely and serves their interests better than federal government.

Probably the most elegant and most influential contribution advocating the superiority of decentralised government has been Tiebout's (1956) article. As an answer to Samuelson's (1954) seminal paper, Tiebout developed the argument that profit-maximising local authorities would compete for residents who "vote with their feet" for their most preferred package of taxes and local public services, thereby truthfully revealing their preferences for public services. He showed that, under certain conditions, the resulting "decentralised competitive equilibrium" is Pareto-efficient.[5] The advantages of decentralisation have lead many people to uphold the so-called "principle of subsidiarity" which states that authority should be decentralised to local government, unless there are compelling reasons for centralisation.[6]

How does the literature on federalism relate to the theory of organisation and to the puzzle stated above? In my view, there is a very close connection between the two.[7] In fact, the puzzle precisely applies to the problem of federalism, and any satisfactory theory of decentralised government should bypass it through a well-founded argument for the strict superiority of decentralisation. Unfortunately, most of the traditional theory of federalism does not satisfy this criterion. A closer inspection of that literature reveals that the superiority of decentralised government typically relies on either unmotivated or implicit

---

4. See also Farrell (1987) and Greenwood and McAfee (1991).

5. Tiebout's argument has been criticised, among others, by Pestieau (1977) and Bewley (1981). Bewley demonstrates that Tiebout's theorem requires extremely strong assumptions, making local public goods essentially identical to private goods.

6. On the principle of subsidiarity see Begg et al. (1993).

7. This view has been advanced independently by Crémer, Estache and Seabright (1994).

constraints on central government. The most commonly imposed constraint is that central government is required to use uniform policies throughout the federation, while decentralised government is able to implement differentiated policies in a heterogeneous federation.[8] It is difficult to see why a benevolent and rational government would use a uniform policy when such policy is inefficient. Once such constraints are removed, the superiority of decentralisation disappears with them.

As a consequence, a literature has grown showing how a benevolent, unconstrained central government can always implement efficient policies in a federation; see e.g. Gordon (1980), Wildasin (1989), Krelove (1992), Piketty (1993) and, in a more general context, Varian (1992). These articles characterise policies which are designed to internalise inter-jurisdictional externalities in a Pogouvian fashion. They make explicit the relevance of the aforementioned puzzle for the theory of federalism. Similarly, Tiebout's theorem is not actually about decentralisation. It merely characterises an optimal mechanism for local public goods economies. A benevolent, unconstrained centre just needs to follow the rule given in Tiebout's article, namely to maximise profits in each locality, thus implementing an optimal allocation.[9] This illustrates very nicely, how the puzzle directly applies to the problem of fiscal federalism. In the following section, I argue that the avoidance of this puzzle depends on the view of government one chooses to adopt.

1.2 Decentralisation and alternative views of government

In the previous section, I argue that the traditional literature on federalism lacks an appropriate foundation. In this section, I assess two ways in which such a foundation and a meaningful trade-off between costs and benefits of decentralised government can be established: This is possible if government is benevolent but there exist constraints on the possible contracts which can be enforced (i.e. in a world of incomplete contracts). A second possible foundation for decentralisation arises if government is seen as self-inter-

8. Other constraints are sometimes imposed on the choice of instruments or on the information available to the centre. For example, even in studies where the uniformity constraint is forcefully rejected such as in De Lecea (1983), the central government is implicitly barred from using intergovernmental transfers in order to reveal local information.

9. This interpretation has been suggested by several authors, including Weldon (1966, p. 231), Breton and Scott (1978, pp. 40-41), Bewley (1981), Hamlin (1991, p. 195) and Seabright (1995, p. 3).

ested rather than benevolent. I review here the recent literature in the light of these two types of answers to the puzzle stated above.

1.2.1. Decentralisation with benevolent government and incomplete contracts

In a world of benevolent government and incomplete contracts, the argument used to explain the puzzle stated in the previous section does not go through any more, since the centre will not always be able to induce efficient behaviour from all agents as its instruments are limited. In this case decentralisation of authority may be a way to make better use of the limited range of enforceable contracts and, thus, decentralisation may strictly dominate centralisation.

A very nice illustration of this can be found in Kehoe (1989). He demonstrates that if the central government is not able to credibly commit to an *ex ante* optimal future policy of capital taxation, then individuals will save too little. The reason for this is that they anticipate that *ex post* it will be optimal for the government to tax capital at a very high rate since such a tax will (*ex post*) be non-distortionary. This time consistency problem is weakened when fiscal authority is allocated to the local governments, since they will *ex post* not be able to set too high rates as this would erode their tax base (= tax competition). Thus, the prospect of (*ex post*) inefficient tax competition makes decentralisation act as a commitment to low tax rates and induces individuals to save more. In this manner decentralisation may strictly dominate centralisation.

Other models where also a *commitment problem* (which is a consequence of incomplete contracts) provides the foundation for decentralised authority can be found e.g. in Melumad and Mookherjee (1991), Schmidt (1991) and Daveri and Panunzi (1995). The latter is explicitly embedded in a model of fiscal federalism. The authors build on the problem of soft budget constraint introduced by Dewatripont and Maskin (1995): A benevolent government must refinance bad projects if the social cost of stopping them in a late stage is too high. The fact that government cannot commit not to refinance bad projects gives wrong incentives *ex ante*. Daveri and Panunzi (1995) argue that under decentralisation the social cost of stopping a bad project as perceived by a local government is lower than the full social cost because of out-migration of some of the unemployed. This implies that decentralisation serves as a commitment to stop bad projects and, therefore, hardens the soft budget constraint. This leads them to conclude that, with low mobility costs, decentralisation may strictly dominate centralisation.[10]

Aghion and Tirole (1994) present another argument for decentralisation. It is based on the incentive to acquire information. Information acquisition is associated with convex costs, so that several agents should participate in it. But, as neither the effort an agent undertakes to acquire information nor an agent's information are contractible, information acquisition is samthing like a public good. The only way for several agents to induce each other to acquire information is for them to share authority (i.e. the right to choose a policy or project). Under shared authority, more agents will become informed which will be beneficial to all of them.[11]

The incomplete contract foundation for decentralised government is, of course, only as convincing as the justification of contract incompleteness. To the extent that the contract incompleteness can be endogenised, this approach can provide a satisfactory justification of decentralisation. If the endogenisation is not possible or not convincing, then the argument will be as doubtful and arbitrary as in the traditional theory of federalism.

1.2.2. Decentralisation with self-interested government

Another view of government one could adopt is that it is composed of self-interested individuals and, therefore, does not necessarily strive to maximise the welfare of its citizens. They may then be better off under an arrangement where political power is decentralised. This is the approach I am taking in part one of this thesis. I develop a number of variations on the theme that fiscal and electoral decentralisation may improve the performance of self-interested, elected officials. My model captures the idea that local government serves the interest of its citizens better, because it is better controlled by them. The interaction between voters and officials is seen as a principal-agent relationship. In chapter 2, I show that voters may prefer to strictly limit the funds they allocate to one office in order to give the officeholder a stronger incentive to try to be re-elected. A second argument for decentralisation is that splitting up the electorate into homogeneous jurisdictions (=electoral decentralisation) eliminates the possibility of officeholders to play off the voters against each other. In chapter 3, I show that voters can set up a bene-

10. Qian and Roland (1994) also develop a model where decentralisation serves to harden the budget constraint of firms.

11. A closely related argument can be found in Crémer (1992). Another type of incomplete contract argument for decentralisation, developed by Melumad, Mookherjee and Reichelstein (1992a and 1992b), is based on what these authors call "limited communication". As I explain in section 1.4, below, the general equilibrium model developed in part two of this thesis also uses an incomplete contract foundation for decentralisation.

ficial competition among officeholders and thus reduce the latter's information rents compared to centralisation.

My results can be seen in the tradition of the early models of self-interested government developed by Brennan and Buchanan (1980) or Niskanen (1971). This literature, however, uses *ad hoc* assumptions which immediately imply the superiority of decentralised government. Brennan and Buchanan (1980), for example, assume that government officials use an exogenously fixed fraction of the public funds under their responsibility for private purposes. The authors do not explain how this can actually be the outcome of political competition in a democracy. Niskanen (1971) and many others build on the so-called Leviathan hypothesis, which assumes that the government's objective is to maximise its own revenue. Again, this assumption is usually not derived from individually rational behaviour.[12] The model developed in part one of this thesis is consistent with this type of government behaviour, but derives it as the equilibrium of a game of repeated elections. In this sense one could interpret the analysis as a "micro foundation" of the public choice argument for decentralised government. Here government behaviour is explained in terms of individually rational decisions. This means that the advantages of decentralisation are derived endogenously.[13]

1.3 Overview of part one

Part one of this thesis is a contribution to the theory of elections with self-interested politicians. If one believes that politicians are driven by the same motives as other members of society, then one is immediately confronted with the following question: What induces politicians to pursue social goals and to use public funds for public rather than for their own purposes? I explore the view that politicians enjoy monetary and non-monetary benefits of being in power. Their motivation to use at least some of the public funds for social

12. The hypothesis of revenue maximising Leviathan government has been tested empirically in a number of studies, without a clear conclusion. See Oates (1985), Eberts and Gronberg (1988), Zax (1989), and Heil (1991).

13. There exists a growing literature supplying a game theoretic treatment of democratic institutions. Recent contributions include Romer and Rosenthal (1979 and 1995), Brams and Fishburn (1983), Myerson (1993a and b), Baron and Mo (1993), Londregan and Romer (1993) and many others. With the exception of Romer and Rosenthal (1995), none of these contributions deals with the issue of decentralisation.

goals is derived from their desire to be re-elected. They seek to be re-elected in order to be able to enjoy also the benefits of being in office in future periods.

In chapter 2, I develop a model of repeated elections that derives the degree of electoral accountability as a political equilibrium. The approach taken here also departs from the traditional theory of electoral competition where political platforms are determined before the election. I take the point seriously that, once elected, politicians cannot be forced to keep their promises made during the election campaigns. Under these conditions, political platforms and pre-election announcements should be seen as empty promises. I am building on Ferejohn's (1986) model of electoral control. The policy outcomes are described as a subgame perfect equilibrium of a game between voters and politicians, where only credible strategies are relevant.

Within this framework, I show that fiscal and/or electoral decentralisation may increase the accountability of elected officials, as reflected by the fraction of funds they use for the provision of public services. One implication of this approach is that, if voters realise that part of the public funds is going to be misallocated, they will vote for lower taxes compared to the first-best with benevolent government. Moreover, it follows that decisions concerning political centralisation or decentralisation should not be left to the politicians, since the latter would choose too much centralisation. This is not in contradiction to the Leviathan hypothesis, but it is derived from first principles.

1.4 Overview of part two

The puzzle stated in section 1.1 is also valid in the context of competitive general equilibrium theory. After the famous Lange-Lerner controversy[14] in the 1930 which is at the roots of modern mechanism design, we know that with complete and competitive markets general equilibrium theory does not say anything about the allocation of authority in the economy, i.e. it is "institutionally empty". In part two of this thesis, I use a competitive general equilibrium model of a private goods economy with endogenous firm formation. The economy is, however, constrained by the fact that the output produced by any production unit is not verifiable. This imperfection leads to incomplete markets and a problem of social insurance which is solved differently under different allocation sys-

14. See Lange (1936) and Lerner (1936).

tems. I compare a decentralised competitive economy with a missing insurance market to a centralised collectivist economy. Due to the non-observability of firm output, the problem is not institutionally empty. On the basis of the same fundamentals (i.e. individual preferences and endowments, technologies and information structure), I show that the two systems yield very different allocations. Either of the two systems may dominate the other in terms of the majority criterion or even the Pareto criterion.

Within this economic environment, I study the politics of transition. Starting from collectivist economy, society has a choice between total reform, gradual reform and no reform. Under total reform, the collectivist sector is closed down and all individuals have to take part in production within the new, competitive sector. Some will choose to become entrepreneurs and employ the others on a competitive labour market. Under gradual reform, individuals have the choice whether to remain in the collectivist sector or to move to the competitive sector. In general, this results in a dual economy, with some individuals being employed in each sector. I show, however, that gradual reform can never be the democratic outcome. This interesting result has several important implications which I discuss in chapter 4. I argue there, that these findings can be used to explain some recent experiences of transition economies, such as

(i) the political polarisation between pro-reformers and conservatives, and

(ii) the democratic return to power of conservatives after only one term with reformers in power.

Although the two parts of this thesis use very different approaches, they are both the result of an attempt to join economic and political thoughts within one formal analysis. Since theoretical work in this area has been surprisingly scarce, I hope that this thesis is a useful contribution to our understanding of the political economy of decentralisation.

PART ONE

Chapter 2

ELECTORAL CONTROL AND DECENTRALISATION

UNDER SYMMETRIC INFORMATION

2.1 Introduction

The main objective of this chapter is to show how decentralisation affects the degree of electoral control in a democracy. It is widely believed that democracy is a good way of controlling rulers. This belief is usually based on the view that the democratic decision process has essentially two major virtues: First, electing candidates on particular platforms seems to result in "satisfactory" or "acceptable" outcomes of the group decision process. It allows all members of society to contribute to the collective choice and is likely to produce political stability. There exists, of course, a vast literature on this aspect of democracy.[1]

A second virtue of democracy lies in its ability to limit the discretion of public office-holders. This aspect has received much less attention in the literature so far.[2] The process of repeated elections gives the voters an instrument to discipline politicians and public

officeholders. The basic tool is that an incumbent who wants to be re-elected has to act (at least to some extent) in the interest of her electorate. Otherwise, at the next election, another candidate will be elected. The model of repeated elections developed below represents the public sector in a democracy as an organisation whose beneficiaries have the authority to elect their agents, i.e. the voters are principals and the government officials are agents.[3] Voters pay taxes and expect a return in the form of a certain qualitative or quantitative level of public services. Incomplete contracts prevail in the sense that it is not possible for all the voters and bureaucrats to write an enforceable contract specifying an optimal reward structure for the officeholder as a function of the level of output(s) produced. I show how decentralisation may improve the accountability of elected officials and thus strictly dominate centralisation, from the point of view of the voters.

I introduce the distinction between fiscal and electoral decentralisation. *Fiscal centralisation* corresponds to a situation where all public funds are managed by a single bureau. In contrast to this, under *fiscal decentralisation* the budget is split and public services are produced by several offices. Under fiscal decentralisation the political system may be or not be decentralised. *Electoral decentralisation* means here that the electorate is split into several groups (or jurisdictions) electing their own officials. The two arguments developed in this chapter can be summarised as follows:

(i) Public officials want to be re-elected in o der to enjoy the monetary and (more importantly) the non-monetary, private benefits of being in office in future periods. This constitutes an incentive for them to limit the amount of public funds they use for their own purposes. Under fiscal decentralisation, private benefits are more important relative to the public funds under the control of a single officeholder. Therefore, from the point of

---

1. "Political stability" here is understood in the sense that no new majority of voters will form to overrule the outcome. The median voter theorem states that, if voters have single-peaked preferences over the possible policies, there exists a plan which will win the vote under majority voting and there will be no other majority opposed to it. The issue of stability in election outcomes is less straightforward if one is not willing to restrict attention to single-peaked preferences and uni-dimensional decisions. For a survey of the literature on the stability of voting equilibria, see e.g. Coughlin (1990) or Mueller (1990).

2. There is now a growing literature on repeated elections, trying to fill the gap: see Ferejohn (1986), Lott and Reed (1989), Austen-Smith and Banks (1991), Banks and Subramanian (1993), Harrington (1993), Myerson (1993a), and others. See also Barro (1973) and Migué and Bélanger (1974) for earlier work in this direction.

3. Everything said here does, in principle, apply just as well to the relationships between shareholders of a corporate firm and the CEO, or between lenders and sovereign borrowers. Related models on repeated lending have been developed by Allen (1983), Bolton and Scharfstein (1991), and Gromb (1994).

view of the voters, fiscal decentralisation leads to a better use of the private benefits which do not cost the voters anything. This, in turn, increases her incentive to be re-elected, and therefore reduces her incentive to misallocate the funds.

(ii) In a centralised organisation with a heterogeneous electorate and majority rule elections, the incumbent can play off the different voters in a way which allows her to be re-elected even if she misallocates most of the public funds. The basic force is that, since she only needs a majority of the votes to be re-elected, the officeholder will always choose to serve the voters with the lowest demands on her. Therefore, if he wants to be considered by the officeholder in her agenda, each voter is forced to keep his demands from the incumbent low. As voters are played off against each other, they are forced to reduce their expectations from the incumbent such that, at equilibrium, the incumbent will always be re-elected without producing any output. This result, which I call the *Ferejohn Paradox*, was first derived in a similar context in Ferejohn (1986, proposition 6). My result 4 shows that Ferejohn's analysis was incomplete in the sense that there are many more equilibria than the one identified by him, but all equilibria are characterized by zero output. The result shows in an extreme form how, under centralisation, the conflicting interests (i.e. the heterogeneity) of the principals leads to a loss of electoral control.

The basic model with one officeholder, one voter, and symmetric information is introduced in section 2.2. Section 2.4 extends the basic model to the case of many voters and many outputs and derives the Ferejohn Paradox. In sections 2.3 and 2.5, respectively, I develop arguments (i) and (ii) for decentralisation mentioned above. In section 2.5, I also discuss several institutional solutions to the Ferejohn Paradox. In particular, electoral decentralisation leading to a homogeneous electorate reduces the possibility for the incumbent to play off the voters against each other. This, in turn, means that higher output levels will be achieved under decentralisation than under centralisation. Further extensions are discussed in section 2.6.

2.2 The basic model of centralised production

In order to formalise the ideas presented above, I use an extension of Ferejohn's (1986) pioneering model of repeated elections to capture the agency cost associated with public goods provision.[4] There is one (representative) voter and several identical candidates for

one office. Their time horizons are infinite with a constant and common time discount factor δ. In each period t, the voter elects one of the candidates to office for that period. He (the voter) lends her (the officeholder or manager) the funds s (per period) for the production of an output of value g_t to the voter, with

$$g_t = \alpha_t \cdot \theta_t \cdot s \ . \tag{1}$$

α denotes the proportion of funds used "efficiently"[5] and θ is a technological random variable, uniformly distributed on the interval **[0,1]**. The realisations θ_t are independent in each period. Both α_t and g_t are assumed to be non-contractible. In this sense the agency contract is incomplete.

The "contracts" between the voters and their representatives are both incomplete and implicit (i.e. self-enforcing), so that the voters have to incur an agency cost, even if there is no asymmetric information. The contracts considered here are incomplete in a further respect. Namely, candidates cannot be asked to pay for the right to be in office. Two justifications can be given for this: First, given the size of the entire public sector in an economy the entry fee would have to be very big if it were to change incentives. Thus it is relatively plausible that an individual's initial wealth would not be sufficient to pay the necessary entry fee. Secondly, Slutsky (1986) shows, within a framework similar to the one below, that if voters were actually allowed to charge an entry fee, they would do so in equilibrium, and the office would not produce any output. It is therefore very likely that voters will find arrangements which rule out entry fees. The contracts considered here are also *implicit*, in the sense that they do not commit any party to take an action at some point in time in the future which, *at that point*, will not be optimal. This is equivalent to saying that the outcome is a subgame perfect equilibrium of the game played between the contracting parties (see Bull, 1980). Similar reasons that prevent a contract from being complete, together with commitment problems on the side of the consumers, justify the assumption of implicit contracts.

---

4. A main difference to Ferejohn is that I introduce a variable size of the budget and consider its effect on the bureau's performance.

5. "Efficiently" means here that the funds are used for the production of the required public services, i.e. in the interest of the voter(s). It is assumed here that although g_t could be a vector, α_t must be a scalar such that the officeholder's allocation decision is one-dimensional. The case where the officeholder chooses a vector α_t is discussed below in sections 2.4 and 2.5.

Throughout the thesis, I am using Greek letters to denote variables which can take values in the interval **[0,1]**. The officeholder's period utility is

$$w + (1 - \alpha_t) \cdot s \tag{2}$$

where w denotes the private benefits of office. This specification of the agency problem is slightly different from the standard effort model where the agent has a disutility of effort independently of the size of the project. Here α is a proportion of s;[6] this does not guarantee an interior solution and, consequently, the boundary constraints will play a prominent role in the analysis. I believe that, in large projects, the manager's "integrity" in allocating the funds is more important than her personal, non-monetary effort. This is equivalent to saying that the relevant problem is not whether she works hard or not, but rather in whose interest she acts. Therefore, I believe that the resource cost model used here is more appropriate in the present context than the effort cost model.

In the symmetric information case discussed here, immediately after the election in period t, θ_t is publicly observed. I assume that the voter uses a *one period retrospective voting rule (RVR1)*[7], i.e. he specifies a cutoff level of output $\hat{g}_t$, as a function of θ_t below which he will not re-elect the incumbent. When g_t is above the cutoff level, then the incumbent is re-elected for period $t+1$. Otherwise, the period t incumbent is voted out of office forever, and a new incumbent takes her place.[8] Since all candidates have identical abilities, it does *ex post* not matter for the voter who will actually be elected. Notice that, at the time of his re-election decision for the period $t+1$ office, the voter knows the realisation of the random variable in period t, θ_t. He can therefore make the cutoff level contingent on it.

---

6. Ferejohn's (1986) formulation is more general and allows for both versions. In his discussion of Ferejohn's model, Slutsky (1986), uses the effort model with quadratic disutility. The resource cost model used here is similar to the one in Bolton and Scharfstein (1990) and Gromb (1994) where the agent reports on the realization of the profits which are not observed by the principal. The main difference here to those latter papers is that the principals can replace the agent. Also, I consider only a restricted range of strategies; see below.

7. On the concept and the evidence of retrospective voting, see Fiorina (1981).

8. In other words, the voter is restricted to use pure Markov strategies. Slutsky (1986) shows that with more-than-one-period RVR's there exist both equilibria where the voter does better or worse than with a RVR1. See also Ledyard (1986). It can be shown that the voters cannot do better by using a *probabilistic* instead of the deterministic RVR1.

Definition: (**1-1**-equilibrium) *A political equilibrium of centralised production with a representative voter is a sequence* $\{\alpha_t, \hat{g}_t(\theta_t)\}_{t=0,\ldots}$, *such that*

(i) the voter uses a RVR1, i.e. in every period t, he sets an optimal cutoff level $\hat{g}_t$ *(as a function of* θ_t*), taking the incumbent's best-response pattern as given;*

(ii) in every period t, the incumbent chooses α_t *such as to maximise her inter-temporal utility, taking* $\{\hat{g}_t(\theta_t)\}_{t=0,..}$ *as given.*

The voter chooses $\hat{g}_t(\theta_t)$ strategically such as to provide optimal incentives to the office-holder. Clearly, the voter would always prefer the output to exceed $\hat{g}_t$ but he takes into account that an excessive cutoff level will have the effect that the officeholder will prefer not to be re-elected and will enjoy the short run benefits of "shirking". Notice also that the voter commits at the beginning of the period to the function $\hat{g}_t(\theta_t)$ while, at the end of the period, he is always indifferent whether to re-elect the incumbent or not, i.e. the RVR1 is time consistent or subgame perfect. This is because all candidates for office are assumed to be of equal ability. The analysis with candidates of different abilities is much more complicated (see Banks and Subramanian, 1991).

Result 2.1: *The unique 1-1-equilibrium is characterised as follows: In every period t, the voter requires output of at least*

$$\hat{g}_t(\theta_t) = min\,[\theta_t s\ ,\ \theta_t \delta\,(w+s)]\ , \tag{3}$$

in order to re-elect the incumbent. The latter will always satisfy this requirement (and be re-elected) by using a proportion

$$\hat{\alpha}_t = min\left[1\ ,\ \delta\left\{1+\frac{w}{s}\right\}\right] \tag{4}$$

of the funds for production and $(1-\hat{\alpha}_t)$ *for her own purposes.*

DEMONSTRATION OF RESULT 2.1: Suppose first that the voter's sequence of cutoff levels, $\hat{g}_t(\theta_t)\}_{t=0,\ldots}$, is given. Denote $v_t(\alpha_t) \equiv w + (1-\alpha_t)s + \delta v_{t+1}$ the value to the officeholder

of being in office in period t. Notice that v_{t+1} is independent of α_t and g_t, because of the RVR1. Then, the period t incumbent's optimal choice of α_t is either equal to $\hat{\alpha}_t(\theta_t) \equiv \hat{g}_t / (\theta_t s)$, which is the minimum α_t which allows her to be re-elected, or equal to *0*. Any other possible choice of α_t cannot be optimal, since it would only entail more disutility for the incumbent without affecting the voter's re-election decision. This implies that the incumbent will want to stay in office if and only if

$$v_t(\hat{\alpha}_t) \geq w + s \quad . \tag{5}$$

Given this best-response pattern for the incumbent, the voter will choose a sequence $\{\hat{g}_t\}_{t=0,\ldots}$ which maximises his utility. For each period, he will try to set the cutoff level as high as possible, taking into account θ_t, and bearing in mind that the incumbent will only have an incentive to be re-elected if condition **(5)** holds. It is in the interest of the voter to make it hold with equality. In each period, he will therefore set

$$\hat{g}_t(\theta_t) = \theta_t \delta v_{t+1} \tag{6}$$

whereby I made use of the above definition of v_t, together with **(5)**. Using **(1)**, equation **(6)** yields

$$\hat{\alpha}_t = \frac{\delta v_{t+1}}{s} \quad . \tag{7}$$

Due to the stationarity of the problem, the value of office, $v=v_t=v_{t+1}$, can be found by solving the following equation[9]

$$v = w + \left[1 - \frac{\delta v}{s}\right] s + \delta v \tag{8}$$

---

9. The fact that equilibrium is stationary follows from the assumption that the voters use a RVR1, together with the stationarity of the environment, i.e. voters and incumbents face the same problem in every period.

which, after solving for v, yields

$$v = w + s \ . \tag{9}$$

Substituting v from **(9)** into equations **(7)** and **(6)** yields the following expressions, respectively:

$$\hat{\alpha}_t = \delta \left(1 + \frac{w}{s} \right) \tag{10}$$

and

$$\hat{g}_t(\theta_t) = \theta_t \delta (w + s) \ . \tag{11}$$

Equations **(10)** and **(11)** are, however, only relevant as long as the boundary constraint, $\alpha \leq 1$, which represents the officeholder's budget constraint, is not binding. Using equation **(10)**, the boundary constraint becomes

$$\frac{w}{s} \leq \frac{1 - \delta}{\delta} \ . \tag{12}$$

Taking **(12)** into account, equations **(10)** and **(11)** yield equations **(3)** and **(4)**, respectively, as stated in result 2.1.[10] ❑

The intuition behind equation **(9)**, which states that the (stationary) value of being in office is simply equal to the private benefit enjoyed by the incumbent in one period plus the value of public funds entrusted to her, is as follows: The incumbent receives *was* in the first instance. v cannot be smaller than this because the incumbent has always the option to walk away with *was*. On the other hand, voters try to keep v as low as possible by their choice of $\hat{g}$. The result is equation **(9)**.

10. When (12) is binding, the stationary value of office is $vs/\delta = w/(1-\delta)$.

Notice that the officeholder's equilibrium "integrity", $\hat{\alpha}_t$, is independent of θ_t. Intuitively, this is because the voter takes θ_t into account already when he chooses $\hat{g}_t$. He adjusts $\hat{g}_t$ such as to require the same $\hat{\alpha}_t$ in every period.

Equations **(10)** and **(11)** reflect that the basic motivation for an officeholder to use the funds for production is that she wants to be re-elected in order to enjoy the future benefits of being in office. These benefits consist in each period of the private benefits, on the one hand, and the misallocated funds, on the other hand.

Not surprisingly, equilibrium output g_t is increasing in δ, w, and s. The comparative statics of result 1 with respect to the budget size s shows that if the private benefits of office w are sufficiently large compared to the funds lent to the incumbent, then the boundary constraint is binding and the officeholder has an incentive to use all the funds for production. But if s becomes relatively large, the incumbent will be tempted to "walk away" with it. Therefore, the voters tend to loose control as s is increased relative to w, as depicted in figure 2.1.

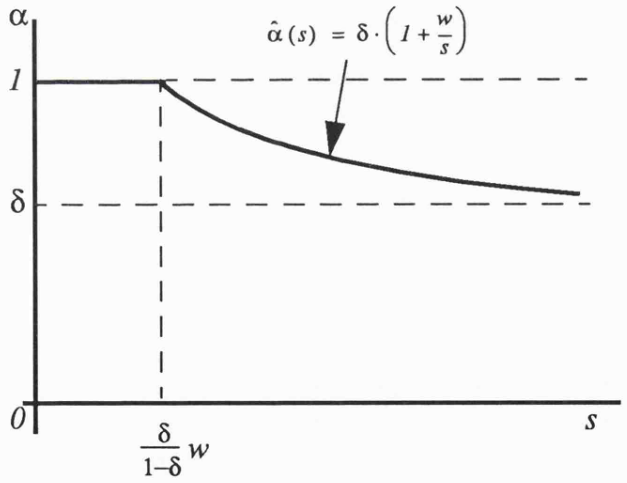

FIGURE 2.1: *The voter tends to lose control as the bureau's budget is increased.*

The same reasoning implies that, keeping s constant, for low values of office w, $\hat{\alpha}$ is strictly increasing until it reaches the upper bound at $\hat{\alpha} = 1$. This is shown in figure 2.2.

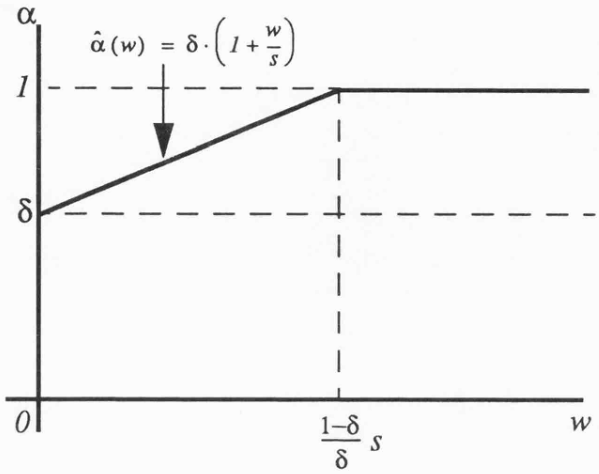

FIGURE 2.2: *Higher private benefits of office induce a higher productivity.*

The comparative statics with respect to the discount factor δ is also intuitive. For small δ, the incumbent does not value the future much, which - given that the only incentive scheme available to the voter is his re-election rule - makes it more expensive for the voter to induce the officeholder to invest the funds in his best interest. Clearly, as the officeholder cares more about the future - in particular about the private benefits of office which she could enjoy - it becomes easier for the voter to control her.

2.3 Fiscal decentralisation I: The role of private benefits of office

The simple framework presented in the previous section shows that, although the production technology has been assumed linear, the particular nature of the agency problem results in diminishing returns to scale beyond a certain threshold, $\hat{s} \equiv \frac{\delta}{1-\delta} w$, as drawn in figure 2.3. This implies that the voter can benefit from splitting the office and electing several officeholders independently. In other words, for given private benefits of office accruing to the officeholder and not too large values of the discount factor (i.e. $\delta < s/(w+s)$), the smallest optimal number of offices increases with the budget size. The reason for this is that fiscal decentralisation allows the voter to use the private bene-

fits of office - which represent a cost free means of giving the incumbent an incentive to try to be re-elected - more efficiently.

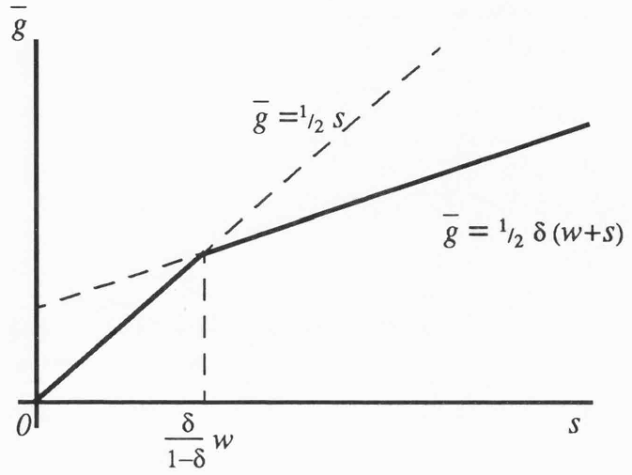

FIGURE 2.3: *Diminishing returns to scale beyond* $\delta\, w/(1-\delta)$.

The crucial role of the private benefits is highlighted even more when they are allowed to vary with the amount of funds allocated to the office. Keeping everything else constant, one might expect the private benefits of each office to be a decreasing function of the total (integer) number of offices, N. This clearly weakens the advantage of decentralisation. But, as long as this function, $w(N)$, is not decreasing too quickly, decentralisation can still be beneficial. The findings can be summarised in the following statement.

Result 2.2: *If the total budget s is sufficiently large relative to the private benefits of office w, and the latter are not decreasing too quickly in the number of offices, then the voter will strictly prefer to have more than one office producing the services. In other words, fiscal decentralisation (i.e. N>1) is strictly beneficial to the voter if*

$$\frac{s}{w(1)} > \frac{\delta}{1-\delta} \quad \wedge \; Nw(N) > w(1), \;\; \exists N > 1 \;\; .$$

DEMONSTRATION OF RESULT 2.2: Denoting s_n the total funds allocated to and g_{nt} the output produced by office n. For simplicity, assume that the budget is split equally under fiscal decentralisation, i.e. $s_n = s/N$. According to result 1, total output g_t available to the voter in each period (in equilibrium) is $\sum_{n=1}^{N} \hat{g}_{nt}$, whereby $\hat{g}_{nt}$ are determined in the

same way as $\hat{g}_t$ in equation (3) with the only difference that, now, w is treated as a function $w(N)$ of the number of offices. Total output in period t will therefore be equal to

$$\sum_{n=1}^{N} \hat{g}_{nt} = N \cdot min\left[\theta_t s_n \ , \ \theta_t \delta \{w(N) + s_n\}\right] \ . \qquad (13)$$

From the first condition stated in result 2.2 (and recalling result 2.1), it follows that total output under centralisation will be $g_t=\theta_t\delta[w(1)+s]$.

In the remainder of this demonstration, I will show that $Ng_{nt}>g_t$ for some $N>1$, which is equivalent to

$$N \cdot min\left[\theta_t s_n, \theta_t \delta \{w(N) + s_n\}\right] \geq \theta_t \delta \{w(1) + s\} \ , \ \exists N \quad . \qquad (14)$$

This condition consists of two possible cases:

Case 1: $\theta_t s_n \leq \theta_t \delta \{w(N) + s_n\}$

In this case, total output under decentralisation is $N\theta_t s_n = \theta s > \theta_t \delta \{w(1) + s\}$ for any $N>1$ because of the first condition stated in result 2.2.

Case 2: $\theta_t s_n > \theta_t \delta \{w(N) + s_n\}$

In this case, total output under decentralisation is:

$N \theta_t \delta \{w(N) + s_n\} = \theta_t \delta \{Nw(N) + s\} > \theta_t \delta \{w(1) + s\}$ for at least some $N>1$ because of the second condition stated in result 2.2.

This shows that, in any of the two cases, output under decentralisation will be bigger than under centralisation. ❑

It is important to notice that decentralisation here simply means setting up more than one bureau, elected by the same electorate. This is what I called *fiscal decentralisation* in the introduction, as opposed to electoral decentralisation. The result illustrates one advan-

tage of fiscal decentralisation which leads to a more efficient use of the incumbent's private benefits of being in office.

Recently, I became aware of a working paper by Paul Seabright in which he studies electoral control and decentralisation using a very similar framework to mine. He also considers incomplete contracts between voters as principals and government officials as agents and finds that the latter's accountability can be improved through decentralisation. In Seabright (1995), as here, voters re-elect an incumbent officeholder if and only if she provides a satisfactory level of public services, $\hat{g}$. The main difference is that in Seabright's model, $\hat{g}$ is determined exogenously whereas here it is chosen optimally by the voters as part of the equilibrium in the re-election game played between the voters and the incumbent.[11]

Result 2.2 can also be interpreted as one of credit rationing. If the repayment of a business loan has to be self-enforcing (because of the hold-up problem), then the lender will not want to extend the amount of the loan beyond the point $\hat{s} \equiv w\delta/(1-\delta)$ where the manager cannot be given the right incentive to repay the full amount. The argument is closely related to, but different from, Allen's (1983). There also, the value of being in office is non-proportional to the budget size (i.e. the loan size). But in Allen's model the reason why the value of office is increasing less than proportionally lies basically in the concavity of the production function and not, as here, the presence of the private benefits of office.

2.4 Many voters, many outputs and loss of control

So far it was assumed that one (representative) voter was electing a manager who had to produce a single output. This section extends the analysis to the cases of many voters (section 2.4.1) and many outputs (section 2.4.2). The framework is thus one of repeated common agency with incomplete contracts.[12]

---

11. In my survey (1995), I compare and relate this paper to Seabright's and to other recent work on the theory of fiscal federalism.

12. Common agency models with complete contracts have been analysed e.g. by Bernheim and Whinston (1986), Berkok (1990) and Galor (1991).

2.4.1. Many voters and a single output

Suppose there are N voters indexed $n=1,\ldots, N$ who use a RVR1 when casting their votes. This means that each voter n sets a minimum required level of output, $\hat{g}_{nt}(\theta_t)$, for re-electing the period t incumbent, taking all other voters' cutoff rules, $\hat{g}_{it}(\theta_t)\ ,\forall i \in \mathcal{E}\setminus\{n\}, \forall t$, and the incumbent's best-response pattern as given. $\mathcal{E}$ denotes the set of all voters. Let $\hat{g}_t^{-n}$ denote the vector of all other voters' cutoff rules. The following definition is a natural extension of **1-1**-equilibrium to the case of many voters.

Definition: *(N-1-equilibrium) A political equilibrium with N voters and M-majority voting is a sequence of vectors* $\{\alpha_t; \hat{g}_{1t}(\theta_t),\hat{g}_{2t}(\theta_t),\ldots,\hat{g}_{Nt}(\theta_t)\}_{t=0,\ldots}$, *such that*

(i) every voter uses a RVR1, i.e. in every period t, each voter sets an optimal cutoff level, $\hat{g}_{nt}(\theta_t)$, *taking the sequence of vectors of all other voters' cutoff rules* $\{\hat{g}_t^{-n}(\theta_t)\}_{t=0,\ldots}$ *and the incumbent's best-response pattern as given;*

(ii) in every period t, the incumbent chooses α_t *such as to maximise her inter-temporal utility, taking* $\{\hat{g}_t(\theta_t)\}_{t=0,\ldots}$ *as given;*

(iii) the period t incumbent is re-elected for t+1 if and only if the demands of a majority of M voters (where N/2 < M ≤ N) are satisfied. The votes of all unsatisfied voters go to the same alternative candidate.[13]

Parts (i) and (ii) are the same as in **1-1**-equilibrium. Part (iii) introduces the majority rule. The simple majority rule corresponds to the special case of a M-majority rule with $M=(N+1)/2$.

Preliminary result: *Denote* $\hat{g}_{Mt}(\theta_t)$ *the M-th lowest component of the vector* $\hat{g}_t \equiv \{\hat{g}_{1t}(\theta_t),\hat{g}_{2t}(\theta_t),\ldots,\hat{g}_{Nt}(\theta_t)\}$. *Any sequence* $\{\hat{g}_{Mt}/(\theta_t s); \hat{g}_t\}_{t=0,\ldots}$ *is a N-1-equilibrium.*

---

13. This definition has a problem, though, for $M>(N+1)/2$. It implies that the incumbent will be replaced by another candidate although the other candidate may actually have less votes in her favour. This inconsistency does not apply when $M \leq (N+1)/2$.

DEMONSTRATION OF THIS RESULT: Consider the problem of a voter n, for any given θ_t and $\hat{g}_t^{-n}$ in period t. To see that voter n is always indifferent about his choice of $\hat{g}_{nt}$, denote $\hat{g}_{Mt}^{-n}$ the M-th lowest component of the vector $\hat{g}_t^{-n}$ and $\hat{\alpha}_{Mt}^{-n} \equiv \hat{g}_{Mt}^{-n} / (\theta_t s)$. First, notice that, given the vector $\hat{g}_t^{-n}$, n is always indifferent between all $\hat{g}_{nt} \geq \hat{g}_{Mt}^{-n}$. This is because any $\hat{g}_{nt} \geq \hat{g}_{Mt}^{-n}$ will not induce the incumbent to change α_t. Clearly, any $\hat{g}_{nt} < \hat{g}_{Mt}^{-n}$ cannot be optimal since it would induce the incumbent to choose $\alpha_t < \hat{\alpha}_{Mt}^{-n}$. Since this argument holds for any vector $\hat{g}_t^{-n}$ and for any voter n, the result follows. ❑

This result shows that, without additional restrictions on the equilibrium concept, no particular outcome can be predicted. As result 2.3 shows, the following assumption is sufficient for a unique equilibrium. Moreover, this equilibrium corresponds to the one with a single voter.

Assumption $(\hat{g}MAX)$ **:**

In any period t, it is the case that

$$\hat{g}_{nt} = min\left[max\left[\hat{g}_{nt}^{*}(\theta_t), \hat{g}_{nt}^{**}(\theta_t), \ldots\right], min\left[\theta_t s, \theta_t \delta (w+s)\right]\right], \forall n \in \mathcal{E}$$

where $\hat{g}_{nt}^{*}(\theta_t)$, $\hat{g}_{nt}^{**}(\theta_t)$, ... *are optimal rules for voter n.*

Assumption $(\hat{g}MAX)$ says that, whenever a voter is indifferent between several optimal choices of cutoff level, he chooses the highest of them. However, his choice will never exceed the level which the incumbent would be willing to fulfil if n was the only voter. This latter level is, of course, the one derived in section 2.2 (equation **(3)**).

Result 2.3: *Under assumption* $(\hat{g}MAX)$ *, the unique N-1-equilibrium is the sequence*

$$\{\alpha_t; \hat{g}_{1t}, \hat{g}_{2t}, \ldots, \hat{g}_{Nt}\}_{t=0,\ldots} ,$$

where $\alpha_t = min[1, \delta\{1 + w/s\}]$ *and* $\hat{g}_{nt} = min[\theta_t s, \theta_t \delta\{w+s\}]$*, for all voters n, for all t, i.e., the unique equilibrium is the same as in the case of a single voter.*

DEMONSTRATION OF RESULT 2.3: Given the sequence of vectors $\{\hat{g}_t^{-n}\}_{t=0,\ldots}$, the incumbent's best-response pattern is now as follows: Set $\alpha_t = \hat{\alpha}_{Mt}$ if

$$(1-\hat{\alpha}_{Mt})\,s + \delta v_{t+1} \geq w + s \ , \qquad \textbf{(15)}$$

and $\alpha_t=0$ otherwise, where $\hat{\alpha}_{Mt} \equiv \hat{g}_{Mt} / (\theta_t s)$, and $\hat{g}_{Mt}$ denotes the M-th lowest component of the vector $\hat{g}_t = (\hat{g}_{1t}, \hat{g}_{2t}, \ldots, \hat{g}_{Nt})$.

Consider the problem of voter n in period t, for any θ_t and for any $\{\hat{g}_t^{-n}\}_{t=0,\ldots}$.

Denote $\hat{g}_{Mt}^{-n}$ the M-th lowest component of the vector $\hat{g}_t^{-n}$, and $\hat{\alpha}_{Mt}^{-n} \equiv \hat{g}_{Mt} / (\theta_t s)$. The following three cases are possible:

(1) $\hat{g}_{(M-1)t}^{-n} > min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)]$

(2) $\hat{g}_{(M-1)t}^{-n} \leq min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)] \wedge \hat{g}_{Mt}^{-n} > min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)]$,

(3) $\hat{g}_{(M-1)t}^{-n} \leq min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)] \wedge \hat{g}_{Mt}^{-n} \leq min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)]$,

where $\hat{g}_{(M-1)t}^{-n}$ is the (M-1)-st lowest component of $\hat{g}_t^{-n}$.

In case (1), for any choice of $\hat{g}_{nt}$, $\hat{g}_{Mt}$ will be greater than $min[\theta_t s, \theta_t \delta(w+s)]$. Thus, condition **(15)** implies that the incumbent will set $\hat{\alpha}_t = 0$, no matter what voter n does. Therefore he will set $\hat{g}_{nt} = min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)]$, out of indifference (because of assumption $(\hat{g}MAX)$). In case (3), condition **(15)** implies that the incumbent will set $\alpha_t = \hat{\alpha}_{Mt}^{-n} = min\,[1 \ , \ \delta\,(1+w/s)]$, no matter what voter n does.

This is because, for any choice $\hat{g}_{nt}$, $\hat{g}_{Mt}$ will be equal to $min\,[\theta_t s \ , \ \theta_t \delta\,(w+s)]$. There-

fore, voter n will again set $\hat{g}_{nt} = min[\theta_t s, \theta_t \delta(w+s)]$, out of indifference (because of assumption $(\hat{g}MAX)$).

In case (2), finally, $\hat{g}_{Mt}^{-n} > min[\theta_t s, \theta_t \delta(w+s)]$ implies that $\hat{\alpha}_{Mt}^{-n} = 0$. In this case, voter n's choice matters. An optimization similar to the one used to derive result 2.1, yields that voter n's unique optimal choice is $\hat{g}_{nt} = min[\theta_t s, \theta_t \delta(w+s)]$. Since this is true for any voter $n \in \mathcal{E}$, result 2.3 follows. ❑

2.4.2. Many voters, many outputs and the Ferejohn paradox

So far, α_t was assumed to represent a one-dimensional allocation decision. In this section the incumbent's allocation problem is across N different products. One can think of N local public goods in N different regions. It is assumed here that there is one representative voter in each region.[14] Under centralisation, the N regions elect a manager to office, in every period, and give her the task to produce an output for each region, i.e. a vector of outputs $(g_1, g_2, \ldots, g_N)$. The manager receives the total tax revenue s and has to produce N different local public goods. The difference to the previous subsection is that the central bureaucrat has to allocate the total amount of the funds she uses for production (i.e. αs) across the N local jurisdictions. For each local public good, the technology is the same as before, with the additional assumption that there is no uncertainty about the technology, i.e. $\theta_{nt}=1$, for all n, for all t. Since this is a model of symmetric information this assumption only simplifies the notation. The production technology is thus simply

$$g_{nt} = \alpha_{nt} \cdot s \quad , \tag{16}$$

where α_{nt} is the proportion of total public funds used for the production of the local public good in region n. This means that the boundary constraint $\alpha_t \leq 1, \forall t$, applies,

---

14. As shown in result 2.3, assumption $(\hat{g}MAX)$ guarantees that a homogeneous electorate with many voters behaves as a single voter.

whereby $\alpha_t \equiv \sum_{n=1}^{N} \alpha_{nt}$. Total public funds are raised from the regions such that, $s \equiv \sum_{n=1}^{N} s_n$, where s_n denotes the funds collected in region *n*.

Section 2.3 focuses on the benefits of decentralisation derived from a better use of the private benefits of office. In order to separate that effect from the effect discussed here, I will assume from now on that $s/w(1) \leq \delta/(1-\delta)$ and $w(1) = N\,w(1)$, for all *N*. I.e., the private benefits of office for the central officeholder are assumed to be sufficiently big to eliminate the advantage of decentralisation discussed in section 2.3.

Definition: (*N-N*-equilibrium) *A political equilibrium with N voters and N services is a sequence* $\{\alpha_{1t},\alpha_{2t},\ldots,\alpha_{Nt};\hat{g}_{1t},\hat{g}_{2t},\ldots,\hat{g}_{Nt}\}_{t=0,\ldots}$ *, such that*

(i) voters use a RVR1; i.e. in every period t, each voter sets an optimal cutoff level, $\hat{g}_{nt}$*, taking* $\hat{g}_{it}$*, for all* $i \in \mathcal{E}\setminus\{n\}$*, for all t, and the incumbent's best-response pattern as given;*

(ii) in every period t, the incumbent chooses $(\alpha_{1t},\alpha_{2t},\ldots,\alpha_{Nt})$ *such as to maximise her inter-temporal utility, taking* $\{\hat{g}_t\}_{t=0,\ldots}$ *as given;*

(iii) the period t incumbent is re-elected for t+1 if and only if the requirements of a majority of M voters are satisfied, where M is a positive integer from the range $N/2 < M \leq N$*. The votes of all unsatisfied voters go to the same alternative candidate.*

This definition is a straightforward extension of *N-1*-equilibrium. Notice that now the incumbent has also to choose how to allocate the funds she decides to use efficiently across the projects, i.e. she chooses a *vector* α_t. The following result extends proposition 6 in Ferejohn (1986).

Result 2.4: (Ferejohn Paradox) *In any N-N-equilibrium under M-majority rule* $(N/2 < M < N)$*, there will be no output from the bureau. All equilibria are characterised by zero output and voters being indifferent about their choice of rules. All sequences of vectors* $\{0,0,\ldots,0,\hat{g}_t\}_{t=0,\ldots}$ *are equilibria, as long as* $\hat{g}_t$ *is zero in any M of its components. Formally,*

$$\hat{g}_t = (\hat{g}_{mt}=0;\hat{g}_{kt}= c_{kt}), \quad \forall m \in \mathcal{A}_t, \ \forall k \in \mathcal{E}\backslash\mathcal{A}_t, \forall \mathcal{A}_t \subset \mathcal{E}, \ \forall c_{kt} \in [0,\infty], \forall t \ ,$$

where $\mathcal{E}$ denotes the set of all voters, $\mathcal{A}$ denotes any subset of $\mathcal{E}$ containing exactly M elements and c_{kt} are real numbers from the specified interval.

DEMONSTRATION OF RESULT 2.4: Since the incumbent only needs a majority of M votes for re-election, $(\alpha_{1t},\alpha_{2t},\ldots,\alpha_{kt})$ will always be 0 in at least N-M components. Moreover, the incumbent will choose to serve those M voters which have the M lowest requirements. This means that, given the sequence $\{\hat{g}_t\}_{t=0,\ldots}$, the incumbent's best-response pattern is as follows: Set $\alpha_{mt} = \hat{g}_{mt}/s$ for all $m \in \mathcal{A}_t$ and $\alpha_{lt} = 0$ for all $l \in \mathcal{E}\backslash\mathcal{A}_t$ if

$$\left(1 - \sum_{m=1}^{M} \hat{\alpha}_{mt}\right)s + \delta v_{t+1} \geq w + s \ , \tag{17}$$

and $\alpha_{nt} = 0$, for all $n \in \mathcal{E}$, otherwise, where $\mathcal{A}_t$ denotes the set of those M voters who are easiest to satisfy in period t, i.e. the voters with the M lowest cutoff levels $\hat{g}_{mt}$.[15]

The remainder of this demonstration shows that $\hat{g}_t$ is an equilibrium if and only if it implies $\hat{\alpha}_{nt} = 0$, for all $n \in \mathcal{E}$ which, in turn, implies $\alpha_{nt} = 0$, for all $n \in \mathcal{E}$. Consider any $\hat{g}'_t$ which does not induce $\alpha'_{nt} = 0$, for all $n \in \mathcal{E}$, and which therefore has at least N-M+1 strictly positive components. Such a $\hat{g}'_t$ cannot be an equilibrium since any of the N-M voters who are not being served could induce the incumbent to serve them by setting $0 < \hat{g}'_{kt} < min[\hat{g}'_{mt}], k \in \mathcal{E}\backslash\mathcal{A}_t, m \in \mathcal{A}_t$. This shows that any equilibrium vector $\hat{g}_t$ will be 0 in at least M of its components. This implies that all voters $k \in \mathcal{E}\backslash\mathcal{A}_t$ are indifferent between any $\hat{g}_{kt} \in [0,\infty]$. ❑

15. If there are several sets of M voters who are easiest to satisfy, then $\mathcal{A}_t$ denotes any one of them.

It should be mentioned that Ferejohn's analysis is carried out under asymmetric information. My result 2.4 shows that the zero output result derived by Ferejohn also holds under symmetric information. Contrary to Ferejohn's claim, however, my result shows that there is a multiplicity of zero output equilibria (this is true also under asymmetric information).

Result 2.4 shows in an extreme form, how the fact that there are many principals with conflicting interests gives the agent the power to play off the principals against each other and extract the maximum rent. The voters' problem is that, if one of them would try to set a higher cutoff level, then he would not be served by the incumbent. The latter always chooses just a majority of voters to serve, namely those with the lowest cutoff levels. The voters will "undercut" each other until their cutoff levels are driven down to zero. Notice that the result holds for any majority rule (strictly) between 50 and 100 per cent, i.e. for any M in the range $N/2<M<N$.

The result is also not affected if inter-jurisdictional benefit spillovers are introduced. In this case, only equation **(16)** has to be re-written as

$$g_{nl} = \alpha_{nl}s + \sum_{m=1}^{M} \beta_{nk}\alpha_{kl}s \ , k \neq n \quad , \tag{18}$$

with $0 \leq \beta_{nk} < 1$, β_{nk} measuring the degree of spillover from region k to region n. Replacing equation **(16)** with **(18)** means that local public goods are allowed to be *impure* local public goods, excluding only the case of a pure global public good (which corresponds to $\beta_{nk}=1$). The reader can easily check that, with this re-formulation, the reasoning underlying result 2.4 still applies. This is due to the fact that the underlying mechanism with regions which are in the minority "undercutting" the regions which are in the majority still works in the same way as without spillovers. In other words, even in the presence of benefit spillovers, for any voter, there always exists a sufficiently small (positive) number μ such that he prefers an allocation with $\alpha_n=x-\mu$ and $\alpha_k=0$ to the allocation $\alpha_n=0$ and $\alpha_k=x$, for some $k \neq n$, for any $x>0$.

The result is, however, puzzling. Although it seems plausible that a centralised bureau in a large federation will be more difficult to control by heterogeneous voters, it does not seem realistic that production breaks down completely and officeholders use all the funds

for their own purposes, while being always re-elected. I believe that this result, which I call the *Ferejohn Paradox,* indicates a weakness of the model. In what follows, I consider some possible ways of avoiding this paradox.

2.5 Ways out of the Ferejohn Paradox, and the Principle of Subsidiarity

John Ferejohn suggested that, "if the voters agree to utilize expected *aggregate output as the criterion,* they will be able to induce the incumbent to provide the level of service as was exhibited in section" 2.2, i.e. $\sum_{n \in \mathcal{E}} \alpha_{nt} = \hat{\alpha}$, $\forall t$. See Ferejohn (1986, proposition 7). Such a criterion is sometimes called a *sociotropic voting rule* (see Fiorina, 1981). But at the same time he pointed at the collective choice problem associated with such a voting behaviour, namely that "voters will disagree among themselves as to which is the best [criterion,] and candidates for their part will try to induce voters to 'defect' from the sociotropic voting rule and vote, instead, on a distributional basis." See Ferejohn (1986, p. 22).

A second way out of the Ferejohn Paradox was suggested by Steven Slutsky. See Slutsky (1986, p. 129). He believes that in reality, for various reasons, *the incumbent's payoff is effectively an increasing function of the output.* This means that officeholders are somehow "intrinsically" motivated to act in the interest of their constituents. While this suggestion certainly has some appeal, its relevance clearly diminishes as the size of the budget is increased, since then the temptation for the officeholder to choose the short-run benefits of shirking becomes more and more obvious. Thus, the forces described in our model will still be relevant with officeholders who are partly intrinsically motivated.

While the two approaches mentioned so far both avoid the Ferejohn Paradox by changing the behavioural assumptions of the model, I now want to discuss some *institutional* changes which solve the paradox without a change in the behavioural assumptions. In the three following subsections, I consider the unanimity rule, electoral decentralisation, and the threat of separation, in turn.

2.5.1. Unanimity rule

The simplest institutional change yielding the same output as in the case of a representative voter is by replacing the majority rule with the *unanimity rule,* i.e., the incumbent is re-elected if and only if *all* the voters' cutoff levels are satisfied. Otherwise another candidate is elected.[16] This means that the definition of *N-N*-equilibrium is still relevant, where unanimity rule is represented by the special case of $M=N$ (see part (iii) of the definition). This prevents the incumbent from playing off the voters against each other (since now each voter is pivotal), and the Ferejohn paradox does not apply.

Preliminary result: *In a N-N-equilibrium under unanimity rule (i.e. M=N), it must be the case that* $\hat{g}_n > 0$ *and* $\hat{\alpha}_n > 0$ *for at least some n, i.e. the Ferejohn paradox does not hold.*

DEMONSTRATION OF THIS RESULT: Consider the problem of any single voter (indexed n), when all other voters' strategies $\{\hat{g}_t^{-n}\}_{t=0,\dots}$ are fixed. If all others would set their cutoff levels equal to zero, i.e. $\sum_{i \in \mathcal{E}\setminus\{n\}} \hat{g}_{it} = 0$, $(i \in \mathcal{E}\setminus\{n\})$, then voter n would face the same problem as if he were the only principal. He would therefore set $\hat{g}_{nt} = s$, which is the cutoff rule derived in the case of a homogeneous electorate (see equation (3)). ❑

The problem is still that there exist a very large number of equilibria. At least we can make the following characterisation.

Result 2.5: *In a strict N-N-equilibrium under unanimity rule (i.e. M=N), all voters will set their cutoff levels* $\hat{g}_n$ *such that* $\sum_{n \in \mathcal{E}} \hat{\alpha}_n = 1$.[17]

DEMONSTRATION OF RESULT 2.5: Consider the problem of voter n, for a given sequence of vectors $\{\hat{g}_t^{-n}\}_{t=0,\dots}$. In each period t, two cases are possible:

16. The problem mentioned in footnote 13 also applies to the unanimity rule, of course.

17. Strict equilibrium excludes equilibria which involve choices out of indifference. See e.g. Fudenberg and Tirole (1991, p. 444).

(a) $\sum_{i \in \mathcal{E} \setminus \{n\}} \hat{g}_{it} > \hat{\alpha} s$ or

(b) $\sum_{i \in \mathcal{E} \setminus \{n\}} \hat{g}_{it} \leq \hat{\alpha} s$,

where $\hat{\alpha}$ is derived by a similar argument as in result 2.1 and is therefore given by equation **(4)**. By the definition of $\hat{\alpha}$, case (a) implies that the incumbent will not want to stay in office. Knowing this, voter n is indifferent about his choice of $\hat{g}_{nt}$. Therefore, case (a) cannot be a strict equilibrium.[18] In case (b), voter n will set $\hat{g}_{nt}$ such as to make the incumbent's re-election constraint

$$w + \left(1 - \sum_{m \in \mathcal{E}} \hat{\alpha}_{mt}\right) s + \delta v_{t+1} \geq w + s \qquad \textbf{(19)}$$

hold with equality, which implies that

$$\hat{\alpha}_{nt} = \delta\left(1 + \frac{w}{s}\right) - \sum_{i \in \mathcal{E} \setminus \{n\}} \hat{\alpha}_{it} \qquad \textbf{(20)}$$

Because of the assumption made at the beginning of section 2.4.2, that $s/w(1) \leq \delta/(1-\delta)$ and $w(1)=Nw(N)$, for all n, equation **(20)** implies that

$$\sum_{n=1}^{N} \hat{\alpha}_{nt} = \delta\left(1 + \frac{w}{s}\right) \geq 1 \ . \qquad \textbf{(21)}$$

Taking into account the bureau's budget constraint, $\sum_{n \in \mathcal{E}} \alpha_{nt} \leq 1$, result 2.5 follows.[19] ❑

---

18. Since it is assumed here that $s/w(1) \leq \delta/(1-\delta)$, case (a) would actually also violate the office-holder's budget constraint and is therefore not feasible.

19. Notice that there exist still an infinite number of strict equilibria, since any sequence $\{(\hat{g}_t/s)\ ;(\hat{g}_t)\}_{t=0,..}$ satisfying $\sum_{n \in \mathcal{E}} \alpha_{nt} = 1$, for all t, is an equilibrium.

Since however, in practice, democratic elections also have the purpose of aggregating preferences, the unanimity rule will not be a practical solution to the problem. Therefore, I now turn to decentralisation as a way to overcome the Ferejohn paradox.

2.5.2. Electoral decentralisation I: The role of a homogeneous electorate

An alternative institutional solution to the problem is that each group of voters (e.g. residents in region *n*) elect their own officeholder so that each officeholder is elected by a homogeneous electorate. Local electorates are homogeneous in the sense that all the voters in one locality consume the same local public service. With a homogeneous electorate and under assumption $(\hat{g}MAX)$, the voters cannot be played off against each other, and electoral control works in the way described in result 2.1, even with many different voters (=regions).

Result 2.6: *Fiscal and electoral decentralisation (i.e. each region electing its own officeholder with an independent budget) restores electoral control. All the voters will set their cutoff levels* $\hat{g}_{nt}^{D} = s_n$, *in all localities n, in all periods t (where superscript D stands for decentralisation).*

This result follows directly from the reasoning developed in the previous sections (i.e. results 2.1 and 2.3). It shows that local public goods should be provided by the lowest level of government which can provide the service. This statement is closely related to the Principle of Subsidiarity, which is sometimes referred to in the traditional literature on fiscal federalism and in the political debate (see e.g. Council of Europe, 1994). Very often, this principle is used ideologically, rather than being derived from first principles. Result 2.6 shows that local government can be more productive because it is not able to play off different groups of voters against each other. If local public services are produced at the lowest level of government, the latter will face a homogeneous electorate which - under assumption $(\hat{g}MAX)$ - will exert electoral control as described in result 2.3. In contrast, higher levels of government facing a heterogeneous electorate will not be subject to effective electoral control, as shown in the Ferejohn paradox. In this sense, result 2.6 (together with result 2.4) provides a foundation for the Principle of Subsidiarity.

If a federation is already in place, breaking it up may, however, seem a radical step which will probably be associated with substantial costs. The following subsection shows how electoral decentralisation may not actually have to take place, if all parties are aware of its effects. The mere possibility of regions separating and becoming independent can be used by the voters as a threat, forcing the central government to produce a strictly positive quantity of all local public goods, and thus avoiding the Ferejohn Paradox.

2.5.3. The threat of separation

Starting from a situation where, in period *0*, a central government is already elected and in office for the whole federation of *N* regions, one can ask how much it will produce of each local public good if it has to fear that regions could leave the federation and become independent. In this subsection, I consider the case where, after each period *t*, any region may decide to quit the federation at zero cost and elect its own officeholder from period *t+1* onwards. For simplicity, a region that quits the federation is not expected to ever join it again (nor any other federation), i.e. separation is irreversible. All the regions which decide to remain members of the federation, cast their vote either for the incumbent or for the alternative candidate for federal office. If a majority of those regions votes for the incumbent, then she remains in office for the (possibly newly composed) federation in *t+1*. If she does not receive a majority, then a new central officeholder takes her place for *t+1*. Thus, the timing of events within each period is: (1) election, (2) collection of funds, (3) production, (4) separation or not.

In this subsection, I allow for a more general technology in the sense that the centre may have access to a better technology than the local governments (this may be e.g. because of economies of scale). The centre's output is described by equation **(16)**, while the local governments have only access to the inferior technologies described by

$$\hat{g}_{nt}^{D} = \gamma \cdot \alpha_{nt} \cdot s_n \quad , \tag{22}$$

where $0<\gamma<1$.

Under these circumstances, the voters will not only have to set a standard of local public services above which they are going to re-elect the incumbent, $\hat{g}_n$, but also one above

which their region is to remain a member of the federation, G_n. This leads to the following notion of equilibrium.

Definition: *(N-N-threat-equilibrium) A political equilibrium for a federation of N regions and N local public services, with the possibility for regions to separate, is a sequence* $\{ \alpha_{1t}, \alpha_{2t}, \ldots, \alpha_{Nt} ; G_{1t}, G_{2t}, \ldots, G_{Nt} ; \hat{g}_{1t}, \hat{g}_{2t}, \ldots, \hat{g}_{Nt} \}_{t=0,\ldots}$, *such that*

(i) each region uses optimal RVR1s for membership and re-election, i.e. in every period t, each region n sets two cutoff levels, G_{nt} *and* $\hat{g}_{nt}$, *taking* G_{it} *and* $\hat{g}_{it}$, *for all* $i \in \mathcal{E} \backslash n$, *for all t, as well as the incumbent's best-response pattern as given;*

(ii) in every period t, the incumbent chooses $(\alpha_{1t}, \alpha_{2t}, \ldots, \alpha_{Nt})$ *such as to maximise her inter-temporal utility, taking* $\{ G_{1t}, G_{2t}, \ldots, G_{Nt} ; \hat{g}_{1t}, \hat{g}_{2t}, \ldots, \hat{g}_{Nt} \}_{t=0,\ldots}$ *as given;*

(iii) the period t incumbent is re-elected for t+1 if and only if the requirements of a majority of M voters are satisfied, where M is a positive integer from the range $N/2 < M \leq N$. *The votes of all unsatisfied voters go to the same alternative candidate;*

(iv) in period t+1 the federation is composed of all the regions n for which $G_{nt} \leq \alpha_{nt} s$ *holds.*

Compared to the definition of *N-N*-equilibrium, parts (i) and (ii) have been extended to include each voter's choice of output requirement, G_{nt}, for remaining a member of the federation.

Result 2.7: *N-N-threat-equilibrium is given by*

$$\alpha_{nt} = \frac{G_{nt}}{s_n}, \forall n, \forall t \wedge G_{nt} = \gamma s_n, \forall n, \forall t \quad ,$$

i.e. under centralisation, as much output is produced as under decentralisation, despite the fact that the centre has access to a strictly better technology.

DEMONSTRATION OF RESULT 2.7: If a region quits the federation, its consumption of local public good will be $g_n^D = \gamma s_n$ from period *t+1* onwards. Therefore, no region will ever accept $g_{nt} < \gamma s_n$ from the centre, i.e. $G_{nt} \geq \gamma s_n$. On the other hand, any $G_{nt} > \gamma s_n$ cannot be credible (i.e. subgame perfect). Since this reasoning holds for any

n, independently of what all other regions do, the separation cutoff levels are uniquely determined, namely, $G_{nt} = \gamma s_n$, for all n, for all $t=0,...,T_n$, where $T_n \leq \infty$ is the last period in which region n is member of the federation. Next, consider $\hat{g}_{nt}$. It clearly has to satisfy $\hat{g}_{nt} \geq G_{nt}$, for all n, for all t. Finally, a reasoning similar to the one underlying the Ferejohn paradox (result 2.4) implies that $\hat{g}_{nt}$ can never be strictly greater than G_{nt}, for a majority of regions. ❑

Of course, *N-N*-threat equilibrium is characterised by the same kind of multiplicity as *N-N*-equilibrium. I.e., any vector $\hat{g}_t$ is an equilibrium, as long as (i) at least M of its components are equal to γs_n, and (ii) $G_{nt} \geq \gamma s_n$, for all n, for all t.

This result shows that the Ferejohn paradox does not hold when regions have the option to quit the federation. Centralisation leads to positive output levels without the need to actually decentralise. However, although the centre could be much more productive than local governments, it will never be strictly more productive than local governments. This can be seen as a version of the Principle of Subsidiarity.

2.6 Conclusion

On the basis of a model of repeated elections that endogenises the degree of accountability of elected officials, I have developed two arguments for decentralisation, from the point of view of the voters. The model is subject to a number of limitations, some of which can be relaxed easily. In this section, I sketch some possible extensions.

First, the crucial variable of the analysis, namely the amount of public funds s, has been taken as given. Suppose the voters choose s before the first period. Then s will be determined by a standard median voter argument. If the control of the bureau is not perfect (i.e. $\alpha<1$), the median voter will react by reducing the amount of funds he allocates to the bureau.

Second, I specified an infinite horizon model of repeated elections. Similar results could be derived with finite horizon. Finite horizon models of repeated elections have been developed e.g. by Austen-Smith and Banks (1989) and Lott and Reed (1989). The advan-

tage of the infinite horizon version is that it yields a simple steady state equilibrium, whereas the finite horizon outcome is driven by the last period effect. In the last period, it is clear that the incumbent will misallocate all the public funds. The voter's strategy has to be chosen such as to extract as much surplus as possible in earlier periods. A more thorough discussion and comparison of the cases of finite and infinite horizon can be found in Gromb (1994).

Third, the linearity of the production function is not crucial for the incentive effects. The results would be affected in an obvious way, if the technology did exhibit increasing returns to scale. The beneficial incentive effects of politicaldecentralisation would, of course, have to be traded-off against the technological disadvantage. Similarly, the linearity of the objective functions is not crucial either.

Finally, I have used a full information model in this chapter. In the next chapter, I show that the main results developed here carry over to the case of asymmetric information. Moreover, chapter 3 develops two new arguments for decentralisation which rely on asymmetric information.

Chapter 3

ELECTORAL CONTROL AND DECENTRALISATION

UNDER ASYMMETRIC INFORMATION

3.1 Introduction

In chapter 2, I described two ways in which political decentralisation may be beneficial to the taxpayers. In this chapter, I introduce asymmetric information between office-holders on the one side and taxpayers/voters on the other side. Government officials producing public services are assumed to have better information than the voters about the production technology. In section 3.2, I extend the basic model developed in section 2.2 by introducing this asymmetry of information. Results 3.1 and 3.2 characterise the political equilibrium with one bureau and a homogeneous electorate. Just as in the case of symmetric information, there exists a critical level of funds, $\hat{s}^{\alpha}$, beyond which the bureau's accountability is diminishing so that fiscal decentralisation is strictly beneficial to the voters. By limiting the size of each office to $s \leq \hat{s}^{\alpha}$ the voters can make sure that

the funds are used in their best interest. However, they cannot prevent the officeholder from receiving an *information rent.*

In section 3.3, comparing the two settings, I find that there is more scope for fiscal decentralisation under asymmetric information than under symmetric information (result 3.3).

This means that there is an intermediate range of the budget size, $\hat{s} \leq s \leq \hat{s}^{\alpha}$, for which under symmetric information the use of the funds is optimal while under asymmetric information it is suboptimal, from the voter's point of view. The reasoning underlying the Ferejohn Paradox (result 2.4) is essentially the same under asymmetric information and does not need to be repeated here. The officeholder can play off the heterogeneous voters against each other and is re-elected in every period without producing any output.

In sections 3.4 and 3.5, I develop two new arguments for political decentralisation. They both rely on asymmetric information between voters and officeholders. The first reasoning is based on the hypothesis that voters are better informed about the conditions under which a local office produces the services than they are about he centre's conditions of production. This is captured by the assumption that the variance of θ_t is smaller on the local level than on the federal level. Section 3.4 shows that, although the technology is assumed to be linear and voters are assumed to be risk-neutral, a reduction in the variance of the technological random variable, keeping its expected value constant, will increase the voters' expected utility (result 3.5). This is essentially due to the fact that a reduction in the voter's uncertainty means that in equilibrium the incumbent will try to go for re-election *more often* and, therefore, will produce public services of a satisfactory quantity (or quality) more often.

Clearly, the opposite hypothesis, namely that voters are better informed about the centre's technology, would lead to a centralisation result. Either of the two hypotheses could be justified, although the former seems somewhat more natural. I would like to emphasize again that the aim here is not to derive the conclusion that decentralisation is always optimal. The aim is rather to develop the foundations of political decentralisation by describing some advantages of decentralisation which have to be balanced with all the well-known advantages of centralisation.

Another way in which decentralisation can improve the productivity of the public sector in the case of asymmetric information is by enabling the electorate to reduce officeholders' information rents. Section 3.5 shows that under electoral decentralisation the

voters can use one period retrospective voting rules (RVR1s) which effectively set up an *electoral yardstick competition.* Some authors have described such "horizontal competition" without a formal model.[1] Within the present framework, the voters can limit the incumbents' information rents by making their re-election rules contingent not only on one office's output but on all observed output levels. To the extent that technological shocks are correlated across jurisdictions, this will enable them to induce more production from the bureau, despite the information asymmetry. Under centralised production where only one office produces all public services the voter cannot use relative performance evaluation, so that they can only achieve the outcome described in section 3.2. I give an illustration of the working of electoral yardstick competition for the special case of perfectly correlated shocks (result 3.6).

3.2 Centralised production under asymmetric information

As in chapter 2, I start with the situation where there is one (representative) voter (i.e. principal) and several, identical candidates (i.e. agents) for one office. Their time horizons are infinite with a constant and common time discount factor δ. In each period t, the voter elects one of several identical candidates to office for that period. He (the voter) lends her (the officeholder or manager) the funds s (per period) for the production of an output

$$g_t^{\alpha} = \alpha_t^{\alpha} \theta_t s \qquad (1)$$

where α^{a} denotes the proportion of funds used "efficiently" and θ is a technological variable, randomly drawn from the interval **[0,1]**. Superscript α refers to the case of asymmetric information. In every period the incumbent receives private benefits w from being in office. The voter uses a RVR1, i.e. he re-elects the period t incumbent for $t+1$ if and only if her output g_t^{α} does not fall below the critical level $\hat{g}_t^{\alpha}$ set by the voter. The new element in this chapter is that the realisation of the technological random variable θ_t is

1. See e.g. Salmon (1987).

privately observed by the incumbent at the beginning of period t. Since θ_t is not observed by the voter, his RVR1 cutoff level $\hat{g}_t^{\alpha}$ cannot be chosen contingent on it.

Results 3.1 and 3.2, below, confirm that the mechanism of electoral control as described in chapter 2 works in a similar way in the case of asymmetric information. The main difference is that, because of her private information, the incumbent will enjoy an information rent. Before stating the results, I define the new equilibrium concept.

Definition: (**1-1^a**-equilibrium) *A political equilibrium with a representative voter and with asymmetric information is a sequence* $\{\alpha_t^{\alpha}, \hat{g}_t^{\alpha}\}_{t=0,\ldots}$, *such that*

(i) the voter uses a RVR1, i.e. in every period t, the voter sets an optimal cutoff level $\hat{g}_t^{\alpha}$, *taking the incumbent's best-response pattern as given;*

(ii) in every period t, the incumbent chooses α_t^{α} *such as to maximise her inter-temporal utility, taking* $\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}$ *as given.*

This definition is essentially the same as that of **1-1**-equilibrium. The only difference is that under asymmetric information, as already mentioned, the voter cannot set his cutoff level contingent on the realisation of the random variable, since he does not observe it. Result 3.1 characterises the political equilibrium under asymmetric information.

Result 3.1: *The unique* ***1-1^a****-equilibrium is characterised as follows: In every period t, the voter requires output of at least*

$$\hat{g}_t^{\alpha} = min\left[\frac{1}{2}s, \frac{w+s}{\frac{2}{\delta}-1+\frac{1}{2}ln2}\right], \tag{2}$$

in order to re-elect the incumbent. The latter will satisfy this requirement by using a proportion

$$\alpha_t^{\alpha}(\theta_t) = min\left[\frac{1}{2\theta_t}, \frac{w+s}{\{\frac{2}{\delta}-1+\frac{1}{2}ln2\}s\theta_t}\right] \tag{3}$$

of the funds (and therefore be re-elected) if and only if $\theta_t \in [\,1/2\,,1\,]$. *Otherwise, the incumbent "walks away" with the funds, i.e. she chooses* $\alpha_t^{\alpha} = 0$.

DEMONSTRATION OF RESULT 3.1: Suppose first that the voter's sequence of cutoff levels $\{\,\hat{g}_t^{\alpha}\,\}_{t=0,\ldots}$ is given. After being elected for period t the incumbent observes θ_t and then chooses $\alpha_t{}^{\alpha}$ optimally. The officeholder's valuation of being in office in period t is

$$v_t^{\alpha}(\alpha_t^{\alpha}) \;=\; w + \left[1 - \alpha_t^{\alpha}\right] s + \delta\, v_{t+1}^{\alpha}\;, \tag{4}$$

where

$$v_{t+1}^{\alpha} \;=\; \begin{cases} w + \left[1 - \alpha_{t+1}^{\alpha}\right] s + \delta\, v_{t+2}^{\alpha} & \text{if} \quad g_t^{\alpha} \geq \hat{g}_t^{\alpha} \\ 0 & \text{if} \quad g_t^{\alpha} < \hat{g}_t^{\alpha} \end{cases} \tag{5}$$

As with symmetric information, only two levels of $\alpha_t{}^{a}$ can be optimal, depending on whether the incumbent prefers to go for re-election or not.[2] This decision, in turn, depends on the realisation of θ_t. The incumbent will want to go for re-election if and only if

$$w + \{\,1 - \hat{\alpha}_t^{\alpha}(\theta_t)\}\, s + \delta\, v_{t+1}^{\alpha} \geq w + s\;, \tag{6}$$

where $\hat{\alpha}_t^{\alpha} \equiv \hat{g}_t^{\alpha} / (s\theta_t)$ represents the minimum proportion of funds the incumbent has to use for production in order to be re-elected. This implies that the period t incumbent will set

$$\hat{\alpha}_t^{\alpha}(\theta_t) \;=\; \begin{cases} \hat{\alpha}_t^{\alpha}(\theta_t) & \text{if} \quad \theta_t \geq \hat{\theta}_t \\ 0 & \text{if} \quad \theta_t < \hat{\theta}_t \end{cases} \tag{7}$$

---

2. Notice that $v_{t+1}{}^{\alpha}$ is independent of $\hat{\alpha}_t^{\alpha}$ and $\hat{g}_t^{\alpha}$.

where

$$\hat{\theta}_t \equiv \frac{\hat{g}_t^{\alpha}}{\delta\, v_{t+1}^{\alpha}} \tag{8}$$

represents the minimum level of θ_t required to make it worth for the incumbent to go for re-election.

Given this best-response pattern for the incumbent, the voter will set $\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}$ optimally, i.e. such as to

$$\max_{\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}} U = \sum_{t=0}^{\infty} \delta^t \pi_t(\hat{\theta}_t(\hat{g}_t^{\alpha}))\, \hat{g}_t^{\alpha} \tag{9}$$

where $\pi_t(\hat{\theta}_t) \equiv Prob(\theta_t \geq \hat{\theta}_t) = 1 - \hat{\theta}_t$. From the first order conditions, using the assumption that θ_t are independent and uniformly distributed, it follows that

$$\hat{g}_t^{\alpha} = \tfrac{1}{2}\delta\, v_{t+1}^{\alpha}\ ,\ \forall t\ , \tag{10}$$

and, using the definition of $\hat{\alpha}_t^{\alpha}$, this yields

$$\hat{\alpha}_t^{\alpha}(\theta_t) = \frac{\delta v_{t+1}^{\alpha}}{2s\theta_t}\ ,\ \forall t\ . \tag{11}$$

Equations **(8)** and **(10)** together imply that the critical level of the random variable at which the incumbent will just want to satisfy the voter's requirement will be $\theta_t = 1/2$, for all *t*. This, in turn, implies that the (*ex ante*) probability of the incumbent being re-elected is $\pi_t = 1/2$ in every period, and every elected candidate can be expected to stay in office for two periods.

The stationary value of office, $v^{\alpha}=v_t^{\alpha}=v_{t+1}^{\alpha}$, under asymmetric information solves

$$v^{\alpha} = w + s + \pi(\hat{\theta}_t) \cdot \{\delta v^{\alpha} - sE\left[\hat{\alpha}_t^{\alpha}(\theta_t) \middle| \theta_t \geq \hat{\theta}\right]\} \qquad (12)$$

where $E\left[\hat{\alpha}_t^{\alpha}(\theta_t) \middle| \theta_t \geq \hat{\theta}\right]$ is the proportion of funds that an incumbent expects having to use for production in order to be re-elected, in periods where she chooses not to "walk away" with the funds. Solving **(12)**, using **(10)**, **(11)** and the fact that $\pi = 1/2$, yields[3]

$$v^{\alpha} = \frac{w+s}{1-\frac{\delta}{2}+\frac{\delta}{4}ln2} \approx \frac{w+s}{1-\frac{\delta}{3}} \quad . \qquad (13)$$

Substituting v^{α} from **(13)** into **(10)** and **(11)**, yields

$$\hat{g}^{\alpha} = \frac{w+s}{\frac{2}{\delta}-1+\frac{1}{2}ln2} \approx \frac{w+s}{\frac{2}{\delta}-\frac{2}{3}} \qquad (14)$$

and

$$\hat{\alpha}^{\alpha}(\theta_t) = \frac{w+s}{\left[\frac{2}{\delta}-1+\frac{1}{2}ln2\right]s\theta_t} \approx \frac{w+s}{\left[\frac{2}{\delta}-\frac{2}{3}\right]s\theta_t} \quad , \qquad (15)$$

respectively.

As in the case of symmetric information, the incumbent's budget constraint has to be taken into account, namely $\hat{\alpha}_t^{\alpha}(\theta_t) \leq 1, \forall\theta_t$, which is equivalent to $\hat{\alpha}_t^{\alpha}(\hat{\theta}) \leq 1$. Using **(10)** and $\hat{\theta} = 1/2$, the incumbent's budget constraint becomes

$$\frac{w}{s} \leq \frac{1}{\delta}-\frac{3}{2}+\frac{1}{4}ln2 \approx \frac{1}{\delta}-\frac{4}{3} \quad . \qquad (16)$$

---

3. In what follows, the approximation *ln2~2/3* is used to simplify some of the expressions.

Taking this constraint into account, equations **(14)** and **(15)** yield expressions **(2)** and **(3)**, respectively.[4] ❑

It is worth noticing that the optimal choice of cutoff level which solves **(9)** involves the following trade-off: On the one hand, a higher cutoff level requires the incumbent to produce more if she wants to be re-elected. But on the other hand, it implies that the probability of the incumbent actually wanting to be re-elected becomes smaller. Due to the specific setting analysed here - in particular the independently and uniformly distributed random variables θ_t - the solution to this optimization is simply expression **(10)**.[5]

3.3 Comparison of electoral control under symmetric and asymmetric information

Not surprisingly, result 3.1 shows that electoral control under asymmetric information is less effective than under symmetric information, from the voter's point of view. This is basically due to the information rent that officeholders enjoy. The information rent is reflected in a higher value of being in office v^{α}, as expressed in equation **(13)**. It can be easily checked that v^{α} is always bigger than the corresponding value derived under symmetric information, $v=w+s$.[6] This is true despite the fact that an officeholder can only expect to stay in office for two periods, on average, before losing it.

Equation **(13)** contains also the finding that an incumbent's expected utility is increasing in the discount factor δ. This means that, although the voter will increase the required

---

4. When constraint **(16)** is binding, the stationary value of being in office is $v^{\alpha}=s/\delta-(w+s/2)/(1-\delta)$.

5. As long as θ_t are independently and identically distributed on the interval [0,M] with distribution $F(\theta_t)$ and density $f(\theta_t)$, satisfying the condition that the function $[1-F(\theta_t)]/f(\theta_t)$ be monotonically decreasing, the optimal θ_t will be independent of both δ, w and s. See Ferejohn (1986, proposition 3).

6. See equation **(9)** in chapter 2.

output level $\hat{g}_t^{\alpha}$ as δ increases (see equation **(14)**), the gains resulting from the fact that with higher δ the future benefits are valued more will always dominate.

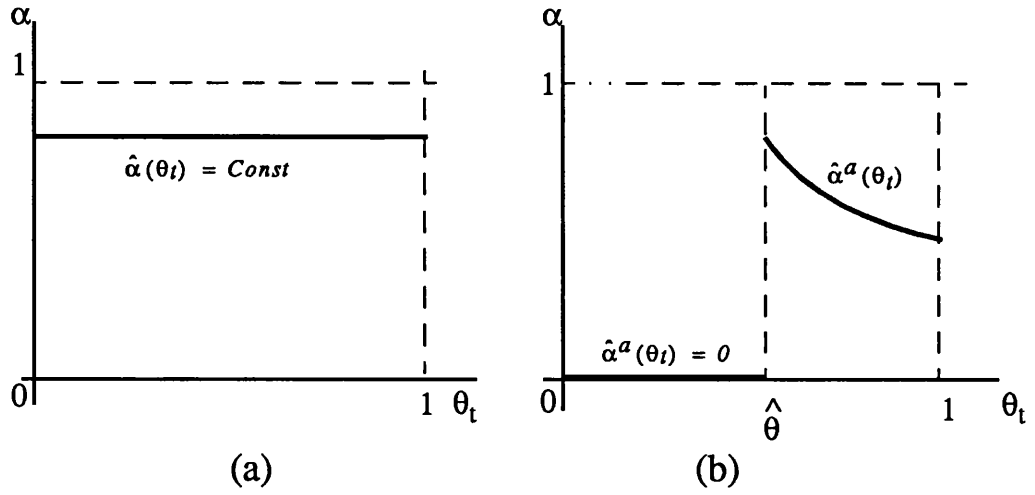

***FIGURE 3.1:** The bureau's productivity as function of* θ*:* **(a)** *symmetric information,* **(b)** *asymmetric information.*

Figure 3.1 compares $\hat{\alpha}^{\alpha}(\theta_t)$ with the corresponding function $\hat{\alpha}(\theta_t)$ derived under symmetric information. It illustrates the information rent accruing to the incumbent, i.e. it shows how the incumbent can reduce her productivity in response to the fact that the voter cannot make the RVR1 contingent on θ_t, such as to reduce the proportion of funds used productively.

Finally, from equation **(2)**, expected output per period, $\bar{g}^{\alpha} \equiv E[\hat{g}^{\alpha}]$, can be expressed as

$$\bar{g}^{\alpha} = \frac{1}{2}\hat{g}^{\alpha} = min\left[\frac{1}{4}s, \frac{w+s}{\frac{4}{\delta} - 2 + ln2}\right] . \qquad \textbf{(17)}$$

The comparative statics of expected output with respect to s shows a similar picture as under symmetric information. An increase in s is initially associated with an increase in expected output with a slope of $1/4$. Beyond a critical level $\hat{s}^{\alpha}$, however, expected output increases at a lower rate. This leads to the following decentralisation result for the case

of asymmetric information. It represents the counterpart of result 2.2 derived under symmetric information.

Result 3.2: *Fiscal decentralisation (i.e. N>1) is strictly beneficial if the total budget s is sufficiently large relative to the private benefits of office w, and the latter are not decreasing too quickly in the number of offices in the federation, i.e. if*

$$\frac{s}{w(1)} > \frac{\delta}{\frac{1}{\delta} - \frac{3}{2} + \frac{1}{4} ln2} \quad \wedge \; N\,w(N) > w(1), \;\; \exists N > 1 \quad . \tag{18}$$

DEMONSTRATION OF RESULT 3.2: Denote s_n total funds allocated to and $\bar{g}_{nt}$ expected output produced by office n. For simplicity, assume that the budget is split equally under fiscal decentralisation, i.e. $s_n=s/N$. From equation **(17)** and the assumption of i.i.d. random shocks θ_{nt}, it follows that expected total output in the federation $\sum_{n=1}^{N} \bar{g}_n^{-\alpha}$ er period, whereby $\bar{g}_n^{-\alpha}$ is equal to the expression in **(17)**, with the specification that w is now treated as a function $w(N)$ of the number of offices. Expected total output in period t will therefore o

$$\sum_{n=1}^{N} \bar{g}_{nt}^{-\alpha} = N \cdot min\left[\frac{1}{4} s_n \,,\, \frac{w(N) + s_n}{\frac{4}{\delta} - 2 - ln2} \right] \quad . \tag{19}$$

From the first condition stated in result 3.2 (and recalling result 3.1), it follows that total output under centralisation will be $g_t = \delta\theta_t[w(1)+s]$.

In the remainder of this demonstration, I will show that $Ng_{nt}>g_t$ for some $N>1$, which is equivalent to

$$N \cdot min\left[\, \theta_t s_n \,,\, \delta\theta_t(w(N) + s_n) \,\right] \geq \delta\, \theta_t(w(1) + s) \;\;, \exists\, N \quad . \tag{20}$$

Combining equation **(19)** with the first condition stated in result 3.2 we know that under centralisation expected output will be equal to $[w(1)+s]/[4/\delta-2-ln2]$. It is now obvious

that equation **(19)**, jointly with the two conditions stated in result 3.2, implies that there exists $N>1$ such that

$$\sum_{n=1}^{N} g_n^{-\alpha} > \frac{w(1)+s}{\frac{4}{\delta}-2-ln2} \quad . \quad \square \tag{21}$$

Results 2.2 and 3.2 show that, for a given level of private benefits of office, w, accruing to the incumbent, there is scope for fiscal decentralisation whenever the public funds s under the responsibility of that officeholder exceed a critical level. This critical level is $\hat{s}$ under symmetric information and $\hat{s}^{\alpha}$ under asymmetric information. The following result compares these two critical levels.

Result 3.3: *There is more scope for decentralisation under asymmetric information than under symmetric information, in the sense that* $\hat{s}^{\alpha} < \hat{s}$.

DEMONSTRATION OF RESULT 3.3: Condition **(16)** holds with equality for s equal to

$$\hat{s}^{\alpha} = \left(\frac{1}{\delta}-\frac{3}{2}+\frac{1}{4}ln2\right)^{-1} \cdot w \quad . \tag{22}$$

Condition **(12)** in chapter 2 holds with equality for s equal to

$$\hat{s} = \frac{\delta}{1-\delta} \cdot w \quad . \tag{23}$$

From this it follows that $\hat{s}^{\alpha} - \hat{s} < 0$ for all $\delta \in (0, 1]$. $\square$

The following intuition can be given for this result. Under both symmetric and asymmetric information the incumbent will not misallocate any public funds, as long as s is very small relative to w. As s is increased, her "temptation to walk away" with the funds also increases, until the point where she starts misallocating a (increasing) fraction of the funds. This point is precisely at $\hat{s}^{\alpha}$ and $\hat{s}$, respectively. Since under asymmetric infor-

mation the incumbent has an information advantage that gives her more discretion, this point is reached earlier, i.e. $\hat{s}^{\alpha} < s$.

3.4 Electoral decentralisation II: Better informed voters

It has often been argued in the literature on fiscal federalism that fiscal and political decentralisation would imply that the local authorities will be better informed about local preferences and local conditions. Therefore, their policies would serve the local needs more efficiently and more accurately. Usually, the literature does not consider the possibility that the centre could use a revelation mechanism in order to reveal the local information. Once this possibility is taken into account, it is not so clear any more why decentralisation should be superior.

Here I will make a different assumption on the information structure. I assume that with fiscal and political decentralisation *the voters* tend to be better informed about the circumstances under which the public officials provide the services. This section confirms the intuition that, under this hypothesis, electoral control will be more effective on the local level than on the federal level. This provides another justification for decentralised government. This section also clarifies further the role of asymmetric information in this model of repeated elections (see result 3.5).

The model developed in section 3.2 is now modified in order to capture a variable degree of uncertainty for the voters by varying the support of the random technological variable. θ is now assumed to be uniformly distributed on the interval $[{}^1/_2-\sigma\,,\,{}^1/_2+\sigma]$, where σ may vary between *0* and ${}^1/_2$. $\sigma={}^1/_2$ corresponds to the case analysed in the preceding sections. $\sigma=0$ corresponds to the case of symmetric information with $\theta={}^1/_2$.

Result 3.4: *Allowing for a variable degree of asymmetric information, the unique* ***1-1^a****-equilibrium is characterised as follows: For small values of* σ*, i.e.* $\sigma \leq 1/6$ *, the voter requires output of at least*

$$\hat{g}_t^{\alpha} = min\left[\left(\frac{1}{2}-\sigma\right)s, \frac{w+s}{\frac{1}{\delta\left[\frac{1}{2}-\sigma\right]} - \left[\frac{1}{2}-\sigma\right] + ln\frac{\frac{1}{2}+\sigma}{\frac{1}{2}-\sigma}}\right], \tag{24}$$

in order to re-elect the incumbent. The latter always satisfies this requirement. For larger values of σ, *i.e.* $\sigma \leq 1/6$, *the voter requires output of at least*

$$\hat{g}_t^{\alpha} = min\left[\frac{1}{2}\left[\frac{1}{2}+\sigma\right]s, \frac{w+s}{\frac{2}{\delta\left[\frac{1}{2}+\sigma\right]} - \frac{1}{2\sigma} + \frac{1}{8\sigma^2}\left[\frac{1}{2}+\sigma\right]ln2}\right], \tag{25}$$

in order to re-elect the incumbent. The latter will satisfy this requirement as long as

$$\theta_t \geq \frac{1}{2}\left[\frac{1}{2}+\sigma\right]. \tag{26}$$

Otherwise, she "walks away" with the funds, i.e. she chooses $\alpha_t^a = 0$.

DEMONSTRATION OF RESULT 3.4: Suppose that the voter's sequence of cutoff levels $\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}$ is given. Following the same argument as in section 3.2, the incumbent's best-response pattern can be summarised by

$$\hat{\theta}_t \equiv \frac{\hat{g}_t^{\alpha}}{\delta v_{t+1}^{\alpha}} \wedge \hat{\alpha}_t^{\alpha}(\theta_t) \equiv \frac{\hat{g}_t^{\alpha}}{s\theta_t} \quad , \tag{27}$$

and she will always set $\hat{\alpha}_t^{\alpha} \leq \delta v_{t+1}^{\alpha} / s$.

Given this best-response pattern and the budget constraint $\alpha^{\alpha} \leq 1$ for the incumbent, the voter will set $\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}$ optimally, i.e. such as to

$$\max_{\{\hat{g}_t^{\alpha}\}_{t=0,\ldots}} U = \sum_{t=0}^{\infty} \delta^t \cdot \pi_t(\hat{\theta}_t(\hat{g}_t^{\alpha})) \cdot \hat{g}_t^{\alpha} \quad . \tag{28}$$

There are two side-constraints to this maximisation. The first comes from the fact that the probability of re-election π_t cannot exceed *1*. This follows from the incumbent's budget constraint $\hat{\alpha}_t \leq 1$, which implies that the incumbent will not be able to satisfy $\hat{g}_t^{\alpha}$ whenever $\theta_t < \underline{\theta}_t$, where $\underline{\theta}_t \equiv \hat{g}_t^{\alpha}/s$. Thus, given $\hat{g}_t^{\alpha}$, the probability that the incumbent will actually go for re-election is

$$\pi_t = \frac{\frac{1}{2} + \sigma - max\left[\frac{1}{2} - \sigma,\ \hat{\theta}_t,\ \underline{\theta}_t\right]}{2\sigma} \quad . \tag{29}$$

In what follows the two cases $\hat{\theta}_t < \underline{\theta}_t$ and $\hat{\theta}_t > \underline{\theta}_t$ are analysed in turn. First, suppose that $\hat{\theta}_t < \underline{\theta}_t$ which is the case when $\delta v_{t+1}^{\alpha} > s$.[7] Given this, no $\hat{g}_t^{\alpha} < g^0$ can be optimal, where $g^0 \equiv (1/2 - \sigma)s$. This is because, within the range $\hat{g}_t^{\alpha} < g^0$, any increase in $\hat{g}_t$ has no effect on π_t. Therefore $\hat{g}_t$ must satisfy $\hat{g}_t^{\alpha} \geq (1/2 - \sigma)s$. Whether an increase of $\hat{g}_t^{\alpha}$ strictly above g^0 is optimal, depends on the sign of the derivative

$$\frac{\partial U}{\partial \hat{g}_t^{\alpha}} = \frac{\partial \pi_t}{\partial \underline{\theta}_t} \cdot \frac{\partial \underline{\theta}_t}{\partial \hat{g}_t^{\alpha}} \cdot \hat{g}_t^{\alpha} + \pi_t \tag{30}$$

at $\hat{g}_t^{\alpha} = g^0$. From the definition of π_t it follows that the derivative of **(30)** at $\hat{g}_t^{\alpha} = g^0$ is positive for $\sigma > 1/6$. This means that $\hat{g}_t^{\alpha}$ must solve the first-order condition to **(28)** which yields

---

7. In this case, the incumbent's individual rationality constraint **(6)** is never binding in equilibrium.

$$\hat{g}_t^{\alpha} = \frac{1}{2}\left[\frac{1}{2} + \sigma\right] s \ . \qquad (31)$$

For $\sigma \leq 1/6$, the derivative of **(30)** is not positive which means that the optimal choice of cutoff level of output for the voter is

$$\hat{g}_t^{\alpha} = \left[\frac{1}{2} - \sigma\right] s \ . \qquad (32)$$

Suppose now that $\hat{\theta}_t > \underline{\theta}_t$.[8] Given this, no $\hat{g}_t^{\alpha} < g^1$ can be optimal, where $g^1 \equiv [1/2 - \sigma]\, \delta v_{t+1}^{\alpha}$. This is because, within the range $\hat{g}_t^{\alpha} < g^1$, any increase in $\hat{g}_t^{\alpha}$ has no effect on π_t. Therefore $\hat{g}_t^{\alpha}$ must satisfy $\hat{g}_t^{\alpha} \geq [1/2 - \sigma]\, \delta v_{t+1}$. Whether an increase of $\hat{g}_t^{\alpha}$ strictly above g^1 is optimal, depends on the sign of the derivative

$$\frac{\partial U}{\partial \hat{g}_t^{\alpha}} = \frac{\partial \pi_t}{\partial \hat{\theta}_t} \cdot \frac{\partial \hat{\theta}_t}{\partial \hat{g}_t^{\alpha}} \cdot \hat{g}_t^{\alpha} + \pi_t \qquad (33)$$

at $\hat{g}_t^{\alpha} = g^1$. From the definition of π_t and the hypothesis $\hat{\theta}_t > \underline{\theta}_t$, it follows that the derivative in **(33)** is positive for all $\sigma > 1/6$. This means that $\hat{g}_t^{\alpha}$ must solve the first-order condition to **(28)** which, in this case, yields

$$\hat{g}_t^{\alpha} = \frac{1}{2}\left[\frac{1}{2} + \sigma\right] \delta v_{t+1}^{\alpha} \ . \qquad (34)$$

Solving equation **(12)** stated in section 3.2 for the stationary value of office yields in this case

8. In this case, the incumbent's budget constraint is never binding, in equilibrium.

$$v^{\alpha} = \frac{w+s}{1-\frac{\delta}{4\sigma}\left[\frac{1}{2}+\sigma\right]+\frac{\delta}{16\sigma^2}\left[\frac{1}{2}+\sigma\right]^2 ln2}\ , \tag{35}$$

and, using (34),this yields

$$\hat{g}_t^{\alpha} = \frac{w+s}{\frac{2}{\delta\left[\frac{1}{2}+\sigma\right]}-\frac{1}{2\sigma}+\frac{1}{8\sigma^2}\left[\frac{1}{2}+\sigma\right]ln2}\ . \tag{36}$$

For $\sigma \leq 1/6$, the derivative in (33) is not positive which means that the optimal choice of cutoff level of output for the voter is

$$\hat{g}_t^{\alpha} = \left[\frac{1}{2}-\sigma\right]\delta v_{t+1}^{\alpha}\ . \tag{37}$$

Solving equation (12) for the stationary value of office yields in this case

$$v^{\alpha} = \frac{w+s}{1-\delta+\delta\left[\frac{1}{2}-\sigma\right]ln\frac{\frac{1}{2}+\sigma}{\frac{1}{2}-\sigma}}\ , \tag{38}$$

and, using (37), this yields

$$\hat{g}_t^{\alpha} = \frac{w+s}{\frac{1}{\delta\left[\frac{1}{2}-\sigma\right]}-\left[\frac{1}{2}-\sigma\right]+ln\frac{\frac{1}{2}+\sigma}{\frac{1}{2}-\sigma}}\ . \tag{39}$$

Equations (31), (32), (36) and (39) yield the equations (24) and (25) which are stated in the result. ❑

On the basis of this result, a comparison of expected output levels under different degrees of asymmetric information is possible. Such a comparison is, however, complicated by

the fact that a variation of σ also entails a variation of the incumbent's budget constraint. In order to simplify, but also to separate the main point of this section from the effects underlying the decentralisation result 3.2, I assume for the following argument that δ, w and s are such that the incumbent's budget constraint is binding for any value of σ.

Result 3.5: *Assume that w/s and δ are sufficiently large for the incumbent's budget constraint to be binding for any $\sigma \in [0, 1/2]$. Then, an increase in σ reduces expected output. Therefore, if the voters' uncertainty about the technology used by the local government is smaller than about the central government's technology, i.e. $\sigma^L < \sigma^C$, then they will prefer the former to produce the public services, i.e. $\bar{g}^{aL} > \bar{g}^{aC}$.*

DEMONSTRATION OF RESULT 3.5: Since the second assertion follows straightforwardly from the first one, I only have to demonstrate that an increase in σ reduces expected output. If $\sigma \leq 1/6$ and the budget constraint is binding, then we know from result 3.4 that the output level required by the voters is $\hat{g}_t^{\alpha} = \left[\frac{1}{2} - \sigma\right] s$. We also know that the incumbent always satisfies it, i.e. $\pi=1$. This implies that expected output is

$$\bar{g}^{\alpha} \equiv \pi \, \hat{g}^{\alpha} = \left[\frac{1}{2} - \sigma\right] s \; . \tag{40}$$

If $\sigma > 1/6$ and the budget constraint is binding, then we know from result 3.4 that the output level required by the voters is $\hat{g}_t^{\alpha} = \frac{1}{2}\left[\frac{1}{2} + \sigma\right] s$, and the probability that the incumbent satisfies it is $\pi = 1/4 + 1/(8\sigma)$. This implies that expected output is

$$\bar{g}^{\alpha} \equiv \pi \, \hat{g}^{\alpha} = \frac{1}{8}\left[1 + \sigma + \frac{1}{4\sigma}\right] s \; . \tag{41}$$

It can be checked easily that in both cases the derivative $\partial \bar{g}^{\alpha} / \partial \sigma$ is strictly negative in the relevant ranges, except for $\sigma = 1/2$, where differentiating **(40)** yields *0*. Therefore, as long as the incumbent's budget constraint is binding for all σ, an increase in σ reduces expected output. ❑

3.5 Fiscal decentralisation II: Political yardstick competition

The previous sections have illustrated that under asymmetric information the agent will get a considerable information rent. In this section, I develop the argument that the voters may benefit from fiscal decentralisation, since it enables them to set up a form of yardstick competition between local governments. The theory of (economic) yardstick competition has been developed by Nalebuff and Stiglitz (1983) and Shleifer (1985). A similar reasoning has been used in the context of fiscal federalism by Breton and Scott (1978) and Salmon (1987). These authors describe both "horizontal" and "vertical" competition. The former has governments on the same hierarchical level competing, while the latter refers to competition between different governments on different levels. In this section, I illustrate how these ideas work within the model of repeated elections developed above. The voters can set up political yardstick competition by using a RVR1 that incorporates *relative performance evaluation.*

To see how electoral yardstick competition works and how it allows the voters to reduce the government's information rent, suppose that s is split in halves and two different officials are elected to produce $g^{\alpha}_{1t} = \alpha^{\alpha}_{1t}\,\theta^{\alpha}_{1t}\,\frac{s}{2}$ and $g^{\alpha}_{2t} = \alpha^{\alpha}_{2t}\,\theta^{\alpha}_{2t}\,\frac{s}{2}$, respectively.[9] In result 3.2, the gains associated with such decentralisation due to the reduction of the budget have already been emphasised. In order to focus on the effects of yardstick competition, I assume in this section that s is in a range where the boundary constraint **(16)** is binding. From result 3.3, we know that the corresponding constraint in the case of symmetric information[10] is also binding. Thus, decentralisation does not entail any gains of the type described in so far. The gains described here arise because the voters can improve incentives for the officeholder in one region by making use of information from other regions. This enables them to reduce the agents' information rents.

This is easiest to see in the case where θ_{1t} and θ_{2t} are perfectly correlated.[11] In this case the voters can actually eliminate the agents' information rents altogether.

---

9. Nothing of essence would change if $n \geq 2$ bureaus would be created.

10. See equation **(12)** in chapter 2.

11. Of course, the random variables are still assumed i.i.d. *over time.*

Result 3.6: *If θ_{1t} and θ_{2t} are perfectly correlated, then through fiscal decentralisation the voters can implement the same outcome as under symmetric information if they can buildup a reputation.*

DEMONSTRATION OF RESULT 3.6: I derive the result by constructing a particular RVR1 and showing that voters can eliminate all information rents by committing themselves to this rule.[12] It is also shown below that this rule can be sustained as a subgame-perfect equilibrium. Suppose that all voters (in both regions) commit to the following RVR1, which I will call **rule H(*b*)**:

Definition: (rule H(*b*))

If $g_k > g_l$, where $k,l=1,2$ and $k \neq l$, then incumbent k is re-elected and receives a transfer $b>0$ while incumbent l is not re-elected.

If $g_k^{\alpha}=g_l^{\alpha}>0$, then both incumbents are re-elected.

If $g_k^{\alpha}=g_l^{\alpha}=0$, then both incumbents are not re-elected.

Consider first incumbent *i*'s ($i=1,2$) decision g_i^{α} (or equivalently α_i^{α}) taking this **rule H** and g_j^{α}, $j \neq i$, as given. Since she strictly prefers to be re-elected, her best responses are as follows:[13]

$$\alpha_i^{\alpha} = \begin{cases} 1 & \text{if} \quad \alpha_j^{\alpha} = 1 \\ \alpha_j + \varepsilon & \text{if} \quad \alpha_j^{\alpha} < 1 \end{cases}, \qquad (42)$$

where ε is a small, positive number.[14] Therefore, in equilibrium we will have $\alpha_i^{\alpha}=\alpha_j^{\alpha}=1$.

---

12. It is assumed here that the two incumbents are not able to collude. In this context, the possibility of collusion would make relative performance evaluation ineffective.

13. This is for $\theta_t>0$. When $\theta_t=0$ the incumbent cannot do anything to be re-elected

14. Here I assume that ε is the smallest increment that is perceived by the voters to make a difference on their utility. This is done to avoid the existence problem (of a best-response) due to the continuous action space. e can be any real number from the open interval $(0,2\delta b/s)$.

The reasoning, however, relies implicitly on the hypothesis that the voters will stick to **rule H(*b*)**. The problem is one of time consistency (or subgame perfection): Once the output is produced, the voters will prefer not to make any transfer to an incumbent. This means that if one officeholder produces more than the other, she cannot expect the voters to pay the announced transfer *b*. This leads to an equilibrium with $g_i^{\alpha}=g_j^{\alpha}=0$. Notice that, contrary to the Ferejohn paradox, this outcome is not Pareto efficient, since whenever $\theta_t>0$, the incumbent would strictly prefer to be re-elected and would, in principle, be willing to use *all* the funds for production. The only way to sustain the efficient outcome described above, is by the incumbents using the following type of trigger strategy:[15]

Set $\alpha_{it}^{\alpha}=1$ *as long as the voters stick to* ***rule H(b)****.*

Set $\alpha_{i(t'+1)}^{\alpha}=0$ *in the period after they deviate from* ***rule H(b)*** *by not making the announced transfer.* [16]

Given this strategy, the voters clearly have an incentive to stick to **rule H(*b*)**, as long as *b* is not too large. Otherwise, the voters will not be willing to make the transfer *b*, even if they expect to receive the maximal output in period $(t'+1)$. That is, as long as

$$b<\frac{1}{c}\delta\cdot\frac{1}{2}s \qquad (43)$$

where $c>0$ denotes the opportunity cost for the voters of one unit of transfer, the voters will want to stick to **rule H(*b*)**. Therefore, if incumbents use trigger strategies of the type just described, and if condition **(43)** holds, then any **rule H(*b*)** is credible. On the other hand, *b* ought to be large enough to provide an incentive for the incumbents, i.e. for a fixed ε, it follows from footnote 18 that *b* must be bigger than $\varepsilon s/(2\delta)$. This fact, together with condition **(43)**, implies that *b* must be in the range[17]

---

15. The argument is related to the one in Chemla (1994) where managers have to build up a reputation for being "soft" in the wage bargaining with the workers in order to induce them to exert a high effort. Here, the "soft" action by the voters is to make the announced transfer at the end of the period.

16. Notice that an incumbent can never threaten to set $\alpha=0$ for more than one period, since after one period she will be out of office.

17. For this range to be well-defined, ε has to be small enough, i.e. $\varepsilon\leq\delta^2/c$.

$$\frac{\varepsilon s}{2d} < b < \frac{\delta s}{2c} \quad . \tag{44}$$

This means that, if b lies in this range, and it happens to be the case that $g_k > g_l$ then the voters in region k will make the promised transfer b in order to keep the reputation of using **rule H.**[18] This is how the same outcome as under symmetric information (which is efficient) can be sustained as a subgame perfect equilibrium under asymmetric information with relative performance evaluation. ❑

3.6 Conclusion and relation to the literature

This chapter has extended the model of chapter 2 to the case of asymmetric information. The intuition about the potential benefits of decentralisation described there carries over to this case. Moreover, two new arguments for decentralisation have been developed in this chapter. First, if voters are better informed about the local conditions of public goods provision, they will be better off under decentralisation than under centralisation. Second, the total information rents accruing to the government can be limited by the use of electoral yardstick competition which is only possible under decentralisation.

The results derived in this chapter are related to the literature on yardstick competition and comparative performance evaluation.[19] While my results confirm the basic intuition which would follow from that literature, the framework considered here is different in three respects: (i) The analysis here is carried out from the point of view of the voter(s), i.e. the principal(s), rather than being concerned with Pareto efficiency. It is clear that in this framework (with all risk-neutral individuals, a common time discount factor and linear effort costs) any outcome will be Pareto efficient. The reason for this is that the problem is essentially the division of a cake of fixed size s between all voters and all candidates for office. Since no funds are thrown away, any outcome is Pareto-efficient.

---

18. Since voters commit to their re-election rule at the beginning of each period, they can choose b close to $\varepsilon s/2\delta$. But, in equilibrium, they will never actually have to make the transfer and the choice of b does not affect their utility as long as it is in the range defined by **(44)**.

19. Major contributions to that literature include Lazear and Rosen (1981), Holmstrom (1982), Green and Stokey (1983), Nalebuff and Stiglitz (1983), Mookherjee (1984), Shleifer (1985) and others.

The question asked here is: How does yardstick competition affect the size of the cake accruing to the voters?

(ii) My analysis differs from the standard literature on yardstick competition also in that I use an incomplete contract framework. Under the simplifying assumptions made here, there would be no role for relative performance evaluation if complete contracts were feasible.

(iii) The problem analysed here is one of repeated elections with infinite horizon. Contrary to the standard literature on yardstick competition, it is thus dynamic. The problem, however, is stationary and does not raise any of the inherently dynamic issues considered in some recent contributions on yardstick competition, such as Holmstrom (1982b), Gibbons and Murphy (1992), Meyer (1995) or Meyer and Vickers (1994).

The latter is an interesting paper, since it contains both incomplete contracts and dynamic considerations. The authors show that comparative performance information may not always be useful when contracts are incomplete. They identify two sources of inefficiency under yardstick competition: Comparative performance evaluation may worsen either the ratchet effect or the reputation effect in a dynamic agency relationship. The fact that these two effects do not appear in my model is part of the explanation for the different results.[20] Another part of the explanation is that my analysis is focused on the principal's payoff instead of Pareto efficiency.

I do not claim, of course, that these two effects do not play a role in the real world. But I abstract from them - as I do from all other arguments which work in favour of centralisation - because my aim is here only to identify the benefits of decentralisation. As I argue in chapter 1, this first step in the direction of a more realistic cost-benefit analysis of decentralisation is not properly founded in most of the literature on fiscal federalism. This is the reason why my analysis is completely one-sided. It is only in a next step that one can incorporate the arguments developed here in a more complete analysis which evaluates the trade-off between the costs and benefits of decentralisation.

---

20. The reputation effect is ruled out by the assumption that θ_t is i.i.d. over time. The ratchet effect is not relevant because s and θ are exogenous.

PART TWO

Chapter 4

CLASS FORMATION AND REFORM

4.1 Introduction and relation to the literature

Societies, at times, face fundamental institutional reforms. The changes in a number of former communist countries which are being transformed into market economies are recent examples. Some workers change their jobs, others become self-employed or unemployed, others still become employers, entrepreneurs or capitalists. New firms are set up, others are restructured, privatised or closed down. Clearly, such fundamental reforms are associated with extraordinary risks on both individual and macroeconomic levels. Many developing countries also face similar, deep structural reforms involving the transition from more traditional, bureaucratic or collectivist to more modern, competitive or market-based institutions.

This chapter attempts to shed some light on the way economic and political institutions interact during the transition from a collectivist to a competitive economy. In particular, I am interested in the following questions: How are the risks associated with transition allocated? What types of individuals will tend to move to the new, competitive sector? Will there be enough entrepreneurs willing to take the necessary risks such that the mar-

ket sector can develop? Under what conditions can transition be politically smooth and irreversible? Under what conditions can it happen that a majority of the population suffers from the transition?

Some of these are old questions that have been dealt with in the early literature on economic development. In his very influential work, Lewis (1954) described the process of development as a gradual transition from a stagnating traditional economy with surplus labour and inefficient use of resources to a modern, efficient and growing economy, through a stage of dualism (with both a traditional and a modern sector). Following Lewis, a large part of the literature on economic development has viewed the process of development in a similar light.[1] Implicitly or explicitly, the process of sectorial migration and transition is seen as a desirable development that needs to be fostered through appropriate policies and which is essentially irreversible.

A more balanced perspective on the desirability of such a transition is reached from Sen's (1966) analysis of production in a dual economy. He gave several reasons why surplus labour, and a wage gap between the traditional and the modern sectors can exist in equilibrium. He also showed that the traditional sector may be more efficient than the modern sector, if the latter is characterised by labour market imperfections.

In this chapter I consider the same type of dual economy as Lewis. Individuals can work either in the collectivist sector (which is characterised by institutionally set wages and surplus labour) or in the competitive sector (within wich a class structure with entrepreneurs and workers emerges and incomes are determined in competitive markets). Within this framework I argue, similarly to Sen, that either of the two systems may dominate the other in terms of welfare. The welfare criteria I consider are the Pareto criterion and the majority rule.

Another related branch of the literature studies rural-urban migration in the presence of urban wage rigidities and unemployment. In their celebrated articles Todaro (1969), and Harris and Todaro (1971) endogenised the sectorial choice of workers who migrate to the sector that yields the highest expected wage for them, balancing the higher earnings opportunities against the higher risks in the city. Two crucial assumptions underlie the Harris-Todaro model: First, all workers are identical with respect to both their skills and

1. See Kuznets (1955), Rostow (1960), Barber (1961), Enke (1962), Fei and Ranis (1964), Todaro (1969), Harris and Todaro (1971), among others.

their reservation wage, and, second, the traditional sector is taken to be intrinsically inefficient and only persists because of a wage rigidity in the modern sector.[2] The model presented below, does not build on these assumptions.

More recently, Banerjee and Newman (1995) take a different approach to migration. In their model, workers differ in their initial wealth. As in the Harris-Todaro model, the production technology is assumed to be more efficient in the urban sector than in the rural sector. In the latter, in turn, credit markets work more efficiently because of the absence of asymmetric information, there. Urban credit market imperfections act as a barrier to migration, since they raise the cost of moving to the urban sector. Banerjee and Newman show that this barrier is particularly high for the rural middle class who may be reluctant to migrate, although this would be socially optimal. The contribution of Banerjee and Newman (1995) is one in a series of recent papers recognising that, despite their productive inefficiencies, traditional institutions may persist because they are more efficient in some other ways.[3] In this paper, I join this line of reasoning by emphasising the comparative advantage of collectivist institutions for social insurance. In contrast to Banerjee and Newman, I derive all differences between the two sectors only from institutional differences - the economic fundamentals remain identical in both sectors.

Motivated by the recent historical events, economists began to study the transition process from a communist to a market economy. Many of the arguments from the 1950's development literature are reappearing now in the context of transition economies. The state-owned (traditional) sector is seen as inefficient, and the benefits of the transition to a market-based (modern) system are usually emphasised. Then, for many authors, transition appears as widely desirable and necessary.[4] In practice, however, in many countries such a transition has now turned out to be more difficult than predicted. The development of the private sector has been progressing very slowly. A large number of people have been suffering rather than benefiting from the reforms. On a political level, in some countries the reforming governments were voted out of office. Consequently, interest in the

2. The Harris-Todaro model has been criticised and extended in different ways. See e.g., Fields (1975, 1989). Standing (1978) argues that reservation wages are higher for the workers who are already employed in the urban sector than for the rural workers contemplating migration. This implies, on the one hand, that a class of long-term urban unemployed will emerge and, on the other hand, that the traditional sector will not persist in the long-run.

3. See also Stiglitz (1990), Varian (1990), Xu (1993) and Townsend (1994), among others.

4. See e.g. Journal of Economic Perspectives (1991), Sachs (1991), Dewatripont and Roland (1992).

costs of transition has now emerged both in the political debate and in the academic literature.[5]

In this chapter, I hope to shed some light on the issues of development and transition by pointing out what I believe to be a fundamental trade-off between allocative efficiency and social insurance. In the model developed below the traditional economy is taken to be a collectivist system in which workers contribute to production and are rewarded independently of their abilities, their talents and their marginal productivities. All production units (viz. firms), whether highly productive or not, employ the same number of workers, and all output is collected and shared (equally) among all the workers and managers. *Since production cannot be observed, such a system involves clearly substantial inefficiencies due to the misallocation of talent (managers are not more talented than the average worker) and labour (its marginal product is not equalised across productive units).* On the other hand, the traditional system provides (social) insurance at a lower cost than is possible in the modern economy.[6] No-one bears risk; both individual ability and plant productivity risks are perfectly pooled in the collectivist society.

In contrast to this, the modern economy is much more flexible. It uses the same resources more productively by (partially) revealing the abilities of individual labourers and the productivities of individual firms. More productive firms make more profits, and more able individuals tend to earn higher incomes. Moreover, the labour market is flexible and workers are allocated such as to equate their marginal product across firms. Following a series of papers by Kihlstrom and Laffont (1979, 1983), Kanbur (1979a, 1979b, 1981), Newman (1991), Banerjee and Newman (1993) and Laussel and Le Breton (1995), I take a general equilibrium approach to the process of firm creation and the emergence of entrepreneurial and working classes.[7] My model of the competitive economy extends Kanbur s (1979a) work by introducing individuals of different entrepreneurial ability and considering the allocation of such ability (or talent).[8] Knowing their type and their risk

5. See e.g. Aghion and Blanchard (1994), Sachs and Woo (1994) or Rasier (1995).

6. I assume that there is no aggregate risk. With aggregate risk, egalitarian risk-sharing in the collectivist system is not fully efficient if workers have different attitudes towards risk. Efficient risk-sharing would require the less risk averse individuals to bear more risk than the more risk averse individuals.

7. Kihlstrom and Laffont follow the Knightian view of the entrepreneur as a risk-taking owner-manager; see Knight (1921). Entrepreneurs bear the risk of being more or less successful in their business. Workers bear no risk. This view can be traced back to Cantillon (1755) and, as Kanbur (1980) argues, is very different from Schumpeter's (1934, 1936, 1954) view of the entrepreneur as an innovator.

aversion, some individuals will choose to become labourers, others entrepreneurs. *Since production cannot be observed, no insurance or stock markets exist and entrepreneurs have to bear the risk of business failure.* Moreover, there is an income risk associated with an individual s entrepreneurial ability. Such *risk-bearing represents the cost of modernity* which has to be weighed against the benefit of increased productivity.[9]

If a majority of individuals expect to be better off in the modern economy, then, starting from a traditional economy, they can vote for a total reform. This would mean an immediate closure of all traditional production activity and a mass "migration" to the private sector. In reality such a "big bang" reform may be neither feasible nor desirable because of the associated costs of restructuring. Here I abstract from all such costs, except for those associated with risk-taking. Even so, I show that a majority may suffer from such a "big bang" reform and regret it.

If a more gradual approach to reform is taken, no-one is forced to leave the traditional sector. In this case, the outcome will be one of the following: (i) It may be that the negative externality of those choosing to migrate on those initially not wishing to leave the collectivist sector is so large that the latter will be induced to also move to the modern sector and, eventually, everyone will end up there. This could be interpreted as a "spontaneous big bang", as opposed to the "political big bang" reform described above. (ii) If the externality is weaker, then a gradual reform process will lead to an equilibrium with a dual economy, where some individuals remain in the traditional sector. In both these outcomes, again, a majority of the individuals may regret the changes.[10]

Within this framework, I also develop several results on the political economy of transition. I show, e.g., that either gradual or total reform or total reform may be chosen by a

---

8. Laussel and Le Breton (1995) also introduce different abilities to the model of Kihlstrom and Laffont (1979). They analyse the allocation of talent when all individuals are risk neutral. Here, individuals differ both in their risk attitudes and their managerial abilities (they are, however, identical as workers). The allocation of talent is studied within a different general equilibrium framework by Murphy, Shleifer and Vishny (1991) and Acemoglu (1993).

9. As mentioned above, the market incompleteness is derived from the non-observability of output. Some authors have made the point that even if output is non-observable, a second-best insurance on the basis of inputs is feasible, if some inputs are observable and the production function is common knowledge; see Grossman, Hart and Maskin (1983). But even if some insurance is available for entrepreneurs, full insurance is not optimal, in general, and all the trade-offs described here remain relevant.

10. For an empirical study of the impact of out-migration on the traditional sector and on the modern sector in a number of developing countries, see Hathaway (1964).

majority, even if it makes a majority worse off relative to no reform. I also show that, if society faces the democratic choice between gradual reform, total reform and no reform, there will always be a political polarisation. I.e., gradual reform will never be the most preferred alternative under majority voting. The relation of these results to the findings of Dewatripont and Roland (1992a and b) are discussed in section 4.8. My results are consistent with some of the recent developments in Eastern Europe and the Former Soviet Union, where moderate reformers have essentially lost to either the radical reformers or to the conservatives. Only in countries like e.g. China or Vietnam, where the option of total reform was ruled out dictatorially, more moderate reforms have been possible.

The remainder of this chapter is structured as follows. After introducing the economic fundamentals of the society, which remain the same in all systems (in section 4.2), and characterising the first-best allocations (in section 4.3), I characterise equilibrium in the collectivist and in the competitive economies (in sections 4.4 and 4.5), respectively. Section 4.6 describes total reform, whereas section 4.7 is devoted to the analysis of gradual reform and dual economy. Section 4.8 contains the main conclusions, compares them with some conflicting results in the literature and sketches some possible extensions.

4.2 The fundamentals of the economy

Individual preferences, abilities and technologies are given exogenously and are independent of the economic system. The differences in the performance of different systems will thus be due to the different institutions associated with each economic system. In this section, I present the fundamental elements of the economy that are invariant across the systems.

There are two types of individuals: talented ones of high managerial ability and untalented ones of low ability. Talented individuals are characterised by a higher success rate in managing production (see below). As workers, all individuals are identical. There is a continuum of individuals, a proportion α of which are talented and $1-\alpha$ are untalented. I assume that α is common knowledge, so that there is no aggregate risk. Individual preferences are represented by a continuous and differentiable *certainty equivalent function* of the form

$$v = v(y, \pi, R) \; ; \qquad (1)$$

whereby y denotes income earned with probability π, and R is an index of risk aversion.[11] R can take any value from the interval $[R^-,R^+]$, with $R^- \geq 0$ and $R^- < R^+$. The certainty equivalent function **(1)** satisfies the following set of (rather weak) assumptions:

(A) $\qquad v_\pi > 0 \; ; v_y > 0 \; ; v_R > 0 ;$

(B) $\qquad v(y, 1, .) = y \quad ; \quad v(y, \pi, 0) = \pi y \; ;$

(C) $\qquad v(., ., \infty) = 0 \; .$

I further suppose that abilities are uncorrelated across individuals. This is captured by the assumption that $P_H(R) \equiv P_L(R) \equiv P(R)$, where $P_H(R)$ and $P_L(R)$ denote the cumulative distributions of risk attitudes in the populations of high and low ability individuals, respectively.

Each individual is endowed with one, indivisible unit of labour time. To set up and operate a technology, a fixed investment of one unit of (managerial) labour time must be sunk. The proportion of managers is denoted ε. Once this has happened, the production unit turns out to be either highly productive or not so. The probability of the unit being highly productive is π_H if the manager is of high ability and π_L if the manager is of low ability, with $\pi_H > \pi_L$. Whatever the institution considered, a plant's output is not verifiable by outsiders.[12] For simplicity, I assume that the less productive units are not viable, or unproductive.[13] If the technology turns out to be viable, it is of the form

$$g(\lambda) = \lambda^\gamma \; ; \; 1 > \gamma > 0 \; , \qquad (2)$$

where g denotes output as a function of λ, the amount of labour time (i.e., the proportion of individuals) being employed in that unit along with the manager.

---

11. With probability $1-\pi$ income is 0.

12. Although it cannot be stolen.

13. I could have assumed a continuum of possible technological realisations, as in Kihlstrom and Laffont (1979) and Kanbur (1979a), without affecting the main argument.

In this model, individuals face two types of risk: From an *ex ante* position they face an ability risk; from an *interim* position (i.e., after having learned their ability) they still face a business risk. An important question in this chapter is: "How do different economic systems deal with these risks?"

4.3 The first-best allocations

Given these fundamentals, society could in principle maximise total output while, at the same time, providing full insurance for its members against both business and ability risks. The first-best benchmark considered here corresponds to the resource allocation that would be chosen by a social planner who had perfect knowledge of individual abilities and risk attitudes, but did not know in advance the identity of the lucky or unlucky entrepreneurs. This means that even the social planner is not able to avoid some managerial labour time being sunk in unproductive projects. In this sense, the notion of first-best considered here corresponds to interim efficiency.

Clearly, first-best requires that workers are allocated to occupations according to their abilities and to production units according to the plants' productivities, such as to maximise total output. Notice, however, that since a worker's ability is not publicly known, the first-best will, in general, not be feasible in the collectivist nor in the competitive economy. Denote ε_H the proportion of talented (high ability) managers in the total population, and ε_L the proportion of untalented (low ability) managers in the total population. Result 4.1 characterises the first-best:

Result 4.1: *In the first-best, workers are allocated to production units such as to equalise their marginal products. The optimal mass of managers is decreasing in* γ*, i.e., the more strongly returns to scale are decreasing the more managers are needed. The precise characterisation of the first-best is different in either of the three following parameter ranges:*

Regime I: $\gamma \geq 1-\alpha$.

Regime II: $1-\alpha > \gamma > [(1-\alpha)\pi_L] / [(1-\alpha)\pi_L + \alpha\pi_H]$.

Regime III: $[(1-\alpha)\pi_L] / [(1-\alpha)\pi_L + \alpha\pi_H] \geq \gamma$.

DEMONSTRATION OF RESULT 4.1: In the first-best, individuals are fully insured and total output is maximised. It is easy to show that all first-best outcomes only differ concerning the *distribution* of consumption. The production function is strictly concave in non-managerial labour input. It follows that every production unit should employ the same amount of labour such that the marginal product of labour is equalised across production units. Non-managerial labour will only be used in viable production units, of which there are $\Omega \equiv \pi_H \varepsilon_H + \pi_L \varepsilon_L$. Total output is then given by

$$G(\varepsilon_H, \varepsilon_L) = \left[\pi_H \varepsilon_H + \pi_L \varepsilon_L\right] \cdot \left[\frac{1 - \varepsilon_H - \varepsilon_L}{\pi_H \varepsilon_H + \pi_L \varepsilon_L}\right]^{\gamma} . \qquad (3)$$

Talented individuals are on average more productive managers ($\pi_H > \pi_L$), but all individuals are equally productive workers. It follows that in the first-best all managerial tasks should be performed by talented individuals, if there are enough of them. This is because, starting from *(0,0)*, an increase in ε_H is more effective than an increase in ε_L. Therefore ε_H is increased up to the interior optimum $(\varepsilon_{HFB}, 0)$. However, if the constraint $\varepsilon_H \leq \alpha$ a becomes binding, then $\varepsilon_{HFB} = \alpha$ and $\varepsilon_{LFB} \geq 0$. Whether $\varepsilon_{LFB} > 0$ depends on the sign of $\partial G(\alpha, 0)/\partial \varepsilon_L > 0$.

Maximising **(3)** with respect to ε_H, at $\varepsilon_L = 0$, yields from the first-order condition $\varepsilon_{HFB} = 1 - \gamma$. Since ε_H cannot exceed α, this is only feasible for $\gamma \geq 1 - \alpha$. Substituting $\varepsilon_H = 1 - \gamma$ and $\varepsilon_L = 0$ into **(3)**, yields a total output in <u>*regime I*</u> of

$$G_{FBI} = \gamma^{\gamma} \cdot (1-\gamma)^{1-\gamma} \cdot \pi_H^{1-\gamma} . \qquad (4)$$

If $\gamma < 1 - \alpha$, two other regimes are possible: As long as

$$1 - \alpha > \gamma > [(1-\alpha)\pi_L] / [\pi_H \alpha + \pi_L (1-\alpha)] \quad (\underline{\textit{regime II}}),$$

it is not optimal to employ untalented individuals as managers, i.e., $\varepsilon_{LFB} = 0$ and $\varepsilon_{HFB} = \alpha$. Total output in *regime II* is given by

$$G_{FBII} = (1-\alpha)^{\gamma} \cdot \alpha^{1-\gamma} \cdot \pi_H^{1-\gamma} . \qquad (5)$$

Only if the production function is very strongly concave with $\gamma < [(1-\alpha)\pi_L] / [\pi_H \alpha + \pi_L (1-\alpha)]$ (*regime III*), it becomes optimal to allocate also some untalented individuals to managerial tasks. In this regime, we have $\varepsilon_{HFB}=\alpha$ and $\varepsilon_{LFB} = 1-\alpha-\gamma[1+\alpha(\pi_H/\pi_{HL}-1)]$, which follows from maximising **(3)** with respect to ε_L, at $\varepsilon_H=\alpha$.[14] Substituting this into **(3)** yields a total output in *regime III* of

$$G_{FBIII} = \gamma^{\gamma} \cdot A^{\gamma} \cdot \left[\pi_H \alpha + \pi_L (1-\alpha-\gamma A)\right]^{1-\gamma} , \qquad (6)$$

where $A=1+\alpha(\pi_H/\pi_L-1)$. ❑

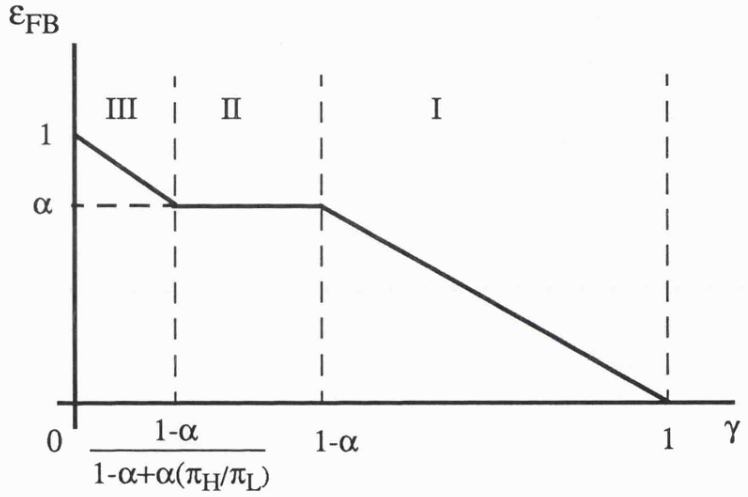

FIGURE 4.1: *The proportion of managers in the first-best.*

Figure 4.1 shows how the optimal proportion of managers, $\varepsilon_{FB} \equiv \varepsilon_{HFB} + \varepsilon_{LFB}$ depends on γ in the three regimes. If γ is large (*regime I*), only talented individuals are allocated to managerial jobs up to the point where all talented individuals work as managers. In an intermediate range of γ (*regime II*), all talented individuals work as managers and all

14. The intuition for this is as follows: For high γ ($\geq 1-\alpha$), the technology is close to constant returns to scale and the number of production units can be kept small in order to save on set-up costs. For low γ ($<1-\alpha$), returns to scale are more strongly diminishing and it is optimal to have more than $1-\gamma$ production units. But the economy runs out of talented managers. Any additional production units have to be run by untalented managers. This is, however, only optimal if γ becomes very small.

untalented as workers. Only if returns to scale are very strongly decreasing (small γ, *regime III*) will also some untalented individuals be allocated to managerial tasks.

4.4 The collectivist economy

I model the traditional economy as a collectivist system (denoted C) that has some similarities with the traditional sector described in Lewis (1954 and 1958) and Fei and Ranis (1964). The traditional sector in the Lewis-Fei-Ranis model, which was intended as a description of an underdeveloped rural area, has the following main characteristics: There is no competitive labour market; workers are paid "institutionally" fixed wages that are independent of (and actually above) the marginal product of labour. The sector is also characterised by surplus labour, i.e., labour that could be withdrawn without reducing total output.

I would like to argue here that this description is more convincing as a description of a collectivist economy where no competitive labour markets operate. My point is that an economy can have a competitive labour market, although it may be underdeveloped, or poor, or agrarian. At the same time an economy with the characteristics of a Lewis-Fei-Ranis traditional economy need not necessarily be poor, nor agrarian. Here, I propose the following view of the collectivist economy:[15] Although the output of one particular firm is not observable by outsiders, it cannot be stolen by insiders (assume for instance that the quality can be either acceptable or zero so that physical quantities cannot escape). After production has taken place, it is pooled and each worker receives an income which is independent of both his ability and his marginal product. In fact, he receives a fixed (e.g. an equal) share of total output. Individuals are allocated randomly to either managerial or productive jobs and to firms. The only variable that is chosen optimally in C is the number of production units that are set up, i.e., the proportion of managers in the population, ε_C. It is chosen by an unspecified authority (which may be thought of as either "tradition" or a "village council" or a "planning bureau") such as to maximise total output, *given* the institutional features just described.

15. The following assumptions are e.g. consistent with the description and the findings on the pre-reform Chinese economy, as reported in White (1982, 1987 and 1988).

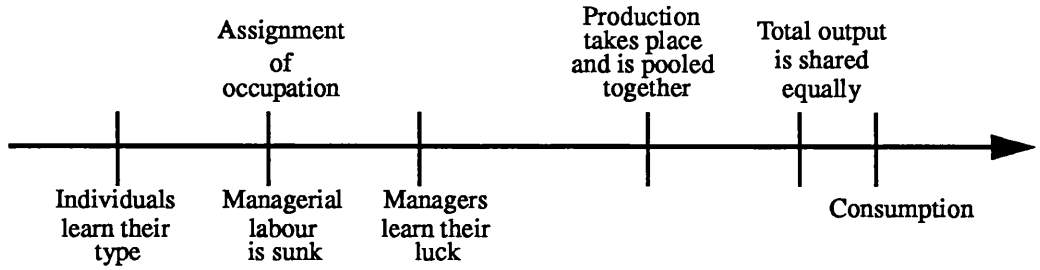

FIGURE 4.2: The sequence of events in a collectvist economy.

Since the output of an individual production unit cannot be observed, no manager has an incentive to reveal her type.[16] Then occupations are allocated randomly and the ratio of talented to untalented managers will be $\alpha/(1-\alpha)$. This means that if ε_C production units are set up, then $[\pi_H\alpha+\pi_L(1-\alpha)]\varepsilon_C$ production units will be viable. The random allocation process also implies that each production unit employs an equal share of the labour force irrespective of whether it is viable or not. Therefore, labour employed in non-viable production units represents *surplus labour*, in the sense that its withdrawal would not reduce output.[17] The sequence of events is summarised in figure 4.2. Total output in C is then given by

$$G(\varepsilon_C) = \left[\pi_H\,\alpha + \pi_L\,(1-\alpha)\right] \cdot \varepsilon_C \cdot \left[\frac{1-\varepsilon_C}{\varepsilon_C}\right]^{\gamma} . \tag{7}$$

Maximising (7) with respect to ε_C yields from the first-order condition

$$\varepsilon_C = 1-\gamma \tag{8}$$

Substituting (8) into (7), yields a total output in C of

16. Assume there is a small benefit of being a manager.

17. Sometimes the term "disguised unemployment" is used instead of surplus labour. The concept and some of the traditional theories of surplus labour are discussed in Kao et al. (1964). See also Lewis (1954) and Sen (1966).

$$G_C = \left[\pi_H \alpha + \pi_L (1-\alpha)\right] (1-\gamma)^{1-\gamma} \gamma^{\gamma} \ . \qquad (9)$$

As the following result shows, more production units are set up in C than in the first-best, but since many of them are not viable, total output is lower.

Result 4.2: *In the collectivist economy less output is produced than in the first-best with either the same number of production units (in regime I) or with more production units (in regimes II and III), i.e.,* $G_{FB}>G_C$, $\varepsilon_{FBI}=\varepsilon_C$, $\varepsilon_{FBII}<\varepsilon_C$ and $\varepsilon_{FBIII}<\varepsilon_C$.

DEMONSTRATION OF RESULT 4.2: In section 4.3, the proportion of managers in the first-best in *regime I* was shown to be $\varepsilon_{FBI}=1-\gamma$. Comparing this with equation **(8)** shows that, in *regime I*, the mass of managers is the same in both the collectivist and the first-best economies. For *regimes II* and *III*, similar comparisons yields that more production units are set up in C, i.e.,

$$\varepsilon_C = 1-\gamma > \varepsilon_{FBII} = \alpha \quad \text{and} \quad \varepsilon_C = 1-\gamma > \varepsilon_{FBIII} = 1-\gamma[1+\alpha(\pi_H/\pi_L-1)].$$

Turning to output, a comparison of equations (4) and (9), shows that the first-best output in *regime I* exceeds total output in C. Similarly, a comparison of **(5)** and **(6)** with **(9)** yields the same finding for *regimes II* and *III*, respectively. ❑

Due to imperfect information, the allocation of talent will in general not be optimal, neither in the collectivist nor in the competitive economy. The collectivist institutions offer maximal insurance to individuals at the expense of productive efficiency. Managers have no incentive to reveal their luck so that some units keep employing labour, although they are unproductive. As the following section shows, the competitive institutions can partly solve this problem but at the cost of risk bearing and consequently insufficient entrepreneurship.

4.5 The competitive economy

The model of the modern economy is based on Kanbur (1979a) with the difference that here the population is heterogeneous regarding managerial ability (or talent).[18] As in C, individuals in the competitive sector first learn their ability which may be either high or low. Second, each individual chooses an occupation and becomes either entrepreneur or

labourer. Then, entrepreneurs set up a firm and discover whether it is viable or not. If it is, they can hire labourers and produce. Entrepreneurs with non-viable firms are not able to produce and are left with zero income.[19] Individuals who choose to become workers, supply their labour on a perfectly competitive market and earn a sure wage. Here the implication of non-observability of output and employment by outsiders is that no insurance market can develop. Figure 4.3 summarises the timing.

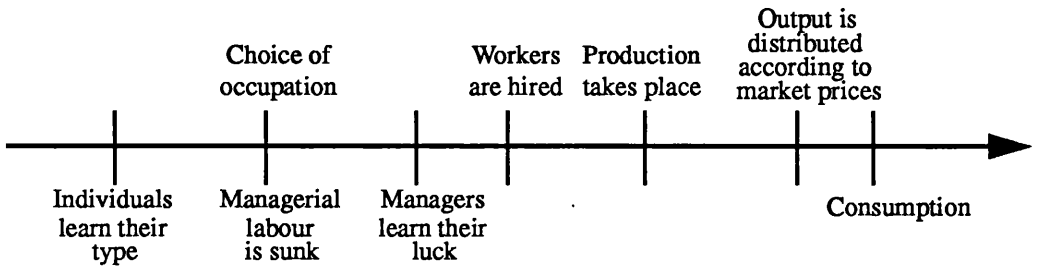

FIGURE 4.3: The sequence of events in the competitive economy.

Given these assumptions, equilibrium in the competitive economy is defined as follows:

Definition 4.1: *Competitive Equilibrium (CE) is a wage rate* w_{CE} *and partitions of both the sets of talented and untalented individuals in workers* Δ_T *and managers (or entrepreneurs)* Γ_T*, for T=H,L, such that:*

(i) each entrepreneur demands λ_{CE} *labourers such as to maximise profits, whereby she takes the wage rate as given, i.e.,*

$$\lambda_{CE} = arg\ max\left[\lambda^{\alpha} - w_{CE}\lambda\right]; \tag{10}$$

(ii) labour supply equals labour demand, i.e.,

$$1 - \varepsilon_{HCE} - \varepsilon_{LCE} = \left[\pi_H\ \varepsilon_{HCE} + \pi_L\ \varepsilon_{LCE}\right]\lambda_{CE}\ , \tag{11}$$

---

18. Sometimes the term "disguised unemployment" is used instead of surplus labour. The concept and some of the traditional theories of surplus labour are discussed in Kao et al. (1964). See also Lewis (1954) and Sen (1966).

19. Notice that they have sunk their labour endowment and cannot work elsewhere.

where ε_{TCE} (for T=H,L) denote the proportions of type T individuals in the total population choosing to be entrepreneurs;

(iii) individuals (indexed k) choose their occupation such as to maximise their expected utility, i.e., it must be the case that for T=H,L:

$$w_{CE} < v\left[r_{CE}, \pi_T, R(k)\right] \Rightarrow k \in \Delta_T \ , \tag{12}$$

and

$$w_{CE} > v\left[r_{CE}, \pi_T, R(k)\right] \Rightarrow k \in \Gamma_T \ , \tag{13}$$

where r_{CE} denotes entrepreneurial income in CE.

Notice that, as in Kanbur (1979a, 1979b, 1981) or Khilstrom and Laffont (1979), no insurance market exists in this economy. This is, of course, a crucial assumption since a competitive equilibrium with complete markets would clearly always be first-best.[20] The assumption of incomplete insurance markets seems, however, reasonable on both theoretical and empirical grounds. Empirically, it can be viewed as a stylised fact that insurance markets tend to be among the last ones to be developed, even in economies with a long-standing history of market-based development. In transition economies, it is easy to argue that insurance markets are probably the most incomplete ones. Theoretically, the incompleteness of insurance markets can be justified within the present framework by the asymmetry of information concerning both individual ability and firm productivity. Risk-bearing is the only device that induces individuals to sort out themselves. Even if insurance markets would develop, they would necessarily be incomplete (since observability is never perfect and monitoring is costly), and the basic trade-offs described in this paper would still be relevant. We can now proceed to characterise competitive equilibrium.

Result 4.3: *CE exists and is unique. The more risk averse individuals of each type, will choose to become labourers, while the less risk averse will choose to become entrepreneurs.*

DEMONSTRATION OF RESULT 4.3: If an agent with risk aversion R and ability T prefers to become entrepreneur, then from part (iii) of definition 4.1 and assumption (A), it must

20. See Kihlstrom and Laffont (1983).

be the case that all agents with ability T and risk aversion less than R will also prefer to become entrepreneurs. This holds for $T=H,L$. Denote $\hat{R}_H$ and $\hat{R}_L$ the risk aversion indices of the indifferent individuals of high and low managerial ability, respectively. Using part (iii) of definition 4.1, $\hat{R}_H$ and $\hat{R}_L$ must solve

$$w_{CE} = v\left(\lambda_{CE}^{\gamma} - w_{CE}\,\lambda_{CE}, \pi_T, \hat{R}_T\right), \text{ for } T=H,L. \tag{14}$$

(Notice that $\hat{R}_T$ may be outside the support of R, $[R^-,R^+]$.) From part (i) of definition 4.1, it follows that

$$\lambda_{CE}(w) = \left[\frac{\gamma}{w}\right]^{\frac{1}{1-\gamma}} . \tag{15}$$

Using the total differences of equation **(14)** for $T=H,L$, together with equation **(1)** and assumption (A), it can be shown that a higher wage is associated with the marginal entrepreneur being less risk averse, i.e., $\Delta\hat{R}_H/\Delta w_{CE} < 0$ and $\Delta\hat{R}_L/\Delta w_{CE} < 0$, respectively. This is shown in equations **(16)** to **(18)** for $T=H,L$:

$$D \equiv v\left(r_{CE}(w), \pi_T, \hat{R}_T\right) - w = 0 , \tag{16}$$

$$\Delta D = \frac{\partial v}{\partial \hat{R}_T}\Delta\hat{R}_T + \left(\frac{\partial v}{\partial \hat{R}_T} - 1\right)\Delta w = 0 , \tag{17}$$

$$\frac{\Delta\hat{R}_T}{\Delta w} = \frac{1 - \frac{\partial v}{\partial w}}{\frac{\partial v}{\partial \hat{R}_T}} < 0 . \tag{18}$$

This, in turn, means that, a higher wage induces smaller proportions of individuals to become entrepreneurs, i.e., $P(\hat{R}_H(w_{CE})) > P(\hat{R}_H(w'_{CE}))$ and

$P(\hat{R}_L(w_{CE})) > P(\hat{R}_L(w'_{CE}))$, whenever $w_{CE} < w'_{CE}$. In order to see that there exists a unique wage, w_{CE}, which also satisfies part (ii) of definition 4.1, define excess supply of labour as

$$Z(w) = 1 - \varepsilon_H(w) - \varepsilon_L(w) - [\pi_H \varepsilon_H(w) + \pi_H \varepsilon_L(w)]\, \lambda_{CE}(w) \ . \qquad (19)$$

It is now easy to check that $Z<0$ for $w=0$ and $Z>0$ for $w \to \infty$. Moreover, $Z(w)$ is continuous and monotonic in the relevant range ($w>0$). $Z(w)$ is continuous because all its parts are continuous: From **(19)**, it is clear that $\lambda(w)$ is continuous. To see that $\varepsilon_H(w)$ and $\varepsilon_L(w)$ are also continuous, re-write part (ii) of definition 4.1 as

$$\lambda_{CE} = \frac{1 - \varepsilon_H - \varepsilon_L}{\pi_H \varepsilon_H + \pi_H \varepsilon_L} \ . \qquad (20)$$

Since a marginal change in w has a marginal effect on the *lhs* of this equation, it must also have just a marginal effect on the *rhs*. This equation also shows that a marginal change in w must have just a marginal effect on the sum ($\varepsilon_H+\varepsilon_L$) and on ε_H and ε_L, separately, since they move in the same direction. (Otherwise, there would be a contradiction to part (iii) of definition 4.1.) From this continuity and monotonicity, it follows that there is a unique value $w_{CE}>0$ which clears the labour market, i.e., $Z(w_{CE})=0$. ❑

This result extends the findings of Kanbur (1979) and Kihlstrom and Laffont (1979) by establishing existence and uniqueness of CE with individuals of different abilities. As the following result confirms, individuals sort themselves according to the traditional Knightian view of entrepreneurship, depending on their risk aversion.

Result 4.4: *In a CE, relatively more talented than untalented individuals become entrepreneurs, i.e.,* $P(\hat{R}_L) > P(\hat{R}_H)$ *. Or, equivalently, the proportion of H types is higher among entrepreneurs than in the total population, i.e.,*

$$\frac{\varepsilon_{HCE}}{\alpha} > \frac{\varepsilon_{LCE}}{1-\alpha} \Leftrightarrow \left\{ \frac{\varepsilon_{HCE}}{\varepsilon_{HCE}+\varepsilon_{LCE}} > \alpha \wedge \frac{\varepsilon_{LCE}}{\varepsilon_{HCE}+\varepsilon_{LCE}} < 1-\alpha \right\} \qquad (21)$$

DEMONSTRATION OF RESULT 4.4: In a *CE*, $\hat{R}_H$ and $\hat{R}_L$ are defined by equation **(14)**. Since $\pi_L<\pi_H$, it follows that $\hat{R}_L > \hat{R}_H$, because of assumption (A). And from this, the result follows. ❑

This result suggests that although CE equalises workers' marginal products across firms[21], in general, it will not achieve a first-best. There *are three main sources of inefficiency of CE.* First, the absence of an insurance market makes it necessary for entrepreneurs to bear risk. This risk lowers expected utility, since individuals are risk averse. Second, also due to risk aversion, CE is normally characterised by a shortage of entrepreneurs (see result 4.6, below). Third, CE is in general characterised by a misallocation of talent, in the sense that relatively too many untalented and/or too few talented individuals become managers. Result 4.5 gives, for each regime, necessary and sufficient conditions for these inefficiencies to vanish and CE to coincide with a first-best.

Result 4.5: *CE coincides with a first-best if and only if:*

- $R^-=0$, *and there are sufficiently many risk neutral individuals, i.e.,* $\alpha P(0) \geq 1-\gamma$, *in regime I,*
- *the most risk averse individual is not too risk averse, i.e.,* $R^+<B_{II}$, *in regime II, where* B_{II} *is defined below, and,*
- $R^-=0$, *and there are sufficiently many risk neutral individuals, and the most risk averse ones are not too risk averse, i.e.,*
  $(1-\alpha)P(0) \geq 1-\alpha-\gamma[1+\alpha(\pi_H/\pi_L-1)]$, *and* $R^+<B_{III}$, *in regime III, where* B_{III} *is defined below.*

DEMONSTRATION OF RESULT 4.5: From result 4.1 we know that there is a unique maximum feasible output and a unique output maximising allocation of talent. This implies that $(\varepsilon_{HCE},\varepsilon_{LCE})=(\varepsilon_{HFB},\varepsilon_{LFB})$ is necessary for CE to coincide with a first-best. It is also sufficient, since labour is allocated efficiently across firms.

21. From part (i) of definition 1, it is clear that the labourers will be allocated to firms efficiently, i.e., their marginal products will be equalised.

From part (i) of definition 1, $w_{CE} = \gamma \lambda_{CE}^{\gamma-1}$. The profit of a successful entrepreneur is $r_{CE} = \lambda_{CE}^{\gamma-1} - \lambda_{CE} w_{CE}$. Notice also that $(\varepsilon_{HCE},\varepsilon_{LCE})=(\varepsilon_{HFB},\varepsilon_{LFB})$ implies that $\lambda_{CE}=\lambda_{FB}$.

Using this, we can write

$$\frac{w_{CE}}{r_{CE}} = \frac{\gamma \lambda_{FB}^{\gamma-1}}{\lambda_{FB}^{\gamma} - w_{FB}\lambda_{FB}} = \frac{\gamma}{(1-\gamma)\lambda_{FB}}. \qquad (22)$$

In <u>*regime I*</u>, the first-best requires $\varepsilon_{HFB}=1-\gamma$ and $\varepsilon_{LFB}=0$. From part (i) of definition 1, we know that in *regime I* every viable firm employs $\lambda_{FB}=\gamma/[(1-\gamma)\pi_H]$ labourers.

This yields

$$w_{CE} = \pi_H r_{CE} \ . \qquad (23)$$

On the other hand, in *regime I*, some H-type individuals must be indifferent between the occupations, i.e.,

$$w_{CE} \leq v(r_{CEI}, \pi_H, \bar{R}). \qquad (24)$$

Since, from assumption (B),

$$v(r_{CEI}, \pi_H, \bar{R}) \leq v(r_{CEI}, \pi_H, 0) \ , \qquad (25)$$

and, from assumption (C),

$$v(r_{CEI}, \pi_H, 0) = r_{CEI}\pi_H \ , \qquad (26)$$

the necessary condition for a *regime I* equilibrium to be consistent with **(23)** (and with the first-best) is that $R_H=\bar{R}=0$. Notice that this condition is also sufficient for

$(\varepsilon_{HCE},\varepsilon_{LCE})=(\varepsilon_{HFB},\varepsilon_{LFB})$, in *regime I*. To see this, notice that, since $\pi_H>\pi_L$, condition **(23)** implies

$$w_{CE} > \pi_H r_{CE} \quad . \tag{27}$$

which guarantees that no L-individual wants to become entrepreneur, i.e., $\varepsilon_{LCE}=\varepsilon_{LFB}=0$. This establishes that, in *regime I*, $R_H=R^-=0$ is necessary and sufficient for $(\varepsilon_{HCE},\varepsilon_{LCE})=(\varepsilon_{HFB},\varepsilon_{LFB})$.

Next, consider <u>*regime II*</u>, where $\varepsilon_{HFB}=\alpha$ and $\varepsilon_{LFB}=0$, implying $\lambda_{FBII}=(1-\alpha)/(\alpha\pi_H)$. This yields the following expression for the wage

$$w_{CE} = \gamma\left(\frac{\alpha\,\pi_H}{1-\alpha}\right)^{1-\gamma} , \tag{28}$$

and, for the profit of a successful entrepreneur,

$$r_{CE} = (1-\gamma)\left(\frac{1-\alpha}{\alpha\,\pi_H}\right)^{\gamma} . \tag{29}$$

For $\varepsilon_{HFB}=\alpha$, it must be the case that even the most risk averse H-individual prefers the risky occupation to the safe wage, i.e.,

$$w_{CE} < v(\,r_{CE}\,,\pi_H\,,R^+) \ , \tag{30}$$

or, using **(28)** and **(29)**,

$$\gamma\left(\frac{\pi_H\,\alpha}{1-\alpha}\right)^{1-\gamma} < v\left(\{1-\gamma\}\left[\frac{1-\alpha}{\pi_H\,\alpha}\right]^{\gamma},\pi_H\,,R^+\right) \ . \tag{31}$$

Now define B_{II}, the value of R^+ for which the *rhs* of **(30)** is equal to the *lhs*. Then **(30)** states that $R^+<B_{II}$ is a necessary condition for CE to coincide with the first-best. To see that it is also sufficient, notice that for $\varepsilon_{LFB}=0$, it must be the case that

$$w_{CE} \geq v(r_{CE}, \pi_L, \bar{R}) \ . \tag{32}$$

But in *regime II* this condition is satisfied for any $R \geq 0$, since not even risk neutral L-individuals would want to become entrepreneurs. This can be seen from the definition of *regime II* given in section 3. It can be re-written as

$$\frac{\gamma}{1-\gamma} \cdot \frac{\alpha}{1-\alpha} \cdot \pi_H > \pi_L \ , \tag{33}$$

which, using **(28)** and **(29)**, implies that $w_{CE} > \pi_L r_{CE}$. And this guarantees that all L-individuals choose to be labourers.

Finally, consider *regime III*, where $\varepsilon_{HFB} = \alpha$ and $\varepsilon_{LFB} = 1-\gamma-\alpha-\gamma\alpha(\pi_H/\pi_L - 1)$. Define the shadow wage, w_{FBIII}, and the shadow profits, r_{FBIII}, as follows:

$$w_{FBIII} = \gamma^{\gamma} \cdot (1-\gamma)^{1-\gamma} \cdot \pi_L^{1-\gamma} \ , \tag{34}$$

and

$$r_{FBIII} = \lambda_{FBIII}^{\gamma} - w_{FBIII}\, \lambda_{FBIII} \ . \tag{35}$$

For CE to coincide with a first-best, w_{FBIII} must satisfy parts (i) to (iii) of definition 4.1. It is easily checked that it satisfies part (i), if $\lambda_{CE} = \lambda_{FBIII}$. Substituting

$$\lambda_{FB} = \lambda_{FBIII} = \frac{\gamma\left[1 + \alpha\left(\frac{\pi_H}{\pi_L} - 1\right)\right]}{\pi_H \alpha + \pi_L\left[1 - \alpha - \gamma - \gamma\alpha\left(\frac{\pi_H}{\pi_L} - 1\right)\right]} = \frac{\gamma}{(1-\gamma)\pi_L} \tag{36}$$

into equation **(21)**, yields

$$w_{CE} = \pi_L\, r_{CE} \ . \tag{37}$$

Since in *regime III* the marginal entrepreneur is of L type, equation **(15)** determines R_{LCE}. By assumptions (A) and (B), for conditions **(15)** and **(37)** to hold simultaneously, it is necessary that $\hat{R}_{LCE} = R^{-} = 0$. N. This condition is, however, not sufficient for $(\varepsilon_{HCE},\varepsilon_{LCE})=(\varepsilon_{HFB},\varepsilon_{LFB})$, in *regime III*, since it does not guarantee that $\varepsilon_{HCE}=\varepsilon_{HFBIII}=\alpha$. The following condition guarantees that, at the wage given by equation **(37)**, all H-individuals become entrepreneurs, as required for the first-best. It states that even the most risk averse H-individual must prefer the risky occupation,

$$\pi_L \, r_{CE} < v(\, r_{CE}, \pi_H, R^{+}). \tag{38}$$

Define B_{III} as the value of R^{+} which makes the *lhs* in **(38)** equal to the *rhs*. Using equations **(34)** to **(36)**, B_{III} is defined by

$$\pi_L \cdot \left[\lambda^{\gamma} - \gamma^{1-\gamma} \cdot (1-\gamma)^{\gamma} \cdot \pi_L^{1-\gamma} \cdot \lambda\right] =$$

$$v(\lambda^{\gamma} - \gamma^{1-\gamma} \cdot (1-\gamma)^{\gamma} \cdot \pi_L^{1-\gamma} \cdot \lambda \, , \pi_H \, , B_{III}) \quad , \tag{39}$$

where λ is given by equation **(36)**. Thus, the second condition which has to hold for CE to coincide with a first-best in *regime III* is $R^{+}<B_{III}$. ❑

Notice that, even if CE does not coincide with a first-best, some individuals may be better off under CE than under the equal-treatment-first-best. But, of course, for each CE, there exists a redistribution of income which makes everyone better off and which is feasible in first-best.

Corollary 4.1: *In the special case where all individuals are risk-neutral, CE coincides with the first-best.*

DEMONSTRATION OF COROLLARY 4.1: It is easily checked that $R^{+}=0$ implies that the conditions given in result 4.5 are satisfied, in all three regimes. ❑

A similar result appears in Kihlstrom and Laffont (1979) in their model, in which all individuals have the same ability. The result does not hold if workers have different productivities. Then, the non-observability of workers' abilities makes competitive equilibrium inefficient, even with all risk neutral individuals. Laussel and Le Breton (1995) show that

there will be too few firms in competitive equilibrium. In the present context, the inefficiencies of competitive equilibrium are also associated with too few firms. But here the underlying cause is the risk aversion of individuals. This is shown in the following result.

Result 4.6: *If the conditions given in result 4.5 are not satisfied, then there will be too few entrepreneurs in CE in the sense that* $\Omega_{FB} > \Omega_{CE}$.

DEMONSTRATION OF RESULT 4.6: Recall that the number of operating firms $\Omega_{CE}=\pi_H\varepsilon_{HCE}+\pi_L\varepsilon_{LCE}$. If the conditions given in result 4.5 are not satisfied, then CE does not coincide with the first-best, i.e., $(\varepsilon_{HCE}, \varepsilon_{LCE}) \neq (\varepsilon_{HFB}, \varepsilon_{LFB})$. The three steps shown below establish that this inequality implies that CE is characterised by too few operating firms relative to the first-best, i.e., $\Omega_{FB}>\Omega_{CE}$.

- Step 1: $\varepsilon_{HCE} > \varepsilon_{HFB}$ *is impossible.*

First, notice that $\varepsilon_{HCE}>\varepsilon_{HFB}$ is only possible in *regime I* and implies that $\varepsilon_{CE}>\varepsilon_{FB}$. In all three regimes, equilibrium wages and profits satisfy

$$w = \gamma\left(\frac{\pi_H\varepsilon_H + \pi_L\varepsilon_L}{1-\varepsilon_H-\varepsilon_L}\right)^{1-\gamma} \tag{40}$$

and

$$r = (1-\gamma)\left(\frac{1-\varepsilon_H-\varepsilon_L}{\pi_H\varepsilon_H + \pi_L\varepsilon_L}\right)^{\gamma}, \tag{41}$$

respectively. From equations **(40)** and **(41)**, it follows that an increase in ε_{CE} is associated with w_{CE} going up and r_{CE} going down. But, taking into account assumption (A), this is not consistent with the equilibrium condition $w_{CE}=v(r_{CE},\pi_H,0)$.

- Step 2: $\varepsilon_{HFB} > \varepsilon_{HCE}$ *implies* $\Omega_{FB}>\Omega_{CE}$.

If $\varepsilon_{HFB}>\varepsilon_{HCE}$ is associated with $\varepsilon_{LFB} \geq \varepsilon_{LCE}$, then step 2 is established. But even if $\varepsilon_{HFB}>\varepsilon_{HCE}$ is associated with $\varepsilon_{LFB}<\varepsilon_{LCE}$ step 2 holds. To see this, suppose that all L-entrepreneurs are risk neutral. (If some were risk averse, the result would be even

clearer.) Then the equilibrium condition $w_{CE}=\pi_L r_{CE}$ must hold. Substituting from equations **(40)** and **(41)** into this condition, yields

$$\varepsilon_{LCE} = (1-\gamma)(1-\varepsilon_{HCE}) - \gamma\,\varepsilon_{HCE}\frac{\pi_H}{\pi_L} \quad . \tag{42}$$

Differentiating Ω_{CE}, using **(42)**, yields

$$\frac{\partial\Omega}{\partial\varepsilon_{HCE}} = -(1-\gamma)(\pi_H-\pi_L) \le 0 \quad . \tag{43}$$

- Step 3: $\varepsilon_{HFB}=\varepsilon_{HCE}$ *implies* $\varepsilon_{LFB}>\varepsilon_{LCE}$.

From equations **(40)** and **(41)**, $\varepsilon_{HFB}=\varepsilon_{HCE}$ and $\varepsilon_{LFB}<\varepsilon_{LCE}$ imply $w_{FB}>w_{CE}$ and $r_{FB}<r_{CE}$. This, however, is inconsistent with equilibrium in all three regimes. To see this, consider first *regime I*, in which $\varepsilon_{HFB}=\varepsilon_{HCE}$ implies the equilibrium relation $w_{CE}=\pi_H r_{CE}=v(r_{CE},\pi_H,0)$. This relation is inconsistent with $w_{FB}>w_{CE}$ and $r_{FB}<r_{CE}$. In *regime III*, $\varepsilon_{HFB}=\varepsilon_{HCE}$ implies the equilibrium relation $w_{CE}=\pi_L r_{CE}=v(r_{CE},\pi_L,0)$ which also is inconsistent with $w_{FB}>w_{CE}$ *and* $r_{FB}<r_{CE}$. Finally, in *regime II*, equations **(40)** and **(41)**, imply that a risk neutral L-individual will choose the risky occupation if and only if

$$\frac{\gamma}{1-\gamma}\frac{\alpha}{1-\alpha}\pi_H < \pi_L \,. \tag{44}$$

As this contradicts the definition of *regime II* given in section 4.3, which can be re-written as

$$\frac{\gamma}{1-\gamma}\frac{\alpha}{1-\alpha}\pi_H > \pi_L \,, \tag{45}$$

it means that step 2 holds also for *regime II*. ❑

Corollary 4.2: *More production units operate in the collectivist economy than in the corresponding competitive equilibrium with the same fundamentals.*

DEMONSTRATION OF COROLLARY 4.2: Follows directly from results 4.2 and 4.6. ❑

Clearly, results 4.2 and 4.5 together imply that if CE achieves a first-best then total output in CE exceeds total output of the corresponding C with the same fundamentals. On the other hand, the inefficiencies due to risk-bearing in CE may be so severe that they exceed the allocation inefficiencies in C, so that wages and total output are higher in the collectivist system. This is demonstrated now.

Result 4.7: *As society becomes more risk averse (in the sense that every member becomes more risk averse), the proportion of entrepreneurs, total output and wages fall. For an infinitely risk averse society the three variables fall to zero.*

DEMONSTRATION OF RESULT 4.7: Suppose that there are only H-individuals in the economy, i.e., $\alpha=1$. (This only makes the result stronger.) Now consider any shift in the distribution $P(R)$ of risk aversion in the population which results in the indifferent individual being more risk averse, i.e., in a higher $\hat{R}_H$ (Every individual becoming marginally more risk averse is sufficient for this.) Since, in equilibrium, condition **(14)** must hold, it follows (using assumption (A)) that $\partial w_{CE}/\partial \hat{R}_H < 0$. From definition 4.1, ε_{HCE} is given by

$$\frac{1-\varepsilon_{HCE}}{\pi_H \cdot \varepsilon_{HCE}} = \left(\frac{\gamma}{w_{CE}}\right)^{\frac{1}{1-\gamma}} \quad \Rightarrow \quad \varepsilon_{HCE} = \frac{1}{1+\pi_H \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}}} \quad . \tag{46}$$

This shows that $\partial \varepsilon_{HCE}/\partial w_{CE} < 0$ and $\partial \varepsilon_{HCE}/\partial \hat{R}_H < 0$. Moreover, total output

$$G_{CE} = (1-\varepsilon_{HCE})^{\gamma} \cdot (\pi_H \, \varepsilon_{HCE})^{1-\gamma} \tag{47}$$

also falls, i.e., $\partial G_{CE}/\partial \hat{R}_H < 0$. Also from condition **(14)** and assumption (A), $R_H = \infty$ implies $w_{CE}=0$, which in turn implies $G_{CE}=0$ and $\varepsilon_{HCE}=0$. ❑

Notice that wages and output in C do not depend on individuals' risk attitudes. This implies that a sufficiently risk averse society is better off with a collectivist system. The remainder of this chapter considers the issue of economic and political transition from a pure collectivist economy to a competitive economy.

4.6 Total reform

Suppose the economy initially takes the form of a collectivist economy of the type described in section 4.4. In such an economy all individuals are certain to earn the same income, w_C. Consider first the case of total reform where the entire collectivist sector is closed down at once and all individuals find themselves in a modern economy of the type described in section 4.5. From corollary 4.1, we know that CE coincides with the a first-best, in the case where all individuals are risk neutral. Therefore, in a risk neutral society individuals will, *ex ante,* unanimously prefer total reform to no reform.

This changes as soon as CE is characterised by risk averse individuals taking entrepreneurial jobs. Then CE becomes subject to the inefficiencies described in the previous section. Increases in risk aversion tend to amplify these inefficiencies. C, meanwhile, remains unaffected by the changes in risk attitudes. As the following corollary shows, there comes a point where individuals prefer no reform to total reform.

Corollary 4.3: *An overall increase in society's risk aversion leads to an increase in the proportion of individuals preferring no reform to total reform. For a sufficiently risk averse society this proportion exceeds 0.5 (i.e., a majority prefers no reform) and eventually reaches 1, for infinite risk aversion.*

DEMONSTRATION OF COROLLARY 4.3: Since C is unaffected by changes in risk attitudes, the corollary follows directly from result 4.7. ❑

Another interesting result regards the political stability of reform. The historical experience from the recent reforms in formerly collectivist economies has shown that the attitude of a majority towards reforms may change dramatically after the reforms are enacted. In many countries the reforming governments were voted out of office at the first elections after the reforms. As the following result shows, this phenomenon can be explained within this framework.

Result 4.8: *It is possible (namely under the conditions given in the demonstration) that a majority of individuals ex ante prefer total reform to no reform, even if the reform makes a majority of individuals worse-off, ex post.*

DEMONSTRATION OF RESULT 4.8: Define the median voter m in the following way: As the wage w falls from ∞ towards 0, more and more individuals will choose to become entrepreneurs. Then, there exists a wage level

and an individual m, such as:

$$w_m = v(r_{CE}(w_m), \pi_m, R(m)) \tag{48}$$

and

$$\varepsilon_H(w_m) + \varepsilon_L(w_m) = \frac{1}{2} . \tag{49}$$

With this definition of m, for a majority of individuals to *ex ante* prefer total reform to no reform it is necessary and sufficient that

$$v(r_{CE}, \pi_m, R(m)) > w_C . \tag{50}$$

Now, notice that if **(50)** holds, if m is a worker and if $w_C<w_{CE}$, then the result cannot be explained, since a majority of individuals will benefit from the reform. Thus, for a majority to suffer from reform, it is necessary that m be an entrepreneur. The remainder of this demonstration derives two different conditions under which the result holds:

(i) Suppose that **(50)** holds, that m is an entrepreneur and that $w_C<w_{CE}$. If also the condition $1-\varepsilon_{HCE}-\varepsilon_{LCE}+\varepsilon_{HCE}(1-\pi_H)+\varepsilon_{LCE}(1-\pi_{LCE})>0.5$ holds (i.e., a majority of the population is composed of either unlucky entrepreneurs or workers), then the result holds, since *ex post* all workers are worse off (as $w_C<w_{CE}$) and will form a majority together with the unlucky entrepreneurs (who receive zero income). The condition can be simplified to $\Omega_{CE}<0.5$.

(ii) Suppose that **(50)** holds, that m is an entrepreneur and that $w_C>w_{CE}$. If also the condition $\varepsilon_{HCE}(1-\pi_H)+\varepsilon_{LCE}(1-\pi_L)>0.5$ holds (i.e., a majority of the population is composed of unlucky entrepreneurs), then the result clearly holds. ❑

This simple result can be discussed in the light of Fernandez and Rodrik (1991). In their paper, they show that a large sector where a majority of people may lose after some

reforms (opening the borders to trade in their case) may be enough to prevent any reform, although a majority of the population would gain from the reform. Here, the argument is reversed. Although a majority will lose, there can be a majority of people willing to take the gamble of the market economy.

In any case, reforms will be primarily supported by potential entrepreneurs since they are the ones who take advantage of the competitive institutions, in case somebody should gain from reforms. So it is no surprise that supporters of change tend to be found among individuals with low risk aversion and high ability.

4.7 Gradual reform and dual economy

So far, I have considered the collectivist and the competitive institutions separately. Clearly, these two allocation mechanisms can coexist with a fraction of the population being employed in the collectivist sector and the remainder in the competitive sector. Such a dual economy represents a natural view of partial or gradual reforms. I will interpret the dual economy equilibrium (DE) defined below, the outcome of gradual reform. Individuals can make a free choice to work either in the collectivist or in the competitive sector and society decides the freedom of this choice to individuals. The new sequence of events is depicted in figure 4.4.

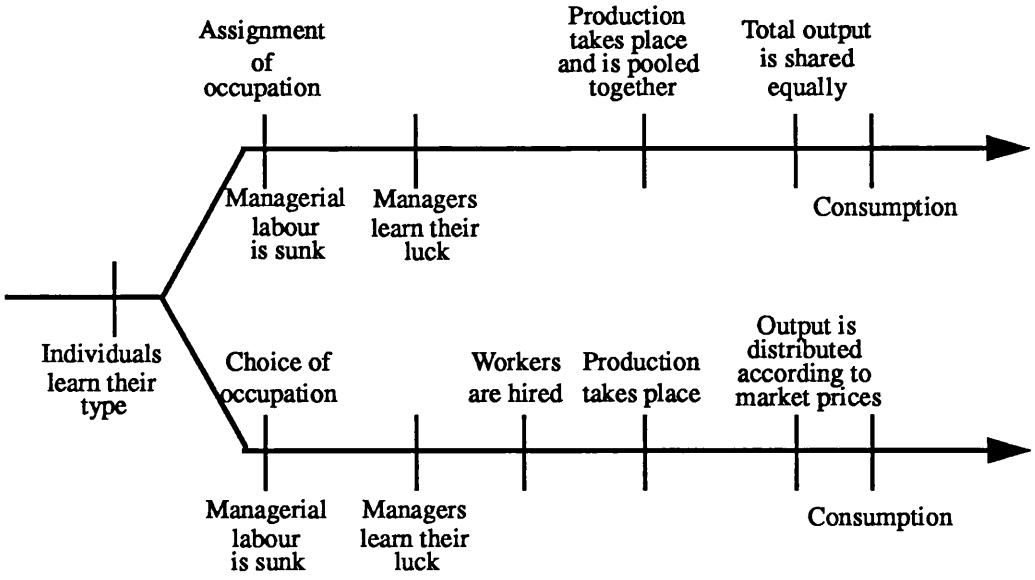

FIGURE 4.4: The sequence of events in the dual economy.

Definition 4.2: *A Dual Economy Equilibrium (DE) is a competitive wage w_{DE}, a collectivist wage w_S and partitions of the sets of talented and untalented individuals in competitive sector workers Δ_{TDE}, competitive sector entrepreneurs Γ_{TDE}, and collectivist sector workers and managers Σ_T (for T-H,L), such that:*

(i) each entrepreneur in the competitive sector demands λ_{DE} workers such as to maximise profits, whereby she takes the wage rate in the competitive sector as given, i.e.,

$$\lambda_{DE} = arg\ max\ [g(\lambda) - w_{DE}\,\lambda]\ ; \tag{51}$$

(ii) labour supply equals labour demand in the competitive sector, i.e.,

$$1 - \sigma_H - \sigma_L - \varepsilon_{HDE} - \varepsilon_{LDE} = (\pi_H\,\varepsilon_{HDE} + \pi_L\,\varepsilon_{LDE})\,\lambda_{DE}\ , \tag{52}$$

where ε_{TDE} (for T=H,L) denote the proportions of type T individuals in the total population choosing to be entrepreneurs and σ_T (for T=H,L) denote the proportions of T individuals in the collectivist sector;

(iii) individuals make their sectoral and occupational choices such as to maximise their expected utility, i.e., for T=H,L it must be the case that

$$v(r_{DE}, \pi_T, R(k)) > max[w_S, w_{DE}] \ \Rightarrow\ k \in \Gamma_T\ , \tag{53}$$

$$w_{DE} > max[w_S, v(r_{DE}, \pi_T, R(k))] \ \Rightarrow\ k \in \Delta_T\ , \tag{54}$$

and

$$w_S > max[w_{DE}, v(r_{DE}, \pi_T, R(k))] \ \Rightarrow\ k \in \Sigma_T\ , \tag{55}$$

where r_{DE} denotes entrepreneurial income.

(iv) When they are indifferent between the two sectors, individuals choose randomly (i.e., independently of their type).

(v) w_S is chosen such a to be maximised, subject to $w_S(\sigma_H + \sigma_L) \leq G_S$, where G_S denotes total output of the collectivist sector.

(vi) ε_S is chosen such as to maximise G_S, given $\sigma \equiv \sigma_H + \sigma_L$.

This definition formalises the idea of gradual reform where individuals have the choice to work either in the collectivist sector or in the competitive sector (see part (iii)). Each of these two sectors, works essentially in the same way as the corresponding pure economies. This is expressed in parts (i) to (iii), for the competitive institutions, and in parts (v) and (vi), for the collectivist institutions. Only part (iv) is new. It states that if an individual is indifferent between the two sectors, her choice will be unbiased, i.e., independent of her type. I am now in the position to demonstrate that the outcome of partial reform as captured by definition 4.2 is uniquely determined.

Result 4.9: *DE exists and is unique.*

DEMONSTRATION OF RESULT 4.9: DE is a 7-tuple $(w_S,\varepsilon_S,\lambda_{SH},\lambda_{SL},w_{DE},\varepsilon_{HDE},\varepsilon_{LDE})$ satisfying definition 4.2. I first take w_S, and ε_S as given, and show that the other five variables are uniquely determined by parts (i) to (iv) of definition 4.2. From part (iii), it must be that $w_{DE}=w_S$. (Otherwise, corner solution which is unique from discussions in sections 4.4 and 4.5.) Let $\hat{R}_{HDE}$ be defined by

$$w_S = max\,[\,w_S\,,\,v\,(\,r_{DE}(w_S\,),\,\pi_H\,,\,\hat{R}_{HDE})\,]\ , \tag{56}$$

and $\hat{R}_{LDE}$ by

$$w_S = max\Big[\,w_S\,,\,v\Big(\,r_{DE}(w_S\,),\,\pi_L\,,\,\hat{R}_{LDE}\Big)\Big]\ . \tag{57}$$

Then, by the monotonicity of $v(.)$ regarding the equilibrium wage, it follows that $\hat{R}_{HDE}$ and $\hat{R}_{LDE}$ are uniquely determined. Therefore, ε_{HCE} and ε_{LCE} are also unique. From result 4.3, this means that $\lambda_{DE}(\pi_H\varepsilon_{HDE}+\pi_L\varepsilon_{LDE})$, i.e., employment in the competitive sector, is unique. This, in turn, determines $(\sigma_H+\sigma_L)$, i.e., employment in the collectivist sector, uniquely from

$$1-(\sigma_H+\sigma_L\,)\ =\ (\varepsilon_{HDE}+\varepsilon_{LDE})\ +\ \lambda_{DE}\,(\pi_D\,\varepsilon_{HDE}\,+\,\pi_L\,\varepsilon_{LDE})\ . \tag{58}$$

Given the unbiasedness assumption made in part (iv) of definition 4.2, the proportion of H relative to L types is the same in the collectivist sector as among the competitive sector workers, such that σ_H and σ_L are unique.

I have now established that for any given pair (w_S,ε_S), the remaining variables are uniquely determined. Then, from part (vi) of definition 4.2, it follows that ε_S is unique from the maximisation. Similarly, from part (v), it follows that w_s is unique from the maximisation. ❑

The definition of DE formalises the outcome of gradual reform. Individuals are free to choose the sector they want to work in and, within the competitive sector, they can also choose their occupation. In general, these choices lead to a DE with some individuals in each sector. But there is also the possibilities of DE coinciding with either CE or C. The former, I would call "spontaneous big bang", the latter "conservative society". The following paragraphs and the example given below, show that these two outcomes are actually possible.

If $w_C<w_{CE}$, then this means that at least some individuals must have moved to the competitive sector. As more talented individuals tend to become entrepreneurs (see result 4.3), proportionally more talented individuals will be in the competitive sector and both average managerial ability and output per worker in the collectivist sector must fall, i.e., $\sigma_H/(\sigma_L+\sigma_H)=\alpha_S<\alpha$ and

$$\frac{G_S(\varepsilon_S)}{\sigma} = [\pi_H \alpha_S + \pi_L (1-\alpha_S)] (1-\gamma)^{1-\gamma} \gamma^{\gamma} < w_C . \qquad (59)$$

But this, in turn, means that even more individuals may leave the collectivist sector until all individuals are employed in the competitive sector. This is the outcome I call a "spontaneous big bang".

Now, consider the case $w_C>w_{CE}$ and $v(r_{CE}(w_C),\pi_H,R^*)<w_C$. In this case, clearly no-one leaves the collectivist sector, and DE coincides with C. This is the outcome I call the "conservative society".

Next, consider the case $w_C>w_{CE}$ and $v(r_{CE}(w_C),\pi_H,R^-)>w_C$. In this case, a competitive sector develops. It is such that $w_S<w_C$, and again the possibility of "spontaneous big bang" arises. This is illustrated in the following example.

Example: Suppose $\gamma=0.5$, $\pi_L=0$ and $\pi_H=1$. From this it follows that $w_S=\alpha_S/4$ and $w_C=\alpha/4$. Suppose risk attitudes are such that a fraction p of the population is risk neutral and $(1-p)$ are infinitely risk averse. This implies that in CE the wage will be

$$w_{CE} = \frac{1}{2}\left(\frac{p\alpha}{1-p\alpha}\right)^{1/2} \qquad \text{if} \qquad p\alpha < \frac{1}{2}\,, \tag{60}$$

which means that $w_{CE}<w_C$. Moreover, all risk neutral H-types become entrepreneurs, if

$$\frac{1-p\alpha}{p\alpha} > \frac{\alpha^2}{4}\,, \tag{61}$$

which always holds, and

$$\alpha_S = \frac{\alpha(1-p)}{1-p\alpha}\,. \tag{62}$$

One can check that for $\alpha=0.5$ and $p=0.11$, it follows that $w_S<w_{CE}<w_C$, which is the situation discussed above, leading to a "spontaneous big bang". ❑

What can be said about society's choice between total reform, gradual reform and no reform through majority voting? As shown above, choosing between C and CE a majority vote may go either way (corollary 4.3), and if CE is chosen, *ex post* a majority may regret it (result 4.8). With the possibility of gradual reform, the social choice is enriched substantially, unless, DE coincides either with C or with CE. In particular, it is *a priori* possible that majority voting is intransitive. The following result and corollary, however, exclude this possibility.

Result 4.10: *DE is never preferred by a majority to both CE and C.*

DEMONSTRATION OF RESULT 4.10: Suppose that DE does not coincide with either CE or C. (Otherwise, the result is not meaningful.) It is easy to show that the median voter m,

as defined in the demonstration of result 4.8, is always part of the winning majority (she is, actually, the only pivotal voter in the society). Notice also that, if DE is different from both CE and C, equilibrium wages are ranked as follows:

$$w_C > w_{DE} = w_S > w_{CE} \ . \qquad (63)$$

The last inequality implies that $r_{CE}>r_{DE}$. With these facts in mind, the rest of the demonstration is very simple: if *m* is a worker in DE, then she prefers C to DE. If *m* is an entrepreneur in DE, she may prefer DE to C, but always prefers CE to DE. ❑

This result shows how the conflict of interest (or class struggle) results in a polarisation of political decisions. I believe that it helps to understand why in many cases - particularly in some former Soviet Republics - gradual reformers tended to lose either to the radical reformers or to the conservatives. There are basically only two ways how moderate reform can be a politically stable outcome in the transition process. Either the democratic choice in the society is restricted by some form of dictate, in which case a majority may opt for gradual reform as opposed to no reform. The second way is if, after total reform has been realised and a majority of the population is made worse-off (see result 4.8 and corollary 4.5, below), they may want to return to a collectivist system. By this time, however, C may not be economically feasible any more, so that they would accept the DE as a substitute.

Corollary 4.4: *Majority rule is always transitive regarding the choice between DE, CE and no reform.*

DEMONSTRATION OF COROLLARY 4.4: There are two orderings which violate transitivity, namely

(i) $DE \succcurlyeq_M CE \succcurlyeq_M C \succcurlyeq_M DE$, and

(ii) $DE \succcurlyeq_M C \succcurlyeq_M CE \succcurlyeq_M DE$,

where $\succcurlyeq_M$ stands for "is preferred by a majority to". But result 4.10 implies that these two orderings are impossible. ❑

This result is important in two ways. Firstly, it makes clear that result 4.8, above, and corollary 4.5 below, i.e., the finding that a majority of voters may regret the decision taken

by a majority in an earlier vote, are not driven by intransitivity, but by the gap between prospects and realisations. Secondly, the transitivity result improves the descriptive relevance of the analysis. Clearly, the results concerning the politics of transition (result 4.8 and corollaries 4.3 and 4.5) would not be very interesting if the political outcome was indeterminate.

The next question raised by the possibility of a dual economy equilibrium is whether it eliminates the possibility of society making the "wrong" choice in the sense of result 4.8. In other words, is it still possible that society chooses either total or gradual reform, although it will make a majority of the population worse-off? In fact, result 4.11 shows that the answer to this question is positive.

Corollary 4.5: *It is possible that a majority of individuals ex ante prefers either gradual or total reform to no reform, even if the reform makes a majority of individuals worse off, ex post.*

DEMONSTRATION OF COROLLARY 4.5: As long as only the dual economy coincides with the competitive economy, the corollary coincides with result 4.8. If the dual economy is distinct from both the competitive and the collectivist economy, it suffices to observe that

$$v(r_{CE}, \pi_m, R(m)) > w_C \tag{64}$$

and

$$0.5 > \varepsilon_{HDE}\pi_H + \varepsilon_{LDE}\pi_L \,. \tag{65}$$

Equation **(64)** states that the median voter is better-off under the dual economy. There is a majority who regrets *ex-post* if there is a majority composed of workers and unlucky entrepreneurs, as required by **(65)**. ❑

4.8 Conclusion

The prototype economy analysed in this paper builds on the standard Arrow-Debreu framework. *I only assume that output and employment in private firms are not observable* by outsiders. Thus, insurance markets will be missing in the competitive economy.

There is no way a private entrepreneur can be insured against her low ability or bad luck. So the potential productive inefficiency of the competitive economy stems from risk aversion. There is one alternative institution to the market: a fixed sharing rule can provide full insurance against individual income risks. But in that case, managers have no incentive to reveal their type or to adjust employment levels to plant productivity. Moreover, managers have no incentive to reveal their luck so that some units keep employing labour, although they are unproductive. This results in inefficient allocations of both talent and labour. The comparison between these two institutions is not trivial. Either one can be better.

As a special case of a collectivist institution, consider a benevolent uninformed central planner. Due to imperfect information, the central planner cannot do better than the outcome described in section 4.4. The allocation of talent and labour cannot be optimal in such a centrally planned economy. Competitive institutions partly solve this problem, but only at the cost of risk bearing. Consequently, the competitive economy is characterised by a shortage of entrepreneurs and an inefficient (though less inefficient) allocation of talent. The first part of this chapter has compared these two second-best outcomes.

In the second part of this chapter, I have used this framework to replicate some stylised facts of transitional economies. First, despite the possibility of a vote in favour of reforms, some countries are still governed according to the old rules, often by the old rulers. It may simply be that the majority of the population is better-off under a collectivist system than with a market economy, at least in the short-run. The second crucial fact is that the political debate in many countries has become polarised on the pro reform/anti reform issue. Moderate options vanished in most countries. It can be understood within our model, since the dual economy is nobody's first choice (result 4.10). The model actually highlights the antagonism between workers and entrepreneurs. Yet it may be that workers are better-off with competitive institutions. If not, they prefer the collectivist organisation, whereas entrepreneurs enjoy the highest profits with competitive institutions.

Result 4.10 is very different from the conclusions derived in Dewatripont and Roland (1992a and b) and Roland (1994). They emphasise the virtues of gradual reform and democracy. In Dewatripont and Roland (1992b), for example, the authors develop a model with benevolent (i.e. social surplus maximising), agenda-setting government which faces democratic constraints. In their model (denoted D-R), the social cost of tran-

sition to a market economy is not based on risk bearing but on the cost of compensating the losers. They show that the benefits of gradualism (=less costly transfers needed) relative to a "big bang" may exceed its costs (=slower transition). This is in contrast to result 4.10 which states that gradual reform is unlikely to be the outcome of a democratic process of transition.

The second main result of D-R concerns the problem of implementing an efficiency-enhancing reform which intertemporally hurts a majority of voters. The authors argue that the government could in large part circumvent the political constraints by designing an appropriate gradual reform. My result 4.8 and corollary 4.5 state a similar result without, however, relying on the assumption that the reforms are efficiency-enhancing. This means that the democratic process may generate reforms, even when they are not efficiency-enhancing.

There are several differences which explain the different results of D-R compared to mine. First, democratic institutions are viewed differently: D-R assume a social surplus maximising agenda-setter while I assume direct democracy. Second, the entire analysis of D-R relies on the assumption that the government can implement income transfers (at a cost), while in my model such transfers are not feasible (mainly because winners and losers are not known to others). This is the crucial difference between the two papers. But there other important differences. While D-R simply assume that the market economy is more productive than the pre-reform economy, I derive the productivities of the two systems endogenously, within a unified framework based on identical economic fundamentals. Finally, my model describes more accurately an economy-wide reform (=transition) within a general equilibrium framework, while D-R's partial equilibrium model captures rather a sectoral structural adjustment of the type currently occurring in some sectors in Western market economies.

The second antagonism analysed in this chapter concerns winners and losers. Among the losers of transition, one can find many unlucky entrepreneurs. This explains why in so many transition countries a majority of people voted initially in favour of reforms, whereas a new majority formed *ex-post* voting for (ex-)communist politicians. If one considers that the first reforms have some degree of irreversibility, the dual economy may become a viable political alternative. Note also that a dual economy, is more likely to be found in countries where reforms where not democratically decided. China and Vietnam offer examples of countries where still a significant collectivist sector survives

despite the presence of some alternative highly competitive institutions. It can also be the case, that even though total reform was never really enacted, it occurred due to the opening of a competitive sector. Russia can be seen as an example of such a situation.

Viewed after a longer period of experience, I would claim that my results seem to better explain the historic events in transition countries than D-R. While D-R seemed to be worried, at the time, that reforms in some Eastern European countries may not take place "because the move towards democracy is not yet credible or solid enough" (p. 300), my analysis may shed some light on the reasons why the "first generation" of radical reformers in most cases are now losing the second round of democratic elections to the reform-critics. At the same time China seems to be on a steady (gradual) reform path, again seemingly contradicting D-R's claim that "the power monopoly of the nomenclatura must be toppled in order to fully make the transition to a market economy" (p. 300).

Of course, my model remains extremely rudimentary and puts the emphasis only on a few aspects of the transition process. Among possible extensions, one could think of introducing aggregate risk, since it is difficult to argue that macro-economic risk would be negligible during transition. Moreover, the simplifying view of the labour market ignores unemployment. It would also be desirable to relax the extreme assumption of non-observability. Imperfect observability should lead to the possibility of imperfect insurance in the competitive sector. Grossman, Hart and Maskin (1983) present a model where partial insurance is offered to entrepreneurs contingent on the amount of labour they hire.

The point which most deserves further research is perhaps the intrinsically dynamic aspect of the transition process. The comparison of the two systems starting from the collectivist economy may not be favourable to the competitive institutions. The main idea is that risks, especially the aggregate ones, are much higher at the beginning of the process. Moreover, a market economy, starting from scratch has to overcome high transaction costs. Within the model, one could assume that individuals know their type with some uncertainty. It is only after a few years that they learn their type more precisely. This intertemporal trade-off should be embodied in the model since it has been argued that things should become worse before they improve.

To consider this argument properly, one would probably need to add some Schumpeterian aspects to the Knightian theory of entrepreneurship. These two views of entrepreneurship, often presented as antagonistic may complement each other to some extent.

The Knightian features are probably more static in nature, whereas Schumpeterian arguments are intrinsically dynamic.

REFERENCES

Acemoglu, Daron (1992) "Reward Structures and the Allocation of Talent", mimeo, LSE.

Aghion, Philippe and Olivier Blanchard (1994) "On the Speed of Transition in Central Europe", Discussion Papers in Economics No. 87, Nuffield Colledge, Oxford.

_____ and Jean Tirole (1994) "Formal and Real Authority in Organizations", Document de travail 37, Institut d'Economie Industrielle, Toulouse.

Allen, Franklin (1983) "Credit Rationing and Payment Incentives", *Review of Economic Studies,* 50: 639-646.

Austen-Smith, David and Jeffrey S. Banks (1989) "Electoral Accountability and Incumbency", in: Ordeshook (1989) *Models of Strategic Choice in Politics,* 121-148.

Banerjee, Abhijit V. and Andrew F. Newman (1993) "Occupational Choice and the Process of Development", *Journal of Political Economy* 101: 274-298.

_____ (1995) "Migration, Integration and Development", mimeo, Columbia University.

Banks, Jeffrey S. and Rangarajan K. Sundaram (1993) "Adverse Selection and Moral Hazard in a Repeated Elections Model", in Barnett et al. (1993): 295-311.

Barber, William (1961) "Disguised Unemployment in Underdeveloped Economies", *Oxford Economic Papers* 13: 193-115.

Barnett, William A., Melvin J. Hinich and Norman J. Schofield (1993), eds., *Political Economy - Institutions, Competition, and Representation,* Proceedings of the Seventh Symposium in Economic Theory and Econometrics, Cambridge.

Baron, David P. and Jongryn Mo (1993), "Campaign contributions and Party-candidate competition in services and policies", in Barnett et al. (1993): 313-353.

Barro, Robert (1973) "The Control of Politicians: an Economic Model", *Public Choice,* 14: 19-42.

Begg, D. et al. (1993) *Making Sense of Subsidiarity,* CEPR, London.

Berkok, Ugurhan (1990) "Common Agency with Adverse Selection: Existence, Uniqueness and Efficiency of Equilibrium", *Economics Letters* 33(4): 319-323.

Bernheim and Whinston (1986) "Common Agency", *Econometrica* 54(4): 923-942.

Bewley, T. F. (1980) "A Critique of Tiebout's Theory of Local Public Expenditures", *Econometrica* 49: 713-740.

Bolton, Patrick and David S. Scharfstein (1991) "A Theory of Predation Based on Agency Problems in Financial Contracting", *American Economic Review,* 80: 93-106.

Bolton, Patrick and Gérard Roland (1994) "The Break-Up of Countries: A Political Economy Analysis", mimeo, Université Libre de Bruxelles.

Brainard, W. C. and F. T. Dolbear, Jr. (1967) "The Possibility of Oversupply of Local 'Public' Goods", *Journal of Political Economy* 75: 86-90.

Brams, S. J. and P. C. Fishburn (1983) *Approval Voting,* Boston.

Brennan, Geoffrey and James Buchanan (1980) *The Power to Tax,* Cambridge.

Breton, A. and A. Scott (1978) *The Economic Constitution of Federal States,* Toronto.

Buchanan, James M. (1950) "Federalism and Fiscal Equity", *American Economic Review* 40(4): 583-599.

Bull, Clive (1987) "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics* 102(1): 147-159.

Cantillon, Richard (1755) *Essai sur la nature du commerce en générale.*

Coase, R. "The Nature of the Firm", *Economica* n.s. 4: 386-405.

Chubb, John E. (1985) "The Political Economy of Federalism", *American Political Science Review,* 79: 994-1015.

Coughlin, Peter J. (1992) *Probabilistic Voting Theory.*

Council of Europe (1994) *Definition and Limits of the Principle of Subsidiarity,* Local and Regional Authorities in Europe, No. 55.

Crémer, Jacques (1992) "Arm's-length Relationships", mimeo, IDEI, Toulouse.

_____, Antonio Estache and Paul Seabright (1994) "The Decentralization of Public Services; Lessons from the Theory of the Firm", Policy Research Paper no. 1345, World Bank.

Daveri, Francesco and Fausto Panunzi (1995) "Decentralization, Mobility Costs and the Soft Budget Constraint", mimeo, IGIER, Milan.

Dasgupta, Partha, Peter Hammond and Eric Maskin (1979) "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility", *Review of Economic Studies* 66: 185-216.

De Lecea, Antonio (1983) On Fiscal Federalism, Information, and Regional Competition, Ph. D. thesis, Université Catholique de Louvain.

Dewatripont, Mathias and Gérard Roland (1992a) "Economic Reform and Dynamic Political Constraints", *Review of Economic Studies* 59(2): 703-30.

_____ and _____ (1992b) "The Virtues of Gradualism and Legitimacy in the Transition to a Market Economy", *Economic Journal* 102: 291-300.

_____ and Eric Maskin (1995) "Credit and Efficiency in Centralized and Decentralized Economies", *Review of Economic Studies.*

Eberts, Randall and Timothy Gronberg (1988) "Can Competition Among Local Governments Constrain Government Spending?" *Economic Review 24(1)*: 2-9, Federal Reserve Bank of Cleveland.

Enke, S. (1962) "Economic Development with Limited and Unlimited Supplies of Labor", *Oxford Economic Papers* 14: 158-172.

Farrell, Joseph (1987) "Information and the Coase Theorem", *Journal of Economic Perspectives* 1(2): 113-129.

Fei, J. and G. Ranis (1964) *Development of the Labor Surplus Economy: Theory and Policy,* Homewood, Illinois: Irwin.

Ferejohn, John (1986) "Incumbent Performance and Electoral Control", *Public Choice,* 50: 5-25.

Fernandez, Raquel and Dani Rodrik (1991) "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty", *American Economic Review* 81(5): 1146-55.

Fields, Gary S. (1975) "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs", *Journal of Development Economics* 2: 165-187.

_____ (1989) "On-the-Job Search in a Labor Market Model", *Journal of Development Economics* 30: 159-178.

Fiorina, M. (1981) Retrospective Voting in American National Elections, New Haven: Yale University Press.

Fudenberg, Drew and Jean Tirole (1991) *Game Theory*, Cambridge (Massachusetts).

Galor, Esther (1991) "A Common Agency With Incomplete Information", *Rand Journal of Economics* 22(2): 274-286.

Gibbons, R. and K. J. Murphy (1992) "Optimal Incentive Contract in the Presence os Career Concerns: Theory and Evidence", Journal of Political Economy 100: 468-505.

Gordon, Roger H. (1980) "An Optimal Taxation Approach to Fiscal Federalism", *Quarterly Journal of Economics* 48(4): 567-610.

Green, Jerry and Nancy Stokey (1983) "A Comparison of Tournaments and Contracts", *Journal of Political Economy* 91: 349-364.

Greenwood, Jeremy and R. Preston McAfee (1991) "Externalities and Asymmetric Information", *Quarterly Journal of Economics:* 103-121.

Gromb, Denis (1994) "Renegotiation in Debt Contracts", mimeo, London School of Economics (FMG).

Grossman, Sanford, Oliver Hart and Eric Maskin (1983) "Unemployment with Observable Aggregate Shocks", *Journal of Political Economy* 91(6): 907-28.

Hamlin, A. (1991) "Decentralization, Competition and the Efficiency of Federalism", *The Economic Record* 67: 193-204.

Harrington, Joseph E. (1993) "The Impact of Reelection Pressures on the Fulfillment of Campaign Promises", *Games and Economic Behavior,* 5: 71-97.

Harris, John R. and Michael P. Todaro (1970) "Migration, Unemployment and Development: A Two-Sector Analysis", *American Economic Review* 60: 126-142.

Hathaway, Dale (1964) "Migration from Agriculture: The Historical Record and its Meaning", in: Eicher and Witt (1964).

Heil, James (1991) "Searching for Leviathan Revisited", *Public Finance Quarterly* 19: 334-346.

Holmstrom, Bengt (1982a) "Moral Hazard in Teams", *Bell Journal of Economics* 13: 324-340.

_____ (1982b) "Managerial Incentive Schemes - A Dynamic Perspective", in *Essays in Economics and Management in Honour of Lars Wahlbeck,* Helsinki.

Journal of Economic Perspectives (1991), Special Issue, Fall.

Kao, Charles, Kurt Anschel and Carl Eicher (1964) "Disguised Unemployment in Agriculture: A Survey", in: Eicher and Witt (1964).

Kanbur, S. M. (1979a) "Of Risk Taking and the Personal Distribution of Income", *Journal of Political Economy* 87 (4): 769-797.

_____ (1979b) "Impatience, Information and Risk Taking in a General Equilibrium Model Model of Occupational Choice", *Review of Economic Studies* 46 (4): 707-718.

_____ (1980) "A Note on Risk Takig, Entrepreneurship, and Schumpeter", *History of Political Economy* 12(4): 489-498.

_____ (1981) "Risk Taking and Taxation", *Journal of Public Economics* 15: 163-184.

Kehoe, Timothy (1989) "Cooperation Among Benevolent Governments May be Harmful", *Review of Economic Studies.*

Kihlstrom, Richard E. and Jean-Jacques Laffont (1979) "A General Equilibrium Theory of Firm Formation Based on Risk Aversion", *Journal of Political Economy* 87(4): 719-748.

_____ (1983) "Implicit Labor Contracts and Free Entry", *Quarterly Journal of Economics, Supplement* XCVIII:: 55-105.

King, David (1984) *Fiscal Tiers: The Economics of Multilevel Government,* London.

Knight, Frank H. (1921) *Risk, Uncertainty and Profit*, New York: Harper & Row.

Krelove (1992) "Competitive Tax Theory in Open Economies: Constrasined Inefficiencies and a Pigovian Remedy", *Journal of Public Economics* 48: 361-375.

Kuznets, S. (1955) "Economic Growth and Income Inequality", *American Economic Review* 45: 1-28.

Lange, Oskar (1936) "On the Economic Theory of Socialism", *Review of Economic Studies* 1.

Laussel, Didier and Michel Le Breton (1995) "A General Equilibrium Theory of Firm Formation based on Individual Unobservable Skills", *European Economic Review* 39: 1303-1319.

Lazear, E. P. and S. Rosen (1981) "Rank Order Tournaments as Optimal Labour Contracts", *Journal of Political Economy* 89: 841-864.

Ledyard, John O. (1986) "Comment on: Incumbent Performance and Electoral Control", *Public Choice,* 50: 129-132.

Lerner, A. P. (1936) "A Note on Socialist Economies", *Review of Economic Studies* 1: 157-175.

Lewis, W. Arthur (1954) "Economic Development With Unlimited Supplies of Labour", *Manchester School* 22: 139-191.

_____ (1958) "Unlimited Labour: Further Notes", *Manchester School* 26: 1-32.

Londregan, John and Thomas Romer (1993) "Polarization, Incumbency, and the Personal Vote", in Barnett et al. (1993): 355-377.

Lott, John R. and W. Robert Reed (1989) "Shirking and Sorting in a Political Market with Finite-Lived Politicians", *Public Choice* 61: 75-96.

Madison, James, Alexander Hamilton and John Jay (1787) *The Federalist Papers*, reprinted 1987, Isaac Kramnick (ed.), London.

Meltzer, A. and S. Richard (1981) "A Rational Theory of the Size of Government", *Journal of Poltical Economy* 89: 914-927.

Melumad, Nahum and Dilip Mookherjee (1989) "Delegation as Commitment: The Case of Income Tax Audits", *Rand Journal of Economics.*

_____, _____ and Stefan Reichelstein (1992a) "A Theory of Responsibility Centers", *Journal of Accounting and Economics.*

_____, _____ and _____ (1992b) "Hierarchical Decentralisation of Incentive Contracts", Working Paper (Stanford University, Stanford, California).

Meyer, Margaret A. (1995) "Cooperation and Competition in Organizations: A Dynamic Perspective", *European Economic Review,* forthcoming.

_____ and John Vickers (1994) "Performance Comparisons and Dynamic Incentives", Discussion Paper, Nuffield Colledge, Oxford.

Migué, Jean-Luc and Gérard Bélanger (1974) "Toward a General Theory of Managerial Discretion", *Public Choice* 3: 27-47.

Mookherjee, Dilip (1984) "Optimal Incentive Schemes with Many Agents", *Review of Economic Studies* 51: 433-446.

Mueller, D. (1990) *Public Choice II*, Cambridge.

Murphy, K., A. Shleifer and R. Vishny (1991) "The Allocation of Talent: The Implications for Growth", *Quarterly Journal of Economics* 106: 503-530.

Musgrave, Richard (1959) *The Theory of Public Finance*, New York.

_____ (1961) "Approaches to a Fiscal Theory of Political Federalism", in *Public Finances: Needs, Sources and Utilization,* NBER, Princeton.

_____ (1971) "Economics of Fiscal Federalism", *Nebraska Journal of Economics and Business* 10.

Myerson, Roger B. (1993a) "Effectiveness of Electoral Systems for Reducing Government Corruption: A Game-Theoretic Analysis", *Games and Economic Behavior,* 5: 118-132.

_____ (1993b) "Proportional r presentation, approval voting, and coalitionally straightforward elections", in Bernett et al. (1993): 107-133.

Nalebuff, Barry J. and Joseph E. Stiglitz (1983) "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *Bell Journal of Economics* 14: 21-43.

Newman, Andrew F. (1991) "The Capital Market, the Wealth Distribution and the Employment Relation", Discussion Paper No. 951, Northwestern University.

Niskanen, W. A. Jr. (1971) *Bureaucracy and Representative Government,* Chicago.

Oates, William (1972) *Fiscal Federalism,* New York.

_____ (1977) (ed.) *The Political Economy of Fiscal Federalism,* Lexington Books.

_____ (1985) "Searching for Leviathan: An Empirical Study", *American Economic Review* 75: 748-757.

Persson, Torsten and Guido Tabellini (1992) "Federal Fiscal Constitutions. Part 1: Risk Sharing and Moral Hazard", Discussion Paper, Stockholm University, Stockholm.

_____ and _____ (1993) "Federal Fiscal Constitutions. Part 2: Risk Sharing and Redistribution", Discussion Paper, IGIER.

Pestieau, Pierre (1977) "The Optimality Limits of the Tiebout Model", in Oates (1977).

Piketty, Thomas (1993) "A Federal Voting Mechanism to Solve the Fiscal Externality Problem", mimeo, DELTA, Paris.

Prud'homme, Rémy (1995) "The Dangers of Decentralisation", *The World Bank Research Observer* 10: 210-220.

Qian, Y. and G. Roland (1994) "Regional Decentralization and the Soft Budget Constraint: the Case of China", mimeo, Ecare, Brussels.

Rasier (1995) *Communist Economies and Economic Transformation* 7(2).

Riker, William (1964) *Federalism: Origin, Operation and Significance,* Boston.

Roland, Gérard (1994) "The Role of Political Constraints in Transition Strategies", *Economics of Transition* 2(1): 27-41.

Romer, T. and H. Rosenthal (1979) "Bureaucrats Versus Voters: On the Political Economy of Resource Allocation by Democracy", *Quarterly Journal of Economics* 93: 563-87.

_____ and _____ (1995) "A Theory of Divided Government", *Econometrica,* forthcoming.

Rostow, Walt W. (1960) *The Stages of Economic Growth: A Non-Communist Manifesto,* Cambridge.

Sachs, Jeffrey (1991) "Sponaneous Privatization: A Comment", *Soviet Economy* 7(4): 317-321.

_____ and Wing Thye Woo (1993) "Structural Factors in the Economic Reforms of China, Eastern Europe, and the Former Soviet Union", *Economic Policy* 9(18): 101-145.

Salmond, Pierre (1987) "Decentralisation as an Incentive Scheme", *Oxford Review of Economic Policy,* 3 (2, Summer): 24-43.

Samuelson, Paul A. (1954) "The Pure Theory of Public Expenditure", *Review of Economics and Statistics* 36: 387-389.

Schmidt, Klaus (1991) "The Costs and Benefit of Privatisation", Ph. D. thesis (chapter 4), University of Bonn.

Schumpeter, Joseph A. (1934) *The Theory of Economic Development; An Inquiry Profits, Capital, Credit, Interest and the Business Cycle,* Cambridge, Mass.: Harvard University Press.

_____ (1936) *A Theoretical, Historical and Statistical Analysis of the Capitalist Process,* abridged edition, New York: McGraw Hill.

_____ (1954) *History of Economic Analysis,* Oxford: Oxford University Press.

Seabright, Paul (1994) "Accountability and Decentralisation in Government: An Incomplete Contracts Model", forthcoming in *European Economic Review.*

Sen, Amartya K. (1966) "Peasants and Dualism With or Without Labor Surplus", *Journal of Political Economy* 74: 425-450.

Shleifer, A. (1985) "A Theory of Yardstick Competition", *The Rand Journal of Economics* 16: 319-327.

Slutsky, Steven M. (1986) "Elections With Incomplete Information. Comments on the papers of Coughlin and Ferejohn", *Public Choice* 50: 105-130.

Standing, Guy (1978) "Aspiration Wages, Migration and Urban Unemployment", *Journal of Development Studies* 14(2): 232-248.

Stiglitz, Joseph E. (1990) "Peer Monitoring and Credit Markets", *World Bank Economic Review,* 351-366.

Tiebout, Charles M. (1956) "A Pure Theory of Local Public Expenditure", *Journal of Political Economy* 64: 416-424.

Tocqueville, Alexis de (1835) *Democracy in America,* Vol. I, translated by George Lawrence, edited by J. P. Mayer, London.

Todaro, Michael P. (1969) "A Model of Labor Migration and Urban Unemployment in Less Developed Countries", *American Economic Review* 59: 138-148.

Townsend, Robert (1994) "Risk and Insurance in Village India", *Econometrica* 62(3): 539-591.

Varian, Hal R. (1990) "Monitoring Agents with Other Agents", *Journal of Institutional and Theoretical Economics,* 153-174.

_____ (1992) " A Solution to the Problem of Externalities when Agents are Well-Informed", mimeo, Michigan.

Weiss, Andrew (1980) "Job Queues and Layoffs in Labour Markets with Flexible Wages", *Journal of Political Economy* 88: 526-538.

Weldon, J. (1966) "Public Goods (and Federalism)", *Canadian Journal of Economics and Political Science* 32: 231ff.

White, Gordon (1982) "Urban Employment and Labour Allocation Policies in Post-Mao China", *World Development* 10: 613-632.

_____ (1987) "The Changing Role of the Chinese State in Labour Allocation: Towards the Market?", *The Journal of Communist Studies* 3: 129-153.

_____ (1988) "State and Market in China's Labour Reform", *Journal of Development Studies* 24: 180-202.

Wildasin, David E. (1987) "Theoretical Analysis of Local Public Finance", in Edwin S. Mills (ed.) *Handbook of Regional and Urban Economics,* Vol. II, Amsterdam.

_____ (1988) "Nash Equilibria in Models of Fiscal Competition", *Journal of Public Economics* 35(2): 229-240.

_____ (1989) "Interjurisdictional Capital Mobility: Fiscal Externality and a Corrective Subsidy", *Journal of Urban Economics* 25:193-212.

_____ (1991) "Some Rudimentary 'Duopolity' Theory", *Regional Science and Urban Economics* 21: 393-421.

_____ (1992) *Urban Public Finance,* New York.

Xu, Chengang (1993) "Risk aversion, Rural-Urban Wage Differentiation and Migration", Centre for Economic Performance, Discussion Paper No. 108, London School of Economics, London.

Zax, Jeffrey (1989) "Is there a Leviathan in Your Neighbourhood?", *American Economic Review* 79: 560-567.