Principled Agents?
The Political Economy of Good Government

Timothy Besley
London School of Economics and Political Science

October 2005
# Contents

0.1 Preface ...................................................... vi

1 Competing Views of Government ........................................ 1
   1.1 The Issues .................................................. 1
   1.2 This book ................................................... 2
   1.3 Background .................................................. 3
      1.3.1 The Size of Government ................................. 4
      1.3.2 Corruption .............................................. 5
      1.3.3 Property Rights ........................................ 8
      1.3.4 Trust and Turnout ..................................... 9
   1.4 Economic Policy Making ...................................... 10
      1.4.1 Foundations ............................................. 11
      1.4.2 Good Policies .......................................... 13
      1.4.3 The Public Choice Critique of Welfare Economics ....... 15
   1.5 Political Economy ........................................... 15
   1.6 Incentives and Selection in Politics ......................... 23
   1.7 Concluding Comments ....................................... 29

2 The Anatomy of Government Failure .................................. 31
   2.1 Introduction ............................................... 31
   2.2 Three notions of Government Failure ......................... 33
      2.2.1 Pareto Inefficiency .................................... 34
      2.2.2 Distributional Failures ................................ 34
      2.2.3 Wicksellian Failures .................................... 37
      2.2.4 Comparisons ............................................ 37
   2.3 An Example: Financing a Public Project ....................... 38
      2.3.1 Private Provision ....................................... 39
      2.3.2 Government Provision ................................... 40
   2.4 Sources of Government Failure ................................ 41
      2.4.1 Ignorance ............................................... 42
      2.4.2 Influence ............................................... 43
      2.4.3 The Quality of Leadership .............................. 49
2.5 Sources of Political Failure ........................................ 51
  2.5.1 Voting ....................................................... 51
  2.5.2 Log-rolling and Legislative Behavior ....................... 53
2.6 Dynamics ......................................................... 56
  2.6.1 Investment Linkages ....................................... 57
  2.6.2 Political and Policy Linkages ............................... 60
  2.6.3 Investment and Politics .................................... 67
2.7 Implications ..................................................... 69
2.8 Concluding Comments ........................................... 73

3 Political Agency and Accountability ............................. 75
  3.1 Introduction .................................................... 75
  3.2 Elements of Political Agency Models .......................... 79
    3.2.1 The Nature of the Uncertainty ............................. 79
    3.2.2 The Motives for Holding Office ........................... 80
    3.2.3 The Nature of Accountability ............................. 81
    3.2.4 Retrospective Voting ..................................... 81
    3.2.5 Model Types ............................................. 82
  3.3 The Baseline Model ............................................ 83
    3.3.1 The Environment .......................................... 84
    3.3.2 Equilibrium ............................................. 85
    3.3.3 Implications ............................................. 86
  3.4 Extensions .................................................... 93
    3.4.1 Polarization and Competition .............................. 93
    3.4.2 Information and Accountability ........................... 97
    3.4.3 The Nature of the Distortion ............................. 102
    3.4.4 Within-Term Cycles ...................................... 106
    3.4.5 Multiple Issues .......................................... 108
    3.4.6 Multiple Two-Period Terms ................................ 112
    3.4.7 Indefinite Terms ......................................... 117
    3.4.8 Multiple Agents ......................................... 120
  3.5 Discussion .................................................... 126
    3.5.1 Civic Virtue and the Quality of Government ................ 126
    3.5.2 Decentralization versus Centralization .................... 127
    3.5.3 Autocracy versus Democracy ................................ 129
    3.5.4 Accountability to Whom? ................................ 130
    3.5.5 Wage Policies for Politicians ............................ 131
    3.5.6 Behavioral versus Rational Choice Models ................ 132
  3.6 Concluding Comments .......................................... 132
CONTENTS

4 Political Agency and Public Finance 135
  4.1 Introduction 135
  4.2 The Model 138
  4.3 Three Scenarios 139
    4.3.1 Pure Adverse selection 140
    4.3.2 Pure Moral Hazard 141
    4.3.3 Combining Moral Hazard and Adverse Selection 144
  4.4 Implications 146
    4.4.1 Equilibrium Voter Welfare 147
    4.4.2 Are Good Politicians Necessarily Good for Voters? 149
    4.4.3 Turnover of Politicians 150
    4.4.4 The Spending Cycle 151
  4.5 Restraining Government 152
    4.5.1 A Direct Restraint on the Size of Government 155
    4.5.2 Indirect Restraints 156
    4.5.3 Summary 163
  4.6 Debt and Deficits 164
  4.7 Governments versus NGOs 168
    4.7.1 Framework 169
    4.7.2 Aid to the Government 171
    4.7.3 Comparisons 172
    4.7.4 Further Issues 174
  4.8 Competence 175
  4.9 Conclusions 176

5 Final Comments 179
0.1 Preface

This book began life as the Lindahl lectures which I delivered at the University of Uppsala in October 2002. I remain honored to have been asked to deliver these lectures. Erik Lindahl developed some powerful ideas in public economics and political economy. He represents the spirit in which this book is written – using ideas in political economy and public economics to understand how to make the world a better place. I am particularly grateful to Soren Blomquist and Bertil Holmlund for their hospitality during my visit and their patience and understanding for the delay in writing up these ideas. I am also grateful to the faculty and students at Uppsala who made my visit such a stimulating one.

The main reason for the delay in writing up the lectures was a creeping ambition to develop some of the ideas more deeply and systematically. For example, it became clear to me that the agency model of politics was less well-understood than I had previously thought and the opportunity to develop a more systematic exposition of its potential was too tempting. The manuscript also presents an opportunity to make a statement on the divide between the so-called benevolent dictator view of the government and the public choice view. Having been schooled in one as a graduate student, I have since been on a journey which has tried to square my belief in the importance of sound (and not necessarily minimal) government with the self-evident proposition that government office is frequently abused. It has become clear to me that incentives in government are first order and should be taught as part of any course in public economics. To the original lectures, I also added a more refined chapter on the idea of government failure.

This book is intended as a contribution to the burgeoning field of political economy – combining ideas from economics and political science. I have tried to keep technicalities to a minimum. But some familiarity with the basic tools of micro-economics is necessary to appreciate some of the arguments. The coverage is far from definitive. The field is still evolving fast. The treatment here focuses mainly on the agency approach to politics which takes asymmetric information in politics seriously. It pays much less intention to spatial models.

I was supremely fortunate in having the opportunity to present some of the ideas in this book to the “Encounters with Authors” series organized by the Harvard Center for Basic Research in the Social Sciences (CBRSS). I am particularly grateful to Jim Alt and Ken Shepsle for their invitation and hospitality. I also benefitted from their numerous insights on this project that they offered while I was there. The participants in that workshop were extremely supportive, giving me countless ideas and helping me to avoid infelicities. I am especially grateful for the feedback that I received at the workshop from Daron Acemoglu, Alberto Alesina, Abhijit Banerjee, John Londregan, Arthur (Skip) Lupia, Adam Meirowitz, Sharun Mukand, Jim Snyder and Enrico Spolaore.

Many people have influenced my thinking on the topics of this book. Throughout my career I have enjoyed the counsel and, above all, friendship of Steve Coate. He
and I have collaborated on a large number of projects and many ideas developed here (especially the good ones!) are probably his. However, they have been etched into my mind by years of discussion with Steve so that I no longer recognize this. Anne Case collaborated with me in my first efforts in political economy when we stumbled on the agency approach while trying to understand determinants of re-election rates among U.S. Governors. Anne has taught me many things which are relevant to the ideas in this book and the part of chapter three on U.S. Governors is based on our joint work. Michael Smart has also collaborated with me in understanding the agency model in public finance settings. Chapter four of this book reports mostly on joint work with him. His work with Daniel Sturm also greatly influenced the formulation of the agency problem that I use in chapter three.

The LSE is a unique place and I have been privileged to serve as a member of the economics department and of the Suntory-Toyota International Centres for Economics and Related Disciplines (STICERD) since 1995. It has provided me colleagues whose curiosity, integrity and energy are second to none. I especially thank my STICERD-based colleagues Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, Markus Goldstein and Andrea Prat with whom I have had many helpful conversations relevant to this book as well as receiving comments on parts of it.

There are many others whose comments and encouragement have helped to improve my understanding of the issues dealt with here. I particularly thank Daron Acemoglu, Bruno Frey, Jim Hines, Clare Leaver, Gilat Levy, Stephen Morris, Roger Myerson, Torsten Persson, Jim Robinson, Paul Seabright, Andrei Shleifer, Daniel Sturm, and Guido Tabellini. Konrad Burchardi, Marieke Huysenstruyt, Masa Kudamatsu, Marit Rehavi and Silvia Pezzini provided excellent research assistance at various points. Masa deserves special mention; he read the entire manuscript carefully and made a host of important comments which lead to substantial improvements.

Last, but not least, I thank my wife, Gillian, and my sons Thomas and Oliver for their forbearance and support.

London
October 2005
Chapter 1

Competing Views of Government

“The aim of every political Constitution, is or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of society; and in the next place, to take the most effectual precautions for keeping them virtuous whilst they continue to hold their public trust.” – James Madison (The Federalist Papers, LVII)

1.1 The Issues

Economic analyses of government divide into two broad camps. One emphasizes government in the public interest. It outlines the range of activities that government can undertake to improve the lives of its citizens. Government provides the underpinnings of the market system by establishing property rights and a means of adjudication through the courts. Government can regulate externalities which private actions fail to internalize. Government can provide public goods which the market will tend to under-provide. Government can regulate abuses of market power in cases where competition is limited. Finally, government can distribute resources towards socially favored groups. The logic behind this has been developed at length and provides the modern theory of the state from a welfare economic point of view.2

At the other extreme are accounts of government seen mainly as a private interest. Government can be a focus for rent seeking in which the power to tax results in private, wasteful efforts to capture the state which then rewards the powerful at the expense of citizens at large. Even in fulfilling its apparently virtuous functions, government can be influenced by organized groups and state officials may receive bribes to act against the interests of the citizens at large. Government may lack sufficient incentives for officials (who are less than publicly spirited) to act in the public interest. The latter may choose

---

1Madison (1961).
2See, for example, Atkinson and Stiglitz (1980).
to divert resources to their own ends or simply slack on the job.\textsuperscript{3}

This book strives to understand a view of government and its potential that lies between these two extremes. Its intellectual origins lie in the so-called Publius view which is today often associated with the name of James Madison, one of the authors of the \textit{Federalist Papers} and an architect of the U.S. Constitution. This view recognizes the potential for government to act in the public interest but understands the tendency for things to go awry. Good government is in part associated with designing an institutional framework which affects the incentives of those who make policy decisions. But good government is not entirely about incentives; it also requires good leaders – persons of character and wisdom. These twin elements of incentives and selection in politics are central themes throughout.

Making progress on these questions requires models of how government allocates resources. At the heart of effective government lies the solution to a principal-agent problem between citizens and government. This approach is a focal point for this book. However, it is located more broadly in the modern political economy literature which develops theoretical and empirical tools for understanding public resource allocation.

The book shares with the Publius view the notion that there are institutional pre-conditions to effective government. But finding them requires an understanding of incentives and the process by which the political class is selected. It is broadly optimistic – good government \textit{is} possible, but only when the preconditions are right.

\subsection*{1.2 This book}

This book comprises four relatively self-contained essays. This first essay provides an introduction to issues studied here and discussions of government from a political economy perspective. We will place this in its historical context and the traditional public economics approach to the role of government. We will also discuss some of the empirical evidence on differences in the quality of government across countries that motivates these ideas. This essay also discusses the themes of incentives and selection in politics in general terms.

Chapter two is an essay on government failure. Most economists now agree that the idea of government failure needs to be placed alongside the idea of market failure in our discussions of government intervention. However, unlike market failure, there is no agreed upon definition of government failure. The economics literature is also obscure on which aspects of government failure are intrinsic to the fact that government has a monopoly of coercive power and which aspects are a consequence of democratic political competition. Chapter two provides an overview of these ideas and develops some simple economic examples to illustrate the main ideas.

\textsuperscript{3}This view has been at the core of the Public Choice. In the context of public finance, see Buchanan (1967).
1.3. BACKGROUND

In chapter three, we turn to political agency models. These explore the consequences of limited information in politics. The chapter develops a canonical model which is then used to discuss an array of issues that arise in such models. As well as providing an introduction to the literature, the chapter also throws up a variety of issues that have not been fully understood to date and merit further work. The chapter also emphasizes the empirical potential of these models in explaining real world policy choices. At the heart of the models is the notion of political accountability—a frequently used concept that often lacks precise usage. The political agency model is an ideal vehicle for thinking about this idea and exploring how accountability works in practice.

Chapter four applies the political agency model to public finance issues reporting on joint work with Michael Smart. The chapter develops a simple and tractable model which can be applied to thinking about the determination of taxes, debt and public spending in an agency framework. The model is used to think about the merits of restraining government in an agency model. It also discusses how agency models can provide insights into the choice between NGOs and government in delivering public services.

Chapter five offers a few concluding comments and observations. One of its main themes is the need to devote more attention to issues of selection in political economy models and it introduces some of the main ideas that are relevant to this.

1.3 Background

The discussions in this book centre around policy making in representative democracies. This achieves special poignancy in view of the fact that there is now an unparalleled consensus about the centrality of the market economy and some form of liberal democracy in allocating public and private resources. Indeed, much of the world is now committed to a broadly capitalist model of production regulated and augmented by a popularly elected government. It was not always thus. From 1945 to around 1990, the world was divided into two broad economic systems. On one side were the planned economies based on total subordination of the economy to the state via economic planning. On the other side were the mixed economy with private goods production dominated by the market and public goods production in government. This economic cleavage was largely paralleled by a political cleavage—planned economies being largely autocratic with most mixed economies committed to some form of liberal democracy. The latter swing towards democracy in recent years is illustrated in Figure 1.1 which uses the Polity IV data set. There is a downward trend from the 1950s through to the mid 1970s with a real “take-off” around 1990. The latter is due in part, but not exclusively, to the fall of the Berlin Wall.

Figure 1.1 about here.
CHAPTER 1. COMPETING VIEWS OF GOVERNMENT

While Democratic values have triumphed in recent years, there is still a huge amount of discussion about the impact of different institutional choices. Moreover, there are many debates about the consequences of democratic policy making for economics. The dominant early traditions in political economy emanating from Virginia and Chicago offered competing analyses of how democracy generates policies and whether there are endemic forms of “democratic failures” that need to be avoided. The key challenge remains understanding how government resource allocation works and uncovering the institutional preconditions to effective democratic governance.

1.3.1 The Size of Government

The remarkable growth in government that occurred in the twentieth century brings the two competing perspectives on government that we began with into sharp relief. Government growth could be seen as confirmation that government is “out of control” with insufficient safeguards in place to restrain its power to tax. However, it could equally be viewed as confirmation of the effectiveness of government in acting in the common interest.

Table 1.1 reproduced from Maddison (2001) illustrates this. Whereas government in the early part of the century consumed only around 10% of GDP, it consumed more than 40% by the end of the century. The trend holds also in Western Europe, Japan and the United States, even though the latter has not adopted the full scale “welfare state”.

Table 1.1 about here

Figures 1.2 and 1.3 look at these patterns from 1960 onwards. They graph the size of government as measured by the percentage of government consumption in GDP in both real and nominal terms, separating out the experience of the (rich) OECD countries and the rest. These figures show that government is larger in rich than poor countries. The size of government in nominal terms has, however, grown as a share of national income in both countries, although with signs of a leveling off towards the end of the period. The graph in real terms shows that the trend is largely flat in rich countries while the drift upwards is still seen in the non-OECD sample. This suggests that the increase in the size of government in nominal terms in largely due to an increase in the cost of providing a given bundle of services along the lines of the unbalanced growth story suggested by Baumol (1967). He argued that government growth in rich countries is significantly affected by the fact that it provides labour intensive services which have not benefitted from labour saving technological change. Borcherding (1985) estimates that around 31% of the rise in the size of government can be explained this way.

---

4See Acemoglu and Robinson (2005) for some important insights about the differences between democratic and non-democratic government.

5Data from the World Development Indicators is only available from 1960 onwards.
1.3. BACKGROUND

How democratization has affected the growth of government is moot. From a theoretical point of view, democracies should put more weight on popular opinion. Acemoglu and Robinson (2005) model democracy as a system of government in which the poor have greater influence which leads to greater redistribution. Mulligan, Gil and Sala-i-Martin (2004) argue that the data give little support to the proposition that democracies have substantially different economic and social policies from non-democracies. However, their data end in 1990, before the important recent increase in the incidence of democracy. In addition, while the Polity IV data may be useful for looking at broad trends, they do a poor job in modeling policy incentives in the aggregate.

Figures 1.4 and 1.5 look at the size of government in democracies and autocracies as measured in the POLITY IV data. They show that there is little difference in terms of government size between democracies and autocracies. The difference in real terms prior to 1990 is largely an income effect – reflecting the larger difference in income levels between democracies and autocracies in this period.

Perhaps the most persuasive lesson from recent research is that democracy comes in many forms and looking for an effect of democracy per se is probably misguided. The important work of Persson and Tabellini (2003) shows that there are distinct patterns in terms of the form of government and policy outcomes. Most notably, there is a robust relationship between both parliamentary (rather than presidential) government and proportional representation (rather than majoritarianism) and the size of government.

The growth of government to its modern proportions constitutes a remarkable act of public trust. Government spends a significant fraction of national income and is expected to do so wisely. Such trust can be supported in one of two ways – either we are confident that government is populated by publicly spirited officials or else there must be sufficient safe-guards in place to curtail self-interested behavior. Either way, it does seem remarkable that this position has changed so markedly over a hundred years. This faith in democracy to deliver benefits in return for taxes, that are remarkably high in any long-run perspective, is a core part of the institutional consensus that has emerged.

1.3.2 Corruption

While it is debatable whether the size of government tells us much about government quality, there is one aspect of government quality that has received a lot of empirical attention, namely corruption. While there is still debate about how far corruption is damaging to the economy as a whole, it is widely agreed to be an important symptom
of low quality of government. That said, corruption is often used _sui generis_ to stand for a variety of distinct problems in state performance, many of which should require separate analyses. One important distinction is that between bureaucratic and political corruption.\(^6\) The institutional structures that perpetuate these may be related, but can also be quite different. For example, the incentive mechanisms for the control of bureaucrats and politicians tend to be rather different.\(^7\)

As we shall see in the next chapter, there are a number of ways to think about the costs of corruption from a normative point of view. Corruption could be viewed primarily as a distributional problem as those who receive illicit returns from the state are rarely socially deserving, i.e. it is pretty hard to think of a social welfare function that would value transfers between public officials and citizens that arise as the product of corruption.

Corruption also leads to resource misallocation. There are many different ways in which this can happen. For example, corruption could lead to misallocation of investment and public infrastructure away from their most productive use. It can also lead to misallocation of talent as self-interested individuals seek rewards in occupations whose returns are inflated by corrupt activities.\(^8\) Corrupt government may also lead to more self-interested individuals choosing to enter public life in order to capture rents.

While corruption is not easy to measure, there are a number of measures now available, the main ones being those from the International Country Risk Guide (ICRG) and Transparency International. A mini-industry has sprung up that correlates corruption with various historical and institutional factors.\(^9\) Some useful empirical regularities have emerged.

Whether as cause or effect, corruption is associated with low income – see Mauro (1995). While it is extremely difficult to establishing the direction of causation here, it is clear that part of being a poor country is having a lower quality government. Moreover, this observation has affected debates about how to promote economic development. The World Bank and other organizations have push a “governance agenda” which puts weight on combating corruption as means of improving government performance throughout the developing world.

Figures 1.6 and 1.7 illustrate these two claims. Figure 1.6 graphs the ICRG corruption measures between the mid 1980s and late 1990s. Over the entire period rich countries had significantly lower corruption levels.\(^10\) As illustrated in Figure 1.7, corruption is also lower in democracies compared to autocracies over the same period.

---

6. Hellman, Jones and Kaufmann (2000) distinguishes "state capture" (corruption to change the laws) from "administrative corruption" (corruption to alter the implementation of laws), and empirically show that the type of firms engaging in state capture is different from that in administrative corruption.

7. See Alesina and Tabellini (2005) for a recent analysis of this.

8. See, for example, Murphy, Shleifer and Vishny (1991).


10. A high score on this index denotes low corruption.
1.3. BACKGROUND

In support of this, there are a number of strong correlations between corruption and other measures of government performance. La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999) investigated the empirical links between corruption and (i) ethno-linguistic fractionalization, (ii) religion, and (iii) legal origins. They find that less diverse societies are less corrupt, as are those with a larger fraction of Protestants. Common law legal origin is also correlated with lower corruption. Even setting aside the question of causation, the policy consequences of these correlations is far from clear. However, they motivate the need to treat quality of government as endogenous and to explore how structural factors affect government performance through their impact on incentives and selection. Some of our findings in chapter 3 which look at determinants of government performance in an agency model are helpful in thinking this through from a theoretical point of view.

The link between openness and corruption has also generated interest. For example, Ades and DiTella (1999) argue that more countries whose economies are more open to trade are less corrupt. This type of analysis is expanded and developed in Bonaglia et al (2001). The main theoretical argument that these studies make is that openness acts as a constraint on the behavior of politicians. In chapter four, we locate such arguments in broader terms, using the political agency model to investigate the quality of government when there is tax competition between nations.

Corruption has also been related to freedom of the media by Brunetti and Weder (2003) and Ahrend (2000) who relate press freedom and corruption in cross-country data. Both of these papers find that press freedom is associated with lower levels of corruption. Djankov, McLeish, et al. (2003) focuses more directly on the effect of media ownership patterns on a variety of outcomes. They develop a remarkable data set on media ownership patterns in 98 countries to do so. The find that corruption is related to state ownership of newspapers. We show in chapter three that such findings can be understood in a political agency framework where press freedom affects the extent of information available to voters.11

Corruption has also been related to political institutions and political outcomes. For example, Persson, Tabellini and Trebbi (2001) consider the link between corruption and political/institutional variables. They test the idea that majoritarian systems and larger voting districts are less prone to corruption, finding strong evidence in favor of this. Triesman (2000) and Persson, Tabellini and Trebbi (2001) document the fact that corruption is negatively correlated with political turnover. In chapter 3, we will return to the links between political turnover that are suggested by the political agency framework which treats both turnover and corruption as endogenous.

Empirical studies of corruption provide one of the most thoroughly researched areas of government quality. That said, the links to theory tend to be quite limited. This

---

11 This is based on Besley and Prat (2004).
CHAPTER 1. COMPETING VIEWS OF GOVERNMENT

shows up the choice of what is treated as endogenous or exogenous. For example, the rate of turnover among politicians is often treated as an exogenous variable whereas many theoretical approaches would determine this endogenously. There are also great difficulties in establishing the direction of causation in cross-country data. Nonetheless, it brings into sharp relief the proposition that, even measured in this crude way, there are significant differences between government quality across the world that cry out for explanation.

1.3.3 Property Rights

Enforcement of property rights has been studied extensively. To motivate this, observe that at the heart of the Weberian view of the state is the notion that the state has a monopoly on the legitimate use of coercion. This coercive power can be used in a benign way to raise taxes to finance public goods and regulate externalities. However, that same power can be used to expropriate wealth. A key aspect of government quality concerns the extent to which this is done. Without suitable restraints on the power of government, it is harder to foster a climate for private investment incentives.

This problem of expropriation arises when government is too strong and overbearing. Alongside government expropriation of wealth is private expropriation in situations in which the government is too weak. In such cases, private contracts may not be enforced which will also have a deterrent effect on trade and investment leading to lower income per capita. Here, the problem is one of private rather than public predation. The role of government is to support market arrangements between private actors and to uphold contracts.

In light of these observations, Djankov, Glaeser et al (2003) see the role of effective government as steering a path between the two evils of authoritarianism and disorder. Effective institutional solutions are those that create the right balance between them. It is clear that political incentives shaped by political institutions provide only a partial picture in understanding these issues. Another key issue is whether there is an effective legal system with sufficient judicial independence to curtail predation by government and uphold private contracts. Djankov, Glaeser, et al (2003) review the extensive empirical evidence relating market performance to the operation of legal systems.

As in the case of corruption, there is a strong correlation between income per capita and enforcement of property rights. This is illustrated in Figure 1.8 using the ICRG measure of “risk of expropriation” by government. The difficult issue is again to establish the direction of causation. In a recent paper, Acemoglu, Johnson and Robinson (2001) suggest the use of settler mortality in colonial times as an instrument for property rights enforcement, arguing that this has lead to institutional legacies. They argue on the basis of this that causation runs from poor property rights to low income. Moreover, the quality of government, in so far as it contributes to property rights enforcement, has long-lived historical roots.
1.3. BACKGROUND

The work, however, says little directly about the exact theoretical route by which this effect is mediated. While Figure 1.9 shows that there is a predictable link between protection of property rights and democracy, it is unclear whether it is causal. In an effort to get at this, Persson (2004) relates settler mortality to modern day political institutions using Acemoglu, Johnson and Robinson (2001)'s variable as an instrument for the form of democracy. He finds that parliamentary (as opposed to presidential) democracies\(^{12}\) and proportional representation (as opposed to majoritarian) systems are growth promoting. In tune with the Acemoglu, Johnson and Robinson (2001) results, these government types also tend to have better protection of property rights.

Even though political institutions are only part of the story, the structure of government incentives and the quality of the political class are an important input into aims to improve the functioning of a market economy, especially in the creation of secure property rights. The models and analysis in chapters three and four which study principal-agent problems are therefore a useful input into this aspect of government quality.

1.3.4 Trust and Turnout

The pattern of democratization illustrated in Figure 1.1 suggests that liberal democracy is riding high as an institution since around 1990. However, there is an underlying concern about the declining health of the more established democracies.\(^{13}\) One major symptom of this is the observed decline in voter turnout.\(^{14}\) This is illustrated for OECD countries in Figure 6 which documents this trend over time by plotting turnout over time for OECD countries. This fell from an average of 84% in 1946-50 to 62% in 1996-2000, the most democratic decline being in the latter part of the period. This OECD experience is mainly driven by the older democracies. A larger sample of countries which includes the new democracies of central and eastern Europe which have enjoyed very high participation rates.

In general, there is a difficulty in assessing the welfare consequences of declining turnout in the absence of any agreed upon theory of voting. The low likelihood that any voter will be decisive in mass elections makes models based on the probability of being pivotal, a questionable basis for the theory of voting. As discussed by Aldrich (1997), political scientists tend to work in frameworks where some extra component of

\(^{12}\) Parliamentary systems that have a vote of confidence procedure for the executive.

\(^{13}\) See, for example, Pharr, Putnam and Dalton (2000).

\(^{14}\) See Levi and Stoker (2000) for a wide-ranging discussion of these issues.
utility (such as social duty) is invoked to explain why people vote in such large numbers. In this case, election turnout could be a barometer for how such feelings of duty extend in the population. To the extent that these are correlated with perceived satisfaction with government, this could create a link between turnout and the quality of government, but the link is tenuous at best. If social duty is the main basis of deciding to vote, then declining turnout could also be linked to a general decline in “social capital”, i.e. a willingness by citizens to privately provide public goods.

Other theories of voting put weight on the role of being informed in affecting the decision to turnout – see Feddersen and Pesendorfer (1996). If elections were purely a common values problem, then nobody would mind if only one voter showed up and caste their vote decisively provided that individual was informed. This argument has less force when there are important ideological divergences between groups of voters. This is a particular issue if the informed voters are mainly from one particular ideological grouping.

To get a feel for whether turnout and declining voter satisfaction with politics and politicians are linked requires more than evidence on turnout alone. The National Election Surveys (NES) from the United States provide an interesting perspective. They ask a variety of interesting questions about attitudes to government which can be correlated with the indicators of what governments do and the pattern of turnout. Of particular interest is the composite variable which gauges trust in government over the period 1958-2002. It is, however, important to bear in mind that these are subjective assessments and there is always the possibility that they are a poor reflection of actual government performance. Figure 1.10 shows a declining trust in government. This figure also looks at turnout and shows that these two are broadly moving in the same direction – both trust and turnout having decline over the period. Trust showed some evidence of recovery towards the end of the period. While only suggestive, this illustrates why declining trust is often linked to declining turnout in many discussions.

Figure 1.10 about here.

1.4 Economic Policy Making

The two views of government discussed above have their parallels in approaches to the study of economic policy. The benevolent government view which focuses on the good that government can do has important figures such as Arthur Pigou, James Meade and James Mirrlees in its ranks. The more cautious view is associated with equally distinguished figures such as James Buchanan, George Stigler and Anne Krueger. The latter place political economy issues at the heart of their approach.

Among the latter group, it is Buchanan who has developed at length the importance
of using constitutional rules to improve the workings of democratic institutions. As we noted above, this is the spirit of the approach used by the Founding Fathers of the United States as expressed in the Federalist papers. This approach argues that there is a need for both a procedural constitution – rules that define the workings of democratic institutions, and a fiscal constitution – rules that define constraints on policy making.

The main purpose of this section is to survey in brief the main elements of the normative approach to policy making. This is important for the discussion in chapter two which tries to understand the idea of government failure rigorously. It is also important as a backdrop to the discussion of political economy issues in policy making.

1.4.1 Foundations

Making sense of good government requires a performance metric. The standard economic notion is social welfare. This is usually invoked in a context where the “inputs” into social welfare are individual utility levels. (The idea that social welfare is solely a function of such data is known as welfarism.) The task is then to provide a means of aggregating these utility levels to form a societal measure of well-being. Once a criterion for social welfare is established, then policies and other social states can be evaluated and ranked. From this, the notion of good government can be defined.

This approach can be applied to policies, political processes and institutions. It provides an intellectual underpinning for ideas of government operating for the common good or in the public interest. Its historical antecedents include classical Utilitarianism as well as Rawls’ theory of justice. But some have taken issue with the coherence of the notion of the common good and hence its usefulness for thinking about good government.

There are two levels of attack. The first questions the very idea of the common good. This would include concerns following the discovery of Arrow’s impossibility theorem. Arrow (1951) examined the possibility of deriving social welfare from individual preferences in the absence of interpersonal comparisons of utility. By producing his impossibility theorem, his work appeared to be a significant set-back in the quest for giving a scientific foundation for social welfare and hence for notions of common good. However, as Sen (1977) has emphasized, one key issue concerns the assumptions made about inter-personal comparisons. If interpersonal comparisons are allowed, social welfare functions can be derived which satisfy Arrow’s axioms. These include Utilitarianism and Rawlsianism.

The second attack on the notion of common good views notions of social welfare, even if logically coherent, as inherently contested. The choice of a social welfare function

---

15 This is particularly true in his joint work with Geoffrey Brennan – Brennan and Buchanan (1980, 1985).
16 These themes are taken up in Chapter 4.
will hinge to a significant degree of value judgements – for example notions of inequality aversion – on which no consensus prevails. The problem is then that there are too many competing notions of the good for an idea of common good to make sense. While the social choice literature has produced elegant axiomatizations of certain kinds of social welfare functions, they are not a particularly helpful way of resolving this issue on a practical level. It is clear that the practical question is whether there are procedures that are regarded as reasonable ways of making collective choices, especially when it allocates coercive power to particular individuals. The remarkable thing about representative democracy is how readily it has been accepted as a legitimate way for this to be done.

The fact that there is no readily agreed upon social welfare function for making social decisions is important for the issue of defining good government. Such concerns are echoed in Schumpeter (1943) who rejected classical notions of democracy that saw the aim of democratic institutions as trying to find ways of pursuing the common good. He recast the problem of democracy in terms of the positive problem of understanding how candidates competed to secure votes, eschewing normative comparisons entirely.

One option is to work without any distributional judgements – focusing exclusively on the implications of Pareto efficiency. As we shall discuss further below, the standard economic model puts weight on Pareto efficiency as a criterion for good policy. The use of Pareto efficiency is broadly uncontroversial in standard economic models. However, it can conflict with other principles that many have invoked as constraints on policies. Respect for individual liberty is one important example. As argued persuasively in Sen (1970), a policy can be Pareto efficient and yet violate the most minimal notions of individual liberty. In fact any side constraint on welfare is likely to have the property of picking Pareto inferior policies (see Kaplow and Shavell (2001)).

Even if this concern about Pareto efficiency as a normative criterion is set aside, there is still the problem that it provides a very weak and ambiguous idea of the common good. There are typically many different Pareto efficient policies with varying degrees of inequality associated with them. It may be an achievement of sorts that policies or institutions guarantee Pareto efficiency, but many policy outcomes become non-comparable on these terms. Having a criterion which allows distributional comparisons would certainly provide a richer theory of good government. But that would require invoking some kind of social welfare function.

Even though it is hard not to have some sympathy with these difficulties, a wholly nihilistic take on normative metrics that invoke some distributional criteria for good government is probably too strong. First, some aspects of distribution, such as helping the poor or particular disadvantaged groups are frequently viewed as broadly desirable. Second, narrowly targeted policies, those that lead to the enrichment on small elite groups are frequently frowned upon and the language of common good can be powerful in getting them thrown out. Indeed, more generally, the calculus of gainers and losers in the policy process often reveals whether there is broad support in societies for particular policies with particular distributional ends.
1.4. ECONOMIC POLICY MAKING

As we shall see in Chapter 2, there is an important tradition in political economy, which has not had much lay in mainstream economics of appraising the case for good government attaching weight to what would happen in the absence of government intervention. This is central to the normative framework put forward by James Buchanan which we discuss further below and which has its historical roots in Wicksell (1896). They advocate a criterion for good government which emphasizes the principle of unanimity rule relative to what would happen in the absence of government. This approach has a libertarian flavour since the government cannot do anything that reduces an individual below his/her status quo (no government) utility. Thus, each individual is given an effective veto over some aspects of government intervention. Below, we will return to a discussion of its implications.

1.4.2 Good Policies

In the traditional welfare economic model which has dominated modern public economics, good government is largely identified with good policy. The latter is typically defined with reference to both efficiency and distribution. Hence, it sets aside worries about the coherence of ideas like social welfare.

Consider a community of individuals (citizens) who have to make a policy choice. There is a set of feasible policies from among which they must choose. Feasibility requires taking into account both technological feasibility and budget balance etc., as well as information constraints on what government knows about its citizens.\footnote{The seminal contributions on incentive compatibility in normative public economics are Mirrlees (1971) and Hammond (1979).}

In the first instance, the appropriate benchmark for good policy is second best Pareto efficiency, i.e., taking into account appropriate restrictions on policy instruments. A whole tradition of policy analysis in this vein has been developed (see, for example, Atkinson and Stiglitz (1980)). The modern theory of planning has also taken such incentive constraints seriously and sees the difficulties in getting citizens to reveal private information as at the heart of designing good policy. The individuals in that community have heterogeneous policy preferences. The aim is to pick a policy from the feasible set.

The study of efficient policies is incredibly powerful. There are many examples of well-known policy rules in economics that are derived from considerations of Pareto efficiency. For example, the Lindahl-Samuelson rule for the efficient provision of a public good says that the sum of marginal rates of substitution should be set equal to the marginal rate of transformation. However, this does not define a unique policy – it is consistent with many different levels of the public good being provided, depending the distribution of resources in society. The Ramsey tax rule is also an example of a second-best Pareto efficient policy. A further important class of second-best policy rules
are variants on marginal cost pricing. In all cases, the different Pareto efficient policy levels vary according to distributional criterion being used, the information available to policy makers and the form of the budget requirement for the industry in question.

Invoking a specific social welfare function permits the study of optimal policy. The exact policy depends on the form of social welfare function that is selected and, to the extent that social welfare is egalitarian, second-best policy models can formalize the idea of an equity efficiency trade off. Society is willing to have a lower level of total income or average utility in exchange for policies that promote the well-being of socially favored groups.

The welfare economic model is often referred to as the “Pigouvian” model of government, an important antecedent being Pigou (1920) who was the first to systematize the idea that intervention in the economy is guided by a benevolent government pursuing social objectives. Pigou also played an important role in recognizing market failure as a motivation for designing “corrective” policies – taxation to reduce pollution levels being an important example.

The welfare economic model can be thought of as generating “rules for good government”, using a systematic model of the economy and what drives human well-being. The model constitutes one of the crowning achievements of twentieth century economics. The welfare economic approach displaced the rather ad hoc policy reasoning displayed in classical economic.

For example, when Adam Smith discussed the role of government in *The Wealth of Nations*, he did so by cataloguing what government should do rather than putting forward a framework. The following passage from the wealth of nations is quite typical:

“According to the system of natural liberty, the sovereign has only three duties to attend to ... first, the duty of protecting the society from the violence and invasion of other independent societies; secondly, the duty of protecting, so far as possible, every member of the society from the injustice or oppression of every other member of it, or the duty of establishing an exact administration of justice, and thirdly, the duty of erecting and maintaining certain public works and certain public institutions, which it can never be for the interest of any individual, or small number of individuals, to erect and maintain...” Smith (1776, 1976) Book IV, Chapter IX.

Even in Book V of *The Wealth of Nations* where Smith develops various principles of good policy, he does not develop an overarching set of principles, but a catalogue of functions.

The great attraction of the welfare economic model is precisely that it does give a unified way of thinking and a set of powerful lessons. Arguably it is this that elevated the position of economists over other branches of the social sciences in the policy sphere. That said, it has tended to say little about the process of policy choice and implementation. To that extent, it is a highly technocratic perspective.
1.4.3 The Public Choice Critique of Welfare Economics

The welfare economic approach to policy has been criticized by those working in the public choice tradition, for failing to consider how actual policy choices are made and implemented. Even if we were to understand what optimal policies are, there need be no guarantee that the kinds of decision making institutions that we observe in reality will bring them about. The public choice critique of welfare economics says that, by failing to model government, it provides a misleading view of the appropriate role for government. (See Buchanan (1970) for a forceful plea for a level playing field.) The landmark contribution in this regard is Buchanan and Tullock (1962) which showed how choices in a democratic system need not conform to any kind of normative ideal.

The main point of the critique is to remind us that the welfare economic case for intervention offers no guarantee that actual governments making real decisions about policies will actually pick the optimal policies according to a welfare economic criterion. While Pigou (1920) is sometimes thought of as one of the main instigators of the welfare economic tradition, it is clear that he was well aware of this point when he argued that:

“It is not sufficient to contrast the imperfect adjustments of unfettered enterprise with the best adjustment that economists in their studies can imagine. For we cannot expect that any State authority will attain, or will even wholeheartedly seek, that ideal. Such authorities are liable alike to ignorance, to sectional pressure and to personal corruption by private interest.”

Pigou (1920) page 296.

While such concerns may sound rather abstract, they have force in real world contexts. A good example concerns designing interventions when corruption is a possibility. In addition to the standard economic benefits from intervention, it is necessary to weigh up the consequences for corruption.

To address these issues, requires models of public resource allocation that enable us to understand when and if government will behave in accordance with welfare economic model. This requires a significant increase in the competence of economists, requiring them to become expert not only in market resource allocation but also in the study of politics and bureaucracy. Moreover, unless economists engage in this analysis, there is a risk that they will become sidelined in many debates about economic policy. Although there is a long-standing interest in these issues, it is only in the past fifteen years or so that mainstream economics has put incentives in government at the heart of the study of policy problems.

1.5 Political Economy

The term “political economy” has been used in many contexts to refer to different intellectual projects. Hence, it is useful to set the newer usage of this term in its
wider historical context. The classical economists used the term political economy synonymously with economics. Sometime in the late nineteenth century, scholars of the economy came to use the term economics apart from political economy and, ultimately, use of the term political economy lapsed in mainstream economics.

Of particular note in this era is the work of John Maynard Keynes’ father – John Neville Keynes – who published his *The Scope and Method of Political Economy* in 1891. On page 34, he identifies three branches of economics: positive science (what is), normative or regulative science (what ought to be) and the art of political economy – which he refers to as “formulation of precepts”. As for John Stuart Mill, it is apparent that John Neville Keynes views the art of political economy as the branch of economics by which practical maxims are formulated. He remarks:

“when we pass ... to problems of taxation, or problems that concern the relations of the State with trade and industry, or to the general discussion of communitistic and socialistic schemes – it is far from being the case that economic considerations hold the field exclusively. Account must be taken of the ethical, social, and political considerations, that lie outside the sphere of political economy regarded as a science." Keynes (1891), page 55.

There is little evidence, however, that studying the art of political economy as described here was of great interest to mainstream economists in the first half of the twentieth century. Nonetheless, the modern political economy literature is re-engaging with the art of political economy as envisaged by the classical economists.

Throughout the twentieth century, the term political economy remained in use in discussions of comparative economic systems – particularly in debates about the relative merits of socialism and capitalism. This brand of political economy was in part the preserve of Marxist thinkers. But it was also evident in Austrian thinkers such as Hayek and Schumpeter.

Political economy considerations surfaced particularly in the market socialism debates of the 1930s where once again they intersected with mainstream economics. Lange and Lerner had proposed a centralized system which could replicate the market system by using a social planner. Hayek’s position turned on two key problems of planned systems: (i) that government is not omniscient and that the informational requirements assumed are too demanding (ii) that it hinged on a model of government that was too optimistic in assuming benevolence. Hayek (1948) argued persuasively that recognizing the role of markets provided the answer to the first of these. The resolution of the second problem was less clear. Clearly some guarantees of liberty were necessary given the limitations on personal freedom under socialism. However, how democratic systems could be organized to improve the workings of the economy and government was left unclear. Resolving this issue would require a model of democracy which Hayek plainly
1.5. POLITICAL ECONOMY

did not have.\textsuperscript{18}

With the fall of socialism, these debates make interesting history of economic thought but offer little of concrete relevance to contemporary economics and politics. But it is clear that modern political economy does have its roots in a prior set of debates in which political and economic issues were jointly influential.\textsuperscript{19}

The immediate post-war period saw debates about what drives government quality being muted in mainstream economics. The move towards systemization and formalization of economics saw mainstream policy economics largely dominated by a technocratic mode descendant from Pigou’s economics of welfare. By systematizing the notion of market failure, Pigou seemed to promise an economic theory of the mixed economy and the role of government. Much economic theory saw the problem of planning as purely technical.\textsuperscript{20}

In continental Europe, the schism between economics and politics was less marked than in the English speaking world. This was particularly clear in the field of public finance which remained imbued with law and political science throughout.\textsuperscript{21} But it was not until the post war period with the creation of the field of Public Choice that these ideas were systematized into a body of understanding and integrated with mainstream economics in the English speaking world. The key contributors in this enterprise were Buchanan and Tullock whose 1963 book \textit{The Calculus of Consent} provides a landmark analysis of problems of log-rolling and implications of democratic governance for taxation and public expenditures.

In some circles the term Public Choice is used to refer to any analysis that links economics and politics.\textsuperscript{22} But here, I am using it more narrowly to represent the work beginning in the Virginia School in the 1950s. This has three distinctive features.

The first is the assumption of rational self-interest in the study of political interactions. Thus, Buchanan says:

\begin{quote}
“Individuals must be modeled as seeking to further their own narrowly defined, in terms of measured net wealth position, as predicted or expected.” (Buchanan (1989, page 20)).
\end{quote}

To most economists, this may seem innocuous. After all, economic agents as rational egoists is a firmly established tradition in a market context. However, below we will argue that setting aside issues of selection in politics constitutes a blind-spot.

\textsuperscript{18}See Boettke (2003) and Boettke and Lopez (2002) for discussion.
\textsuperscript{19}Referring back to the quote above from John Neville Keynes, it is clear that he saw debates about the merits of socialism as falling under the “art of political economy”.
\textsuperscript{20}See, for example, Heal (1973).
\textsuperscript{21}The excellent collection edited by Musgrave and Peacock (1958) brought these contributions to the attention of the English speaking world.
\textsuperscript{22}For example Mueller (2003).
The second key idea in public choice analysis is the importance of constitutions as constraints on self-interest. Here, Buchanan says:

“To improve politics, it is necessary to improve or reform rules, the framework within which the game of politics is played. There is no suggestion that improvement lies in the selection of morally superior agents who will use their powers in some “public interest”” (Buchanan (1989, page 18)).

In this sense, he is firmly interested in the project of designing effective political institutions.

The third key aspect of Public Choice is its normative framework. Economists have tended to work with a particular (broadly Utilitarian) framework in which good and bad outcomes are seen in terms of their impact on individual’s utilities taken as an indicator of well-being. Various proposals have been made for how to trade these off to get measures of “social welfare” which allows the analyst to engage in policy debates about good and bad policies. But the Public Choice approach is rooted in a quite different normative tradition – one that goes back to classical eighteenth century views of the state (particularly John Locke). The main idea is that the legitimate domain of the state is related to what freely contracting individuals would be willing to agree to, but only that.

While Buchanan has been a champion of these ideas, it was the Swedish economist Knut Wicksell who first applied these ideas in a concrete policy setting – the provision of public expenditures. (see Wicksell (1896)). He studies the problem of public provision via unanimity rule and observed that with benefit taxation, the allocation would obey the contractarian ideal. This approach conflicts with a standard welfare economic framework which appeals to some other authority (the guardian of social preferences) as the arbiter of the justness of the allocation.

This intellectual framework gives rise to the Public Choice critique of welfare economics observing a conflict between its recommendations and the kind of idealized world that Buchanan and Wicksell envisaged. On the whole, the case for intervention is less permissive than the welfare economic view. Moreover, the framework of the analysis has a libertarian flavour.

As we will discuss in the next chapter, the Public Choice approach also offers a particular slant on the concept of political failure – the allocation of resources in democratic process which does not meet Wicksell’s test. Moreover, it was a key insight of Buchanan and Tullock (1963) that there is no guarantee that a system of representative government based on majority rule would be immune to such failures.

The Public Choice approach has been influential in thinking about many of the broad issues concerning the proper role of government. However, it is less associated with the creation of specific models for the study of politics.23

23The Leviathan model of Brennan and Buchanan (1980) is perhaps the main exception. This
**1.5. POLITICAL ECONOMY**

Chicago political economy views politics as a process of competition for support with policies tending towards those that maximize political support offered either through votes or direct monetary transfers. Its pioneers are Gary Becker, Sam Peltzman and George Stigler. These tend to downplay the significance of particular institutional differences in the policy process. The Chicago approach tends to be associated with relatively reduced form models which does make the models fairly easy to use in complex policy environments. While policies can be distorted by political resource allocation, one of the main intellectual thrusts of the Chicago approach is a tendency for policies to be efficient. This comes from the observation that support maximizing politicians have a tendency to prefer efficient policies. As we shall see, this turns out to be important in thinking about different notions of government failure.

The most influential economic model for the study of political resource allocation is that due to Downs (1957). While Downs’ book was filled with many important ideas, the main one that caught on among economists was a justification for the idea that politics would converge to the preferences of the median voter. Downs described politics in the language of competing firms called “parties” where customers were voters. He observed that if parties cared only about winning, then they would have an incentive to converge to the centre. Similar ideas were also being developed in Black (1958) who recognized the importance of preference restrictions (single-peakedness) to this prediction.

While it came to dominate economists’ approaches to political economy, there are deep-seated problems with the approach. First, the reason that parties pick the median outcome in simple models is that this outcome is a Condorcet winner, i.e. an outcome that beats all others in pairwise comparisons. In the absence of a Condorcet winner, there can be cycles. This would happen if there are three alternatives \{A, B, C\} where A can beat B in a simple majority vote, B can beat C and C can beat A. Such Condorcet cycles present an insurmountable problem for the Downsian approach since one party can always win an election by proposing something different from the other party whatever that party is proposing. This matters since in just about any interesting policy problem – particularly those with multiple policy dimensions – no Condorcet winner exists.

Countless papers have been written elaborating this point and trying to propose ways around it. But the bottom line is clear. There is relatively little to commend median voter predictions from a theoretical point of view, except in very special circumstances.

\footnote{See, for example, Becker (1983), Peltzman (1976) and Stigler (1971).}

\footnote{Mulligan, Gil and Sala-i-Martin (2004) even argue that a failure to find a significant difference in policies between democracies and non-democracies is consistent with the Chicago approach. This is because the latter puts so little weight on the important of voting per se in determining policy outcomes.}

\footnote{This term is named after the french aristocrat the Marquis de Condorcet.}

\footnote{This is closely linked to the fact that the idea of a median outcome does not make much sense for a multi-dimensional distribution.}

\footnote{See Mueller (1995) for a discussion of many of these.}
Even so, the model has gained enormous influence among economists and is presented as a cornerstone of political economy analysis to generations of students.

There is another important theoretical problem with Downs’ approach. The model assumes that citizens care about policies while politicians are infinitely pliable—adopting any position to get elected. But if politicians have even a little preference for policies then they will have an incentive to renge after the election. Thus, the model needs to build in reasons why the policy pledges of politicians are credible. One way to approach this is by supposing that politics is a repeated game in which individuals build reputations as in Alesina (1988). However, he shows that this will not typically result in complete convergence. Another is to see credibility as coming from picking candidates with appropriate policy preferences as in Osborne and Slivinsky (1996) and Besley and Coate (1997).

The Downsian approach held much more appeal for economists than political scientists. The latter had long been aware of the evidence from polling data suggesting systematic divergence between median preferences and policy outcomes on key dimensions. The model could offer little insight into where convergence might happen and where it would be absent.

The final problem with an agenda building on the Downsian model is the fact that it is not particularly useful in looking at institutional differences. Indeed if politics is about seeking out median preferences among the electorate, there would be little scope for institutional structure in shaping preference aggregation. There is plenty of good evidence that structures matter in practice and hence that something is at work beyond voter preferences in determining policy outcomes.

The more recent literature has not solved the problem of studying political competition in the absence of a Condorcet winner posed by the Downsian model. But it has made sure to keep this firmly in the background. There are some new modeling approaches, but the approach is not built around any kind of dominant political paradigm. A few key approaches are, however, gaining popularity.

Part of the difficulty in the Downsian paradigm is the fact that there is little institutional restriction on policy proposals. It is very difficult to get a stable point when any policy can be proposed by any political actor at any time. By adding more institutional structure to a model, the degree of freedom open to political actors is diminished and it may be easier to understand policy formation. This idea was a key insight of Shepsle and Weingast (1981) who discuss how restrictions of the structure of proposal power within a legislature can be used to generate a stable point in a multi-dimensional policy space. Roemer (1999) restricts proposal power by modeling within-party conflict. Such restrictions improve the odds of developing a model that predicts an equilibrium outcome in a particular policy context, providing a basis for empirical analysis.

---

29 See, for example, Weissberg (1976).
30 See, for example, Persson and Tabellini (2003).
1.5. POLITICAL ECONOMY

Proposal power is also at the heart of the “agenda setter” model of Romer and Rosenthal (1978).

Another way to create the possibility of an equilibrium is to relax the requirement that the equilibrium be in pure strategies. A number of authors have investigated this (see, for example, Banks and Duggan (2000)).31 Predicting the outcome now becomes an involved process and lacks the simplicity of the median voter outcome. Hence, the Downsian model loses many of its attractions after going down this route.

Probabilistic voting features in many recent contributions. This recognizes that there are random shocks to voter intentions which make the mapping from policy choices into political outcomes uncertain. This simple device is powerful in making concrete progress in studying political strategy.32 The influential monograph by Persson and Tabellini (2000) makes extensive use of this device in exploring the policy implications of different models. This approach often assumes that there are some fixed and some pliable policy dimensions with competition taking place on the latter.33

Old style political economy paid little attention to the selection of politicians. For example, the Downsian model sees policies, not politicians, as the currency of political competition. But in a representative democracy, it is politicians who are elected and are charged with making policy. This idea has been formalized recently by Osborne and Slivinski (1996) and Besley and Coate (1997). These models suppose that citizens elect politicians who then implement their preferred policy outcomes. An implication of the candidate centred view of political competition discussed above is that the identity of candidates matter to policy outcomes.34 Such models can work in complicated policy environments although equilibria may also be in mixed strategies.

Models of extra-electoral policy making are important as well in the political economy literature. Recent contributions have been heavily influenced by Grossman and Helpman (1994) who formulated the problem of lobbying using an approach in which policy favours are auctioned to the highest bidder. Policy outcomes then reflect the “willingness to pay” of organized lobbies. This approach has provided a much more transparent way of thinking about lobbying compared to the previous generation of models which typically had a black box “influence function”.35

The recent literature has a core concern with empirical testing of ideas. There is a wealth of data to be exploited as well as scope to generate new data sets. There

31 See also Myerson (1993).
32 It also helps to overcome some of the technical difficulties associated with finding an equilibrium point in a Downsian model which has an inherent ‘discontinuity’ in the payoff function around the point at which a party switches from winning to losing or vice versa. A probabilistic voting model tends to make the probability of winning a smooth function of policy choices over some range.
33 Lindbeck and Weibull (1987) is an important precursor. They assume that parties care solely about winning. Calvert (1985) and Wittman (1977) consider parties with policy preferences.
34 Lee, Moretti and Butler (2004) has looked at close elections (i.e. those determined by a few points) and argue that the data support the candidate centred view of politics for U.S. elections.
35 See, for example, Becker (1983).
are many studies that look at cross-country variation – exploiting the many differences in institutions that we see between national governments.\textsuperscript{36} The great advantage of this is that the extent of institutional variation is vast creating many possibilities for comparisons of institutions. However, on the downside, such institutions tend to be relatively fixed over time and there are many sources of heterogeneity across countries which it is difficult to control for in a convincing manner. The difficulty then lies in discerning the difference between the effect of institutions on outcomes versus some other unmeasurable factor that is correlated with institutions. This can only be overcome with extreme ingenuity.

Another class of studies exploits variation within countries – where there are differences in politics across sub-jurisdictions.\textsuperscript{37} This is not immune to the problems of unobserved heterogeneity discussed in relation to cross-country studies. The fact that many institutions remain fixed over time is also an issue. However, there are sometimes cases where a change in institutions can be exploited or some suitable interaction with a time-varying factor exploited. More generally, sub national data probably suffer less than cross-country data in having highly heterogeneous cross-sectional units. But such studies typically have less variation in interesting outcomes and institutions to exploit.

Finally, there is scope for increasing collection of bespoke data sets to examine specific policy issues. Economists have long undertaken household survey work to investigate economic behavior. There is similarly a tradition of collecting data sets to examine political behavior – voting, activism etc. But only rarely have the two been put together to get a more complete picture. There is growing interest in doing so and in developing pictures of how policy choices evolve. Bespoke data sets can also be used to supplement standard data from official sources.

Public resource allocation has both short and long run effects on the economy. One distinctive feature of the recent political economy literature is the attention that it pays to the dynamics of politics and economics. A key aspect of democratic political life is that governments are typically short-lived while the consequences of many policies are not. Kydland and Prescott (1977) observed that even benevolent governments would have an incentive to make promises that were not credible – for example promise low taxes to encourage investment and subsequently renege on the promise. But the problem is much worse with short-lived government.

A variety of issues have been studied in models that emphasize this feature of political life. A key example is the incentive to incur public debt as a strategic measure to constrain future governments.\textsuperscript{38} The political business cycle is another example. Accounts of government incentives to inflate the economy before an election have been

\textsuperscript{36} Persson and Tabellini (2003) is an excellent compendium of what can be achieved using such sources.

\textsuperscript{37} See Besley and Case (2003) for a review of what has been achieved for cross-state variation in the U.S.

\textsuperscript{38} See the discussion and references in Persson and Tabellini (2000).
around for a long while. But only fairly recently has it been understood how to think
about this when voters are not being systematically fooled.\(^{39}\)

It is also now clear that long-run patterns of development are tied up with the process
of political development. Problems of state failure are endemic in low income countries
and their study has been to central to appreciating the forces that shape economic
development.\(^{40}\)

It is clear from this brief overview, that there is already a lot of work on issues of
incentives in politics – most of the approaches to political resource allocation discussed
above are concerned to understand how particular assumptions about the structure of
political institutions and organization shape policy choice. For the most part, the
conflict of interest being resolved in the models discussed here is that between different
groups of citizens who have divergent policy interests – the classic spatial model of
politics is the “ideal type” of this kind. Politics is then predominantly about the
problem of preference aggregation. In the next section, we look models which focus
more on the conflict of interest between citizens and government – political agency
models.

1.6 Incentives and Selection in Politics

Most of the analysis of this book – especially that in chapters three and four – is fo-
cussed on how elections can resolve the conflict of interest between citizens and govern-
ment. When political authority is delegated to politicians, then they hold the public
trust between elections. Constitutions offer only limited provisions for the control over
politicians – punishing them only for the grossest of abuses. The main sanction for
poor performance is electoral – those who perform badly will not be re-elected. Hence,
politics is about achieving *accountability* of politicians to voters.

The importance of elections as an accountability mechanism has long been recog-
nized. For example, Alexis de Toqueville viewed U.S. presidential elections in this way
when he noted that

“The President is chosen for four years, and he may be re-elected, so
that the chances of a future administration may inspire him with hopeful
undertakings for the public good and give him the means to carry them into
execution,” de Tocqueville [1835, 1994], page 121.

In similar vein, Madison recognizes this in the Federalist papers when he notes that:

\(^{39}\)Rogoff (1990) uses a dynamic model with imperfect information to develop a “signalling theory” of
equilibrium business cycles.

\(^{40}\)See Acemoglu and Robinson (2003) for a recent insightful discussion of these issues.
“the House of Representatives is so constituted as to support in the members an habitual recollection of their dependence on the people. Before the sentiments impressed on their minds by the mode of their elevation can be effaced by the exercise of power, they will be compelled to anticipate the moment when their power is to cease, when their exercise of it is to be reviewed, and when they must descend to the level from which they were raised; there forever to remain unless a fairful discharge of their trust shall have established their title to the renewal of it.” Madison, Federalist Papers No. LVII.

Information is important in thinking about electoral accountability. Voters who hold politicians to account are likely to do so more effectively when they can observe what they are doing. This brings to the fore the role of information providers such as the media and civil society (think tanks and policy analysts) in increasing accountability.

As we shall see in chapter 3, it is useful to think about the distinction between formal and real accountability. A politician is formally accountable if there is some institutional structure that allows the possibility of some action to be taken against him or her (such as being voted out of office) in the event that he/she does a poor job. But there is no guarantee that this ensures that such accountability mechanisms are used effectively. Real accountability requires that those who hold politicians to account have sufficient information (for example about the politician’s action) to make the system work.

This approach to studying political life was pioneered by Barro (1973) and Ferejohn (1986). They showed how the threat of not being re-elected could curtail rent extraction by politicians. The focus, therefore, is on elections as an incentive mechanism. The problem of opportunism is then essentially like a problem of moral hazard in the contracting literature.

The second generation models – such as Rogoff (1990) and Coate and Morris (1995) have looked at the implications of agents who differ in their type, thus adding concerns of adverse selection. Elections then serve two key roles – creating incentives and selecting the best candidate.

The idea that there are some individuals who are better suited to public life than others is an interesting one. The great American political scientist V.O. Key certainly regarded political selection to be an important issue when he remarked

“The nature of the workings of government depends ultimately on the men who run it. The men we elect to office and the circumstances we create that affect their work determine the nature of popular government. Let there be emphasis on those we elect to office.” V.O. Key (1956), page 10.

Selection could matter for two main reasons – differences in competence between individuals and differences in motivation. With motivation comes concerns about integrity,
1.6. INCENTIVES AND SELECTION IN POLITICS

honesty, altruism and policy interests. The quote from Madison at the beginning of this chapter suggests that he had both in mind. Having “wisdom” is akin to competence and “virtue” is akin to motivation.

The idea that individuals differ in terms of their competence as politicians is similar to the idea that individuals have match-specific skills in models of the labour market. Whether this is innate or acquired is far from clear. It is also unclear whether previous career experience is a reliable guide to whether an individual has greater competence. Madison’s reference to wisdom suggested that he had in mind the need to have politicians with generally high levels of human capital and innate ability.

Dealing with differences in motivation provides more of a departure from standard economics. We are used to assuming, for most aspects of market behavior, that individuals are self-interested. Whether this should be extended to their behavior in public life is debatable. As we noted above, Buchanan puts weight on the idea that ruthless self-interest is the correct assumption to make. The idea of modeling political choices from the premise that actors are self-interested is also characteristic of the Chicago approach to political economy. Thus, Peltzman (1980) assumes that “political preferences are motivated purely by self-interest.” (page 16). The self-interest model is also thought to underpin the growth of government that we observed above. For example, Holsey and Borcherding (1998) comment that the “political paradigm, .., views public services as selfishly redistributive (i.e., transfers of purely private goods to politically favored groups), hence government expenditures are determined by the most influential agents.” (pages 565-566).

The idea that self-interest assumption should be applied to the conduct of public life is far from new. The philosopher and economist David Hume, a contemporary of Adam Smith, argues eloquently in favor of this when he says:

“In contriving any system of government and fixing several checks and controls of the constitution, every many ought to be supposed a knave and to have no other end, in all his actions, than private interest. By this interest, we must government him, and by means of it, notwithstanding his insatiable avarice and ambition, cooperate to the public good” Hume (1742).

Smith himself had a somewhat schizophrenic view. He viewed markets as the perfect outlet for self-interest channeled in the common good. But his theory of moral sentiments attached great weight to altruism in non-market contexts. Hayek proposed the following extension.

“Smith’s chief concern was not so much with what man might occasionally achieve when he was at his best but that he should have as little opportunity

\[41\text{Although explaining the growth of government requires that the self-interest of politicians has increased over time due, say, to a decline in civic virtue.}\]
as possible to do harm when he was at his worst. It would scarcely be too much to claim that the main merit of the individualism which he and his contemporaries advocated is that it is a system under which bad men can do least harm. It is a social system which does not depend for its functioning on our finding good men for running it, or on all men becoming better than they are now, but which makes good use of men in all their given variety and complexity, sometimes good and sometimes bad” Hayek (1948), 11-12.

This has echoes of Madison – the key question is how institutions shape actions and how they can – if at all – promote selection of those most suited to public office. Of course, accepting that individuals are broadly self-interested is consistent with a wide variety of utility functions for politicians. For example, liking policies that help the poor can be regarded as a form of self-interest in so far as politicians gain satisfaction from doing this. Hence the real issue is how narrowly this assumption is interpreted. It is clear that very often self-interest is defined in terms of something like consumption of private goods.

Applied to politics, it is this kind of narrow self-interest motive which underpins accounts of political corruption. It is this notion of self-interest which underpins the agency models of Barro (1973) and Ferejohn (1986). There is little doubt that this kind of self-interested behavior is important in practice and its importance has been recognized increasingly in recent discussions of the role of the state. The most egregious examples – the overwhelming kleptocracies of the world – have been in non-democratic systems where electoral sanctions are limited or non-existent.

There are three other models of motivation in politics: ego-rent, policy preference, and fiduciary duty. We discuss each in turn.

Political economy models often assume that politicians seek solely to win office. Perhaps the best known political model founded on this is Downs (1957) which assumed that parties do whatever it takes to win office. The desire to win could be interpreted as another manifestation of self-interest. However, conceptually, the desire to hold office could be distinct from narrow self-interest. Moreover, the psychology literature attaches weight to the way in which winning an election can bolster self-esteem which would create a reward to holding office which was neither dependent on policy nor on extracting private goods. This model has been invoked by Rogoff (1990) and Maskin and Tirole (2000) among others to describe the preferences of politicians who are otherwise benevolent. This is often referred to as an “ego-rent” derived as a non-pecuniary benefit from being in office. Pure ego rents are distinct from a monetary rent since reducing the possibility of malfeasance while in office does not diminish the re-election incentives of agents who earn them.

As we mentioned above, the motive to enter for office could be to influence policy.42

1.6. INCENTIVES AND SELECTION IN POLITICS

As citizens, politicians are affected by policies and hence have self-interested reasons to be policy motivated. This view encompasses at one extreme a very narrow view in which individuals care about specific issues such as the environment or tax policy. However, policy preference can also be mediated through vaguer ideas like ideologies – world views that shape a broader set of policy predilections.

One important model of motivation in politics that has received less attention in modeling the behavior of politicians, but which in line with the models developed in the chapters below, is the fiduciary model. The term fiduciary comes from the Latin verb _fidere_ – to trust. It is most often applied in describing the duties of trustees and directors of company boards.\(^{43}\)

Fiduciarity has two main aspects – a duty of care and a duty of loyalty. The idea of political office as trusteeship runs through Madison’s essays in the Federalist papers who refers to leaders as upholding public trust. In modern parlance, the duty of care could be interpreted as refraining from moral hazard. The duty of loyalty is less clear in some models, but broadly refers to the duty of politicians to act in the interest of citizens at large – eschewing narrow self-interest. The models that we develop in chapters 3 and 4 have politicians who if elected would fulfill their fiduciary duty on behalf of voters. These are contrasted with narrowly self-interested politicians.

Just why some individuals take their fiduciary duties seriously while others do not is not clear. Performing a fiduciary duty could be thought of as a form of _intrinsic motivation_. This idea, which originated among psychologists has only recently been brought into economic thinking.\(^{44}\) It refers to a case where actions are driven by internal factors and are pursued “for their own sake” rather than because of some well-defined external reward. Thus, a politician who upholds his duty of loyalty may be willing to do so even though he is forgoing an increase in his utility by doing so. On this view, being a trustworthy politician is a type rather than a consequence of incentives. The role of elections is to find ways of sorting in such politicians rather than incentivizing them.

To assume that politicians who take their fiduciary duties seriously are completely immune to influence by incentives is perhaps too strong. The extent of incentives may depend on whether they agree wholly with the ends that they are pursuing. Besley and Ghatak (2005) develop a model of agent motivation in which the extent to which agents take care depends on the extent to which they agree with the _mission_ being pursued by an organization. Thus, a politician could be much more motivated when he/she agrees with the cause. This naturally creates some ambiguity in the exercise of the duty of loyalty in such cases. Those voters whose mission preferences are most closely aligned with politicians will tend to receive more attention from incumbents. This suggests that citizens may care about which sort of fiduciary politician they select.

\(^{43}\)This trustee view of political representation is most closely associated with Edmund Burke who put it forward in a speech to the electors of Bristol in 1774. It also found favor in the writings of J.S. Mill.

\(^{44}\)See, for example, Frey (1997), Murdock (2002), Benabou and Tirole (2003) for examples.
CHAPTER 1. COMPETING VIEWS OF GOVERNMENT

Invoking broader notions of motivation is consistent with wide-ranging evidence on behavior in public and private life – especially that emerging from laboratory experiments. It has long been known that narrow assumptions about human motivation have not fared well in explaining individuals’ contributions to public goods when these are studied in the laboratory. The evidence from experimental work reviewed by Ledyard (1998) catalogs many important anomalies to narrow self-interest models. Even in the context of market interactions, Fehr (2002) documents how concerns about fairness in resource allocation are needed to explain behavior.\footnote{Moving beyond narrow self-interest is necessary to create a satisfactory theory of voting. The possibility that individuals vote on the basis of narrow self-interest runs foul of the fact that the probability of being decisive is so low in mass elections. This has lead to a wide varieties of alternatives being proposed which are based some kind of broader motivation such as duty (see Aldrich (1997) for a review). This notion of duty fits very well with the fiduciary model that we are suggesting.}

Once it is recognize that politicians are heterogeneous in important ways, it is necessary to develop models that explain who will be selected. Let us suppose that there are some individuals who will make better politicians and that all citizens would prefer a better politician in office – politician quality is a valence issue. Then, we can study the process by which political selection takes place and understand why some low quality politicians can survive. The citizen-candidate approach of Osborne and Slivinsky (1996) and Besley and Coate (1997) provides a useful framework for thinking about this. They model the entry, voting and policy making stages when citizens can choose to become candidates for political office.

This approach suggests three main reasons why some low quality politicians get into office. First, there is the possibility of restrictions on entry and voting. If the cost of political campaigns differs across candidates of different quality, then either poor or high quality candidates may be deterred. Political elites may also prefer corruptible candidates in order to preserve their rents. If they control entry (through party systems) then high quality candidates may also be deterred. This mechanism can also work by controlling the voting process to prevent high quality candidates from being elected, even if they stand.\footnote{Poutvarra and Takalo (2003) develop a model in which the value of holding office impinges on candidate quality via its effect on election campaigns. Gehlbach and Sonin (2004) apply a citizen candidate framework to ask when economic elites (such as businessmen) will choose to run for political office. Running for office is in this world an alternative to lobbying for influence. They argue that business candidates lead to greater misuse of public office.}

The second reason for low quality candidates is informational. Bad politicians may enter since there is some chance that the voters will not be able to identify them during election campaigns. This has been studied by Caselli and Morelli (2004). This results in bad candidates entering as rent seekers if there are significant private benefits from running for office.

The third reason for existence of low quality politicians is developed in Besley and Coate (1997) which considers the implications of coordination problems among
voters. Suppose that there are two low quality candidates running for office who are polarized in the policy space. Then entry by high quality candidates can be deterred if these candidates fear that voters would be unable to coordinate on the high quality candidates. Otherwise, the result would be for entry by high quality candidates to increase the probability that a candidate with the opposite policy preference will win.

If the process of entry, campaigning and voting cannot weed out low quality politicians, then observing their performance while in office will play a role. But the initial situation is one of uncertainty about the type of the incumbent (adverse selection). Hence, combining this model with one where incentives matter requires taking the twin informational problems of moral hazard and adverse selection in politics seriously.

To summarize, while it is clear that incentive issues are important in understanding conflicts of interest between government and citizens, there are good reasons to think that selection is important too. This could be selection of more competent politicians or those who will take their fiduciary duty seriously. Much of this book is devoted to understanding how politics and public resource allocation works when both incentives and selection are taken seriously.

1.7 Concluding Comments

If government is to work well, then problems of incentives and selection must be dealt with. In situations where either good politicians are selected and/or incentive problems are dealt with, we have a world approximating the standard normative model of government. At the other extreme, there is a preponderance of self-interested politicians who use public office to further their personal ends. To the extent that the incentive mechanisms available are weak, this will result in low quality government. In this sense, understanding the political economy of public resource allocation in settings where selection is an issue is a useful step along the way to understanding the possibility of benevolent government.

Following in the footsteps of the pioneers whose work we refer above, there is now a considerable body of knowledge to draw upon. The remaining chapters of this book are largely synthetic. They develop frameworks which are useful to explaining the main ideas. However, by developing their own frameworks, there is some hope of bringing unity to the area. But this is a preliminary report on work in progress.

The modern political economy literature is having some success in extending the competence of economists in the direction of understanding policy processes as well as policy outcomes. The project coheres well with Madison’s vision in the Federalist papers. This is not a project about restraining government, but of understanding the institutional preconditions for government to work. With this in mind, we turn to the project at hand.
Table 1.1: Total Government Expenditure as Per Cent of GDP at Current Prices:
Western Europe, the United States and Japan, 1913–1999.

<table>
<thead>
<tr>
<th>Country</th>
<th>1913</th>
<th>1938</th>
<th>1950</th>
<th>1973</th>
<th>1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>8.9</td>
<td>23.2</td>
<td>27.6</td>
<td>38.8</td>
<td>52.4</td>
</tr>
<tr>
<td>Germany</td>
<td>17.7</td>
<td>42.4</td>
<td>30.4</td>
<td>42</td>
<td>47.6</td>
</tr>
<tr>
<td>Netherlands</td>
<td>8.2</td>
<td>21.7</td>
<td>26.8</td>
<td>45.5</td>
<td>43.8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>13.3</td>
<td>28.8</td>
<td>34.2</td>
<td>41.5</td>
<td>39.7</td>
</tr>
<tr>
<td>Arithmetic Average</td>
<td>12</td>
<td>29</td>
<td>29.8</td>
<td>42</td>
<td>45.9</td>
</tr>
<tr>
<td>United States</td>
<td>8</td>
<td>19.8</td>
<td>21.4</td>
<td>31.1</td>
<td>30.1</td>
</tr>
<tr>
<td>Japan</td>
<td>14.2</td>
<td>30.3</td>
<td>19.8</td>
<td>22.9</td>
<td>38.1</td>
</tr>
</tbody>
</table>

Source: Maddison (2001)
Note: The data for the Netherlands is for 1913 rather than 1910.
Figure 1.1: Democracies in the world, 1950-2003
Source: POLITY IV
Notes: A country in a given year is classified as a democracy if variable POLITY2 in the POLITY IV dataset is greater than zero. The percentage of democracies is calculated as the ratio of the number of democracies to the number of all countries in the POLITY IV dataset for each year. Note that the POLITY IV dataset only includes countries with the population of more than 500,000 in 2003. Also note that the number of all countries in the dataset changes over time due to the independence of colonies, the break-up of a country, and the integration of countries.
Figure 1.2: Size of government (in nominal terms) in high-income OECD countries and the others
Source: World Development Indicators 2005
Notes: Size of government (in nominal terms) is measured as the percentage of general government final consumption expenditure in current local currency unit over GDP in current local currency unit. The simple average for each group of countries is calculated. High-income OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (the unified Germany before 1990), Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. The data for year 2003 is excluded as eight out of the 23 high-income OECD countries are missing. For the sake of comparability to Figure 2b, country-years for which the data on the size of government in real term is unavailable are dropped.
Figure 1.3: Size of government (in real terms) in high-income OECD countries and the others
Source: World Development Indicators 2005
Notes: Size of government (in real terms) is measured as the percentage of general government final consumption expenditure in constant local currency unit over GDP in constant local currency unit. The simple average for each group of countries is calculated. High-income OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (the unified Germany before 1990), Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. The data for year 2003 is excluded as eight out of the 23 high-income OECD countries are missing. For the sake of comparability to Figure 2a, country-years for which the data on the size of government in nominal term is unavailable are dropped.
Figure 1.4: Size of Government (in nominal terms) between democracies and autocracies
Source: World Development Indicators 2005 and POLITY IV
Notes: Size of government is measured as the percentage of general government final consumption expenditure in current local currency unit over GDP in current local currency unit. The simple average for each group of countries is calculated. A country in a given year is classified as a democracy if variable POLITY2 in the POLITY IV dataset is greater than zero and as an autocracy if POLITY2 is zero or negative. Note that a country that is a democracy in some year can be an autocracy in another year. For the sake of comparability to Figure 3b, country-years for which the data on the size of government in nominal term is unavailable are dropped.
Figure 1.5: Size of Government (in real terms) between democracies and autocracies

Source: World Development Indicators 2005 and POLITY IV

Notes: Size of government is measured as the percentage of general government final consumption expenditure in constant local currency unit over GDP in constant local currency unit. The simple average for each group of countries is calculated. A country in a given year is classified as a democracy if variable POLITY2 in the POLITY IV dataset is greater than zero and as an autocracy if POLITY2 is zero or negative. Note that a country that is a democracy in some year can be an autocracy in another year. For the sake of comparability to Figure 3a, country-years for which the data on the size of government in nominal term is unavailable are dropped.
Figure 1.6: Corruption in high-income OECD countries and others


Notes: Corruption Index (variable “corruption in government” in the dataset) ranges in values 0-6 with higher values indicating less corruption. The simple average for each group of countries is calculated. High-income OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (West Germany before 1990), Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States.
Figure 1.7: Corruption in democracies and autocracies


Notes: Corruption Index (variable “corruption in government” in the dataset) ranges in values 0-6 with higher values indicating less corruption. The simple average for each group of countries is calculated. A country in a given year is classified as a democracy if POLITY2 variable is greater than zero, and as an autocracy if POLITY2 variable is zero or negative. Note that a country that is a democracy in some year can be an autocracy in another year. The International Country Risk Guide dataset covers years 1982-1997.
Figure 1.8: Property rights for high-income OECD countries and others
Notes: Property Rights Index (variable “risk of expropriation” in the dataset) ranges in values 0-10 with higher values indicating more secure property rights. The simple average for each group of countries is calculated. High-income OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (West Germany before 1990), Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. Although the dataset covers years 1982-1997, years 1982 and 1983 are excluded because 21 of the 23 high-income OECD countries are missing in the dataset for these years.
Figure 1.9: Property rights for democracies and autocracies
Notes: Property Rights Index (variable “risk of expropriation” in the dataset) ranges in values 0-10 with higher values indicating more secure property rights. The simple average for each group of countries is calculated. A country in a given year is classified as a democracy if POLITY2 variable is greater than zero, and as an autocracy if POLITY2 variable is zero or negative. Note that a country that is a democracy in some year can be an autocracy in another year. The International Country Risk Guide dataset covers years 1982-1997 though figures for years 1982 and 1983 are dropped because about half of democracies are missing for these two years.
Figure 1.10: National Election Turnout in high-income OECD countries, 1946-2000
Source: International Institute for Democracy and Electoral Assistance (IDEA) (www.idea.int/vt)
Notes: “Vote/Registration” is the ratio of the total number of votes cast to the number of registered voters. “Vote/Voting Age Population” is the ratio of the total number of votes cast to the population over the age of 18. The five-year period is chosen so that each country, except for Greece, Portugal, and Spain during their dictatorial rules (see below), held at least one election in each period, which avoids the composition bias. The five year average turnout for high-income OECD countries is calculated as follows. For each country and each five year period, the turnout (variables Vote/Registration or Vote/Voting Age Population in the IDEA dataset) for legislative elections is averaged. For countries whose form of government is presidential (e.g. the United States), the average turnout for presidential elections is calculated separately, and averaged with the legislative counterpart, in order to ensure that the same weight is given for each type of election. The simple average over countries is then calculated for each five year period. This makes sure that those countries that held more elections within each period than others are not over-represented. High-income OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (West Germany before 1990), Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. There were no national elections in Spain for 1946-1976, Greece for 1965-1973, and Portugal for 1946-1974. The data on the number of registered voters (hence Vote/Registration) is not available for Greece for 1946-50 and the United States for 1946-63.
Figure 1.11: Trust in government and turnout (simple average across US states)

Sources: Trust in government is taken from NES (National Election Surveys); Turnout obtained from Michael McDonald for 1952-1998 and his website (elections.gmu.edu) for 2000-2002.

Notes: Turnout is the ratio of the total vote cast for the highest office to the voting age population. Shown in the figure is the simple average over each presidential year and its following mid-term election year with the same weight on each state. Trust in government (variable CF0656 in the NES dataset) is the aggregate index of survey responses to the following four questions: “How much of the time do you think you can trust the government in Washington to do what is right?”, “Would you say the government is pretty much run by a few big interests looking out for themselves or that it is run for the benefit of all the people?”, “Do you think that people in the government waste a lot of money we pay in taxes, waste some of it, or don’t waste very much of it?”, and “Do you think that quite a few of the people running the government are crooked, not very many are, or do you think hardly any of them are crooked?”. The survey was conducted each other year when national elections were held. The simple average across US states shown in the figure is calculated as follows: the simple average is taken across respondents in each state; then the simple average across states is calculated for each survey year; finally, each presidential election year figure is averaged with the figure for the following mid-term election year (except for 1958 where the survey in 1956 did not take place). For both turnout and trust in government, the states of Alaska and Hawaii and the District of Columbia are not included.
Chapter 2

The Anatomy of Government Failure

“(E)fficient government ... is a standard by which performance can be measured, similar to corresponding standards by which performance of households and firms in the private sector are assessed. Actual performance will differ among governments and periods of time, but efficient conduct and constructive leadership are not beyond reach.” Musgrave (1999, page 34).

2.1 Introduction

Government failure is a term that is often used, but rarely defined. The basic and highly intuitive idea is that there are systematic reasons why government fails to deliver the kind of service to its citizens that would be ideal. It is invoked, in particular, as a reason to be doubtful about the usefulness of the standard welfare-economic recommendations for government intervention. An analyst will frequently say that government failures need to be weighed against market failures in making the case for government intervention according to welfare economic prescriptions.

However, unlike its sister notion of market failure, no systematic account can be found of this idea in the political economy literature. This chapter tries to remedy this by discussing alternative notions of government failure. It will do so through some general discussion as well as developing a simple example to illustrate the main ideas.

It is important to distinguish government failure from a narrower concept of political failure. Government failure refers to problems that arise when one actor in the economy (the state) monopolizes the legitimate use of force. Political failure refers to the narrower idea of problems that arise when power to control this monopoly is allocated in democratic political systems. Political failures as described here are therefore a subset of government failures.
Two examples will help to make this distinction clear. The first example is the problem of imperfect information and the provision of public goods. It is well-known since the seminal work of Clarke (1971) and Groves (1973) that the inability of governments to measure accurately the valuations that individuals place on public goods may lead to a suboptimal level of government provision. This implies that the Lindahl-Samuelson rule for public goods provision cannot be achieved. The problem is generic to the operation of government and is likely to arise (to some degree) under any system of government. Hence, imperfect information of this form may constitute a government failure. But it has nothing to do with politics.

For the second example, consider the problem of finding rules for the operation of a legislature charged with making decisions about public resource allocation. Suppose that the legislature concentrates the power to set the policy agenda to a single individual, with the others able to vote on his or her proposals. Suppose that this agenda setter represents a particular district, but through self interest, fails to internalize the impact that spending in his/her own district has on other districts. This will lead to a sub-optimal pattern of resource allocation. It may be possible to construct a Pareto improvement in public resource allocation (assuming that suitable compensatory transfers can be made). Imperfections in the operation of legislative bodies can result in political failures.

It is often not important whether a particular problem in the functioning of government is a government failure or is specifically associated with democratic resource allocation. However, there are two reasons to be interested especially in political failure. First, studying political failure may give a sense of the potential drawbacks of democracy which we argued in the last chapter is coming to be a dominant institution. As with studies of markets which motivated ideas of market failure, it is useful to know how democracy really works. Second, studying political failures may give concrete insights into how democratic systems of government may be improved, in particular how changing the rules of the game can lead to improvements in resource allocation.

The standard way in which economists discuss market resource allocation provides a model for thinking about government failure. As we discussed in the previous chapter, the primary criteria of assessment are equity and efficiency. While the term market failure is typically reserved to describe situations where market resource allocation results in Pareto inefficiency, government failure, as we shall discuss below, has been used rather differently. In fact, there are three main notions that the literature has suggested. Only one of these corresponds straightforwardly to the standard definition of market failure.

As with market failure, it would be useful to understand just how prevalent government failure is. There are those, for example Wittman (1997) who have argued that democratic systems will tend to produce efficient results. However, the Public Choice tradition as typified by Buchanan and Tullock (1962) sees the world quite differently. In part, these conflicting views stem from using the term government failure rather differently. Either way, with such conflicting claims, it is necessary to look closely at these
2.2 THREE NOTIONS OF GOVERNMENT FAILURE

issues.\(^1\)

The title of this chapter borrows unashamedly from Bator (1958) which pulled together various ideas of market failure for the first time. He defines his quest as “an attempt ... those phenomena which cause even errorless profit- and preference-maximizing calculation ... to fail to sustain Pareto-efficient allocation.” (page 352). He argues that non-appropriability, non-convexity and public goods are the main sources of market failure. In principle, we would like to catalog government failures in terms of similarly simple categories.

The procedure we will adopt is similar. We will assume that actors in government and political processes maximize their payoffs under appropriate constraints. We will then explore when this leads to policy outcomes and patterns of private influence activity that are Pareto inefficient. The overall aim is to identify the basis of inefficiencies. We will also broaden the set of possible normative criteria to include cases where political resource allocation is inequalitarian. Finally, we will consider the implications of certain non-consequentialist criteria.

As we discuss in greater detail below. Ideas about government failure are central to understanding constitution design. However, it also plays a central role in thinking about economic policy reform and reasons why apparently beneficial reforms are not undertaken.\(^2\) By appreciating the specific form that a government failure may take, it may be possible to understand when reforms are likely to work in the way that their architects intend.

2.2 Three notions of Government Failure

In this section, we discuss three ways of defining policy outcomes as government failures. The first parallels the classic definition of market failure – Pareto inefficiency. The second allows the possibility that the political process produces an “undesirable” distributional outcome. The third is due to Wicksell and has been developed in the writings of James Buchanan – it is based on whether a particular intervention Pareto dominates what would happen in the absence of government.

As we shall see in the example studied in the next section, these notions of government failure can be applied to the policy outcome and to the policy process, i.e., to any resources used up in the decision making process. The latter refer, in particular, to standard “rent-seeking” inefficiencies. Thus, we need to study the set of policy outcomes

\(^1\)Acemoglu (2003) in many ways shares the ambition of this chapter. He frames the issue as understanding why the Coase theorem fails in a political arena. Much of the discussion of government failure here – especially the examples of Pareto inefficiencies – could be framed in this way. Acemoglu (2005) pursues the more ambitious task of look an institutions in a policy setting in order to understand institutional failure (in the language of this chapter).

\(^2\)See Rodrik (1996) for an excellent discussion of such issues in general.
and a set of private and public actions that are made to achieve this outcome. A
government failure might then arise either because the policies selected are poor or
because the means of picking (even good) policies is very costly. A system of government
free from government failure will pick good policies and policy processes. It will also
encourage efficient private actions to affect policy outcomes.

2.2.1 Pareto Inefficiency

The most obvious definition of government failure to an economist is based on *Pareto
efficiency*. This most clearly parallels the textbook case of market failure. This is
motivated by a long tradition in public economics which has studied (in particular)
Pareto efficient taxation and public spending.\(^3\) For example, the Ramsey tax rule and
Lindahl-Samuelson rule for provision of public goods are policy rules that follow from
the characterization of Pareto efficient policies.

The output of such an exercise is society’s *Utility Possibility Frontier*. This character-
izes the set of government policies and private resource allocation decisions where
an individual cannot be made better off without another being made worse off. Being
Pareto inefficient means operating inside this frontier. Pareto inefficiency is the econo-
mists’ free lunch – it should be possible to pick a different set of policies and/or private
decisions so as to make every citizen better off.

Market failure is defined as a situation where markets fail to achieve an allocation on
the Pareto frontier. Applied to government failure, this approach says that government
fails when *policies* result in a society being inside its Pareto frontier. Given the obvious
extension of market failure analysis in this way, it is somewhat surprising that this notion
of government failure has received so little attention in the literature. It was suggested
as the appropriate benchmark for government in Besley and Coate (1997) and developed
further in Besley and Coate (1998).

Government failure as Pareto inefficiency is in many ways a weak criterion. As
observed in Besley and Coate (1997), any situation where some citizen is given the right
to pick policy from the feasible set will be trivially Pareto efficient in a static setting
since it is certainly not possible to make the policy maker better off. However, we will
see that there are some interesting non-trivial examples of government failure defined
this way in dynamic models.

2.2.2 Distributional Failures

As we discussed in the previous chapter, Pareto efficiency is frequently deemed too
weak a criterion for normative analysis. After all, a political equilibrium in which a
dictator transfers the whole of society’s resources to himself and a few of his cronies

\(^3\)See, for example, Atkinson and Stiglitz (1980).
2.2. THREE NOTIONS OF GOVERNMENT FAILURE

can be efficient. Yet few would regard it as a satisfactory state of affairs and many think that this could constitute a government failure. The only way to tackle such concerns is to bring in distributional concerns and to judge the policy outcome and process accordingly.

Some authors, for example Mueller (1996), seem to regard transfers to farmers in rich countries, heavily subsidized public projects (such as the development of Concorde) or geographical targeting of public goods to particular regions as prima facie evidence of government failure. If such transfers are made inefficiently, then this collapses back to the first definition of government failure. However, assuming that this is not a source of Pareto inefficiency, it is clear that there is an implicit appeal to particular social welfare function against which the policy produced by government is being judged. These examples therefore lie in the realm of distributional failures.

Thus, to make this operational requires some kind of distributional metric, i.e. a social welfare function. However, unless there is a fair degree of consensus on social preferences, the conclusions from such an analysis will likely be controversial. Since there is little reason to suppose that any kind of political process will maximize a well-defined social welfare function, then the danger is that any policy picked in a democracy will result in a government failure according to this definition. Hence, this notion of government failure risks having no bite at all.

However, this is probably too pessimistic. There might, for example, be rather broad agreement that any decent government should limit the extent to which the government officials use the state for the purposes of self-enrichment, however efficiently they choose to do it. The experience of kleptocratic dictators such as Mobutu in Zaire or Marcos in the Philippines underline the generalized outrage that is felt when government is used as a vehicle for self-enrichment.

Another approach to thinking about government failure which invokes an element of distribution is social surplus. The notion of social surplus only makes sense as an efficiency criterion in the case of “transferable utility”, i.e. utility that is linear in money. It can then be motivated in terms of a compensation test – of the Hicks-Kaldor variety – where the gainers can compensate the losers. Thence if a policy choice maximizes social surplus there is no direction of policy change in which the gainers can compensate the losers. Conversely, moving towards a surplus maximizing policy will guarantee that gainers from the change can compensate losers.

As is well understood, this logic only seems compelling when compensations are actually paid in which case it corresponds to a Pareto improvement. Moreover, payment of compensation can happen only with a very rich set of policy levers, including lump-

---

4See, for example, Mueller (1996, page 23).
5Proponents of probabilistic voting models have sometimes suggested that particular social welfare functions are maximized in political equilibrium. (See Coughlin (1992) for a discussion.) However, they rest on strong assumptions and it appears unlikely that technological assumptions are at the heart of the distributional conflict implicit in political competition.
sum taxes and transfers. If compensations are not actually paid, then the appeal of surplus maximization is less obvious. It is then best thought of as a distributional rather than an efficiency criterion. Indeed when preferences are linear in money, the social surplus optimum and the Utilitarian social welfare optimum frequently coincide.

While surplus maximization often has appeal given the notion of “making the cake as big as possible”, it is better to think of it as a kind of social welfare maximization exercise with a particular social welfare function. In the sequel, we will use it this way. In fact, in our core examples, we will use it as our main example for the investigation of distributional failures. This is because we make no pretence that compensations needed to generate a Pareto improvement are being paid.

In more pragmatic terms, there is a sense in which it serves as a powerful force for criticism in the context of government failures. If policies are captured by political elites or special interests and result in lower social surplus, then there will typically be more sympathy in reducing the political advantages of these groups to benefit the broader group of citizens. When employed in this context, the social surplus criterion can indeed be powerful since it can demonstrate that a small group is profiting at the expense of a much wider group. Thus weighing up the consumer surplus losses from tariff protection against the benefits generated for stakeholders within an industry can have influence on policy debates about protectionism.

As we have already mentioned, distributional considerations are also important in thinking through the implications of politicians earning rents from holding office. There are models that put this at centre stage – including the agency models that we discuss in chapter 3. Another example is the Leviathan model of Brennan and Buchanan (1980) where politicians are assumed to maximize the tax revenue that they can extract from citizens while diverting some tax revenues to private use.

Just how to build a constraint that specifies a government failure in this context is moot even though to many the prima facie case is clear. However, it is worth bearing in mind that the decision to enter politics is likely to be a function of the rewards on offer to politicians and may affect the competence of those who choose to stand for office (with more competent politicians having a higher opportunity cost). Hence, specifying that there should be no reward to holding political office is probably too strong.

One criterion for government failure would investigate whether it is feasible to obtain the same policy outcome with fewer rents. A government failure would then constitute an outcome with “excessive” rents relative to this benchmark. The surplus maximization criterion could also be useful here in demonstrating that enriching the politician is imposing a significant cost on the voters. Many would not find the case for compensating the politician for his loss of rents compelling and hence one of the objections to surplus maximization falls away.
2.2. THREE NOTIONS OF GOVERNMENT FAILURE

2.2.3 Wicksellian Failures

Our third notion of government failure is drawn from the writings of Wicksell (1896). He uses a criterion somewhat outside the standard welfare economic model and is best thought of as a rights based approach which is derived from classical liberalism. At the heart of this is the notion that policy outcomes and political decisions should lead to an outcome that Pareto dominates what would be achieved without government. The idea behind this idea is seen clearly in the following passage:

“If any public expenditure is to be approved, whether it be a newly proposed or an already existing one, it must generally be assumed that this expenditure as such...is intended for an activity useful to the whole of society and so recognized by all classes without exception. If this were not so ..., then I, for one, fail to see how the latter can be considered as satisfying a collective need in the proper sense of the word. ... It would seem blatant injustice if someone should be forced to contribute toward the costs of some activity which does not further his interests or may even be diametrically opposed to them” Wicksell (1896, page 89).

The main motivation for this idea is to think of government (like the market) as a process of exchange which results in Pareto improvements over some status quo point. Alternatively, it could be approached from a contractual point of view, thinking of citizens as signing up to a grand contract that defines what government will do, with every citizen having veto power over the contract.

This notion limits the extent to which the policy process can lead to redistribution of resources between its citizens. For example, it rules out pure redistribution except in so far as the losers feel altruistic towards the beneficiaries and hence benefit as altruistic donors from redistribution to others. For this reason, the approach is often thought of as providing a conservative way of judging the legitimacy of government intervention. This leads to a marked contrast with the main stream welfare economic tradition. One concern with the approach is the fact that an initially unjust allocation of resources would be perpetuated through history and could not legitimately be changed by government.6

2.2.4 Comparisons

We will now look at how these ideas of government failure relate to each other in an abstract sense. The example that we develop below will give this more precise content.

6This was clearly recognized by Wicksell (1896) who notes that “It is clear that justice in taxation tacitly pre-supposes justice in the existing distribution of property and income.” (page 108). He goes on “if there are within the existing property and income structure ... priveleges .. in open contradiction with modern concepts of law and equity, then society has both the right and duty to revise the existing property structure.” (page 109). Just how this is done, is left to something of a fudge and he entertains some kind of qualified majority rule without specifying the exact procedure or rule.
Here, we develop a graphical representation.

As we noted above, the first two criteria of government failure are based on a standard welfare economic approach. They are likely to be nested in the following sense – pretty much any reasonable social welfare function will also regard a Pareto inefficient policy choice as a failure too. Hence, government failures based on Pareto inefficiency tends to be a strict subset of those based on a broader social welfare criterion.

This is illustrated in Figure 1 which illustrates a social welfare function defined on the utilities of two citizens 1 and 2. Point $A$ is the full optimum according to these social preferences. If point $B$ were attained through choice of government policy, then this would be deemed a government failure using these social preferences. However, it is not Pareto inefficient. This makes it clear just how widespread government failures based on distributional preferences are likely to be. Any point away from $A$ is a government failure.

The Wicksellian criterion provides quite a different slant on government failures. Figure 2 illustrates the difference between this criterion and that based on Pareto inefficiency. Suppose that at the status quo (no government), the economy would operate at a point like $A$. This is inside the Pareto frontier representing the possibility that, say, by fixing market failures, the government can make everyone better off. Suppose that point $B$ is the outcome after government intervention. Point $B$ is now on the Pareto frontier and hence is (second best) efficient. However, it does not constitute a Pareto improvement over point $A$. Hence, if chosen by government, it would constitute a Wicksellian government failure. However, it would not be a government failure according to the Pareto efficiency definition as there is no scope for improving government efficiency. Now consider point $C$. According to the Wicksellian definition, it is not a government failure as it a Pareto improvement relative to $A$. However, the definition based on second-best Pareto efficiency would regard it as a government failure. It is possible to make all citizens better off beginning from this point.

Using Wicksell’s definition, a government can intervene efficiently in the welfare economic sense and yet still create a government failure. Indeed, a Wicksellian government failure is possible even if the outcome generated in political equilibrium is social welfare maximizing according to an agreed social welfare function.

### 2.3 An Example: Financing a Public Project

As we have shown, these three notions of government failure are distinct. But whether this matters can only be assessed by thinking about concrete policy problems. This
2.3. AN EXAMPLE: FINANCING A PUBLIC PROJECT

is at its starkest in thinking about policies aimed at redistributing income. Suppose that the government does this in the most efficient way. In the case of full information, this would be using lump-sum taxes while otherwise, it would use an optimal income tax or any other appropriate optimal tax system. This is normally studied with reference to a specific social welfare function and any government that did not use this social welfare function would be deemed to have failed on the basis of distributional failure while the Pareto efficiency criterion for government failure would deem any tax system which is optimal for some set of welfare weights to be free of government failure. The Wicksellian definition would reject any form of redistribution – so if the economy were Pareto efficient without government intervention, this would mean that there is no legitimate government intervention of this form.

Such abstract discussions of the logic of government failure suggest that pure redistribution is not going to be a case where ideas of government failure have much bite unless there are some predefined notions of acceptable income redistribution. Hence, we will focus for the remainder of the discussion on cases where government can intervene when there is some form of inefficiency – specifically private “under-provision” of a public good. This canonical example will allow us to have a richer discussion which does not run into the dead end that we have found for pure redistribution. The competing notions of government failure can then have real bite.

Suppose that a community of \( N \) individuals has to make a single social decision – whether to build a public project. We denote the decision to build the project by \( e \in \{0, 1\} \) where \( e = 1 \) denotes the case in which the project is constructed. If the project is built, the community must decide how to finance it. Here, we will assume that if the government funds the project then it uses a head tax (equal per capita financing) if the project goes ahead.

There are two groups of citizens – those who value the project and receive utility \( b \) from it and those who do not value the project, receiving a utility of zero. The citizens who value the project comprise a fraction \( \gamma \) of the population. All citizens have an income of \( y \) and the project costs \( c \) to implement. Assume that \( y > \frac{c}{N} \), each citizen has an endowment sufficient to pay their per capita cost.

2.3.1 Private Provision

Before proceeding to public provision, observe that we can motivate public provision of the project as fixing a classic market failure. Suppose that the project is to be funded through private subscriptions where each citizen contributes \( s_i \) (\( i = 1, \ldots, N \)) to its cost. If \( \sum_{i=1}^{N} s_i \geq c \), then the project goes ahead and each citizen has any surplus returned to

\footnote{This is a simplification as such tax systems are not seen in practice. However, the main points that are illustrated using the example do not hinge on this. It is key that there is no optimal lump-sum taxation. (See below for more on this.)}
them on an equal sharing basis, i.e. they get \( \frac{1}{N} \left( \sum_{i=1}^{N} s_i - c \right) \). If \( \sum_{i=1}^{N} s_i < c \), then the project does not go through.

Suppose that we look for a Nash equilibrium in contribution levels. Then, our first observation is that there is no equilibrium in which any citizen who does not like the project makes a positive contribution. However, there are a variety of Nash equilibria where those who value it make a contribution. In each of these, all of the contributors must be pivotal so that the value of the contributions just adds up to \( c \).

As long as \( b < c \), then there is always a Nash equilibrium where \( s_i = 0 \) for all citizens. There may also exist a Nash equilibrium where \( s_i = \frac{c}{N} \) if \( N \gamma b > c \). This Pareto dominates the zero provision equilibrium in this instance and hence failure to achieve it leads to a classic “market” failure.

To motivate government intervention, therefore we will focus on the case where there is zero private provision. This discrete case is arguably artificial given that an efficient equilibrium exists. It rests here on a coordination failure due to the free-rider problem rather than the free-rider problem per se. However, inefficiency in private provision of public goods is generic. Hence, it is reasonable to overlook this issue here and the gain from the simplicity of the example more than outweighs this (slightly) artificial feature.

### 2.3.2 Government Provision

Suppose now that government finances the project and uses a head tax. To create a social welfare calculus, we adopt a Utilitarian perspective which is identical to the social surplus criterion in this context. In this case, the Samuelson rule applies to optimal public provision.

Figure 3 gives the payoffs to all of the parties in this instance.

<table>
<thead>
<tr>
<th></th>
<th>Citizens who value the project</th>
<th>Citizens who do not value the project</th>
<th>Social Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e = 0 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( e = 1 )</td>
<td>( b - \frac{c}{N} )</td>
<td>( -\frac{c}{N} )</td>
<td>( \gamma Nb - c )</td>
</tr>
</tbody>
</table>

Figure 3: Citizens’ Payoffs and Social Welfare Under Government Provision

Observe that \( e = 0 \) and \( e = 1 \) are both Pareto efficient policy choices in this setting. This is because we have supposed that the government has to tax all citizens. Thus those who do not value the project are worse off under government provision. This implies that a system of political resource allocation that costlessly decides whether to have the project go ahead or not cannot constitute a government failure using the Pareto criterion.

Using our posited social welfare criterion, the project is worthwhile if the sum of benefits to all citizens exceeds the resource cost which boils down to:

\[ N \gamma b \geq c. \]
If this condition holds, then any political mechanism in which the project fails to go ahead constitutes a government failure. (The converse would be true if we assumed that $N\gamma b < c$.)

Turning now to Wicksell’s unanimity test, observe that this always fails if $e = 0$. This is because the group who does not favor the project have a payoff of $-c/N$ as they do not value the project. Hence any model of political resource allocation in which the project goes ahead would generate a government failure. However, if the government could levy a benefit tax in the amount $c/N\gamma$ on all those who value the project, then going ahead with the project would indeed generate a realized Pareto improvement if $N\gamma b > c$. In fact this achieves the same payoff as with the “good” Nash equilibrium in this case. Except for where we make specific mention, we assume that $N\gamma b > c$ for the remainder of this.

The simple example makes clear why the Wicksellian model is likely to lead to the most conservative criterion for government intervention when the benefits of public intervention are unevenly distributed among the citizens. Unless benefit taxation can be used to fund public projects, then the case for government intervention will be extremely limited. In contrast, Pareto efficiency does not have any bite in this situation whereas social surplus allows trade-offs between the payoffs of gainers and losers.

The example also illustrates why the set of instruments available to government matters for a government failure. If we had assumed that lump-sum taxation were feasible for government, then the Wicksellian criterion and social surplus could always be fulfilled together in this model by suitable use of such transfers. Moreover, there would be a unique Pareto optimal policy which coincided with the surplus maximizing outcome. Thus, interesting conflicts between competing notions of government failure require us to work with a world where there are restrictions on policy instruments.\footnote{The uniqueness of the Pareto optimal outcome with lump-sum transfers is also an artifact of the assumption of transferable utility.}

This example provides a useful building block that we will now use to study public resource allocation and the reasons for government failure. While simple and stylized, it does embody many of the ideas that make the study of government intervention in political economy settings interesting. In particular, there are gainers and losers from government intervention with limited policy instruments for compensating losers.

### 2.4 Sources of Government Failure

This section discusses three aspects of government failure using the example above to illustrate their implications. We begin by discussing government failures that arise whether or not a democratic political process is used to make policy decisions. These problems are due to ignorance, the use of private influence and the quality of leadership.
In each case, we begin by supposing that there is a leader in office who is charged with deciding whether or not to set $e = 1$. He/she is assumed to have their own preference over the project outcome. We do not model the process of leadership selection. We then study how this policy authority will be exercised and its welfare consequences.

### 2.4.1 Ignorance

Governments who intervene in the economy typically lack the kind of omniscience that simplistic models of the policy process invoke. This was the basis of Hayek’s critique of state planning that we discussed in section 5 of Chapter 1. However, it runs through the modern welfare economic literature based on mechanism design. At some level, it is hardly deep or surprising that some forms of ignorance can lead to policy mistakes being made, at least when compared to the omniscient outcome. In our example, a government that was ignorant of $b$ or $c$ would be unlikely to make the correct decision all of the time over whether to go ahead with the project. This would be true regardless of the criterion being used to evaluate the quality of public decision making. However, it is equally unclear, a priori, whether ignorance leads to a systematic bias in policy decisions towards too much or too little government intervention.

Ignorance can clearly lead to a Pareto inefficient policy, at least when compared to the case of full information. It can also lead to social decisions that fail specific welfare criteria and to situations where policies are implemented that do not Pareto dominate the status quo.

While ignorance is undoubtedly a pervasive feature of the policy landscape, the most interesting issues concern situations where the quantity of information is endogenous and when different individuals are differentially informed about policy. In these cases, the question is whether particular decision making processes are better at eliciting and managing information. In general the quality of policy outcomes will fall short of what happens with perfect knowledge – the issue is whether all information is incorporated in making social decisions. However, the relevant criterion is now second-best, i.e. information constrained efficiency.

One key difference between democratic and non-democratic systems of government lies in the way in which information is collected and disseminated in the policy process.\(^9\) This could mean democratic systems of government less prone to government failure if it leads to policies that are more reflective of common underlying information about policy needs. Suppose, for example, that the government is uncertain about the fraction of individuals in the population who value the project (the parameter $\gamma$ in our example). Then, direct voting over policy may be a good way of revealing common values when citizens are differentially informed.

To look at these issues in more detail would require us to develop a specific model of

---

\(^9\)These issues are explored in Feddersen and Pesendorfer (1997).
information aggregation in the political process. This raises interesting and important issues. We will not, however, be addressing them to any extent in this book. However, clearly the form in which social decisions are made has a bearing on the way in which information can be incorporated in the policy process. Laffont (2000) provides many insights in political processes viewed from this perspective.\(^\text{10}\)

The agency models discussed in chapters three and four take informational issues seriously. However, they focus on situations where the government retains an informational advantage about policy. In the current setting this might be because \(b\) or \(c\) is known to the government and not to the citizens. This does not, however, create a problem unless the policy maker also faces a conflict of interest – has an incentive to pick a policy outcome that does not suit voters. Hence, to make the model interesting, we will also need to assume that politicians are not always benevolent. One way of doing this is to suppose that they face external influence in choosing policy. It is to the issue of influence that we now turn.

### 2.4.2 Influence

Whether democratic or not, governments are subjective to influence from powerful organized groups. This will lead, in general, to policy benefits being skewed towards such groups. In this section, I discuss how this can be brought into the example above and discuss their implications for identifying government failures. I consider two models. The first is “pure” corruption whereby influence is purely redistributive. The second is costly “rent-seeking” where resources are used up in the policy process.

**Corruption**

We define pure corruption as a situation where a monetary payment – a bribe – is paid to the policy maker to influence the policy outcome. It is not particularly surprising that, in such circumstances, corruption can change the alignment of citizen and leader preferences. Whether this is for good or ill depends on the notion of government failure under scrutiny.

To be specific, suppose that the policy maker can earn a private monetary rent of \(r > 0\) for setting \(e = 1\) regardless of whether the project is a good idea. We suppose that this payment \(r\) is a transfer made a subset of “organized” citizens. We will not model the transfer process explicitly. However, the general considerations of this section could be approached using the menu auction model due to Bernheim and Whinston (1986) which was first applied to political influence by Grossman and Helpman (1994). The basic idea in that approach is that government auctions off a policy to the highest bidder(s).

\(^{10}\)It is difficult to say a priori how problems of ignorance by government bear on government failures from either a distributional or Wicksellian point of view. This would depend heavily on the specifics of the situation.
Each of the bidders offers a menu which specifies a payment in exchange for a policy outcome. Here, for simplicity, we fix the size of the transfer made and the identity of those who make it.

Suppose that

$$ r > c/N. $$

In this case, the policy maker will implement the project regardless of his personal preference for or against it since the transfer he receives exceeds any share of the taxes that he will have to pay. The policy outcome is, therefore, always $e = 1$.

The utility of the citizens depends on whether they finance the transfer. Suppose that a fraction of $\beta < \gamma$ of citizens who favour the project collectively finance this transfer on an equal per capita basis. Those who favor the project and share the cost of the transfer, therefore, get utility of

$$ b - c/N - r/N\beta $$

while those in favor who do not pay receive:

$$ b - c/N. $$

Citizens who do not favor the transfer receive a payoff of $-c/N$.

Assuming that $b - c/N - r/N\beta > 0$, corruption of the form modeled here cannot generate a Pareto inefficient policy outcome in this setting. The transfers made are individually rational so that both the citizens and the policy makers are better off than they would be in the absence of corruption. The effect is purely a movement around the Pareto frontier. In general, we should expect this for pure bribery – those who pay and receive the bribe ought to be better off and hence it is unlikely that this can be Pareto inferior.\(^{11}\)

Since we have assumed that $\gamma N b > c$, corruption here actually increases social surplus relative to any policy which generates $e = 0$. Of course, things could go the other way. Had we assumed that $\gamma N b < c$, the opposite would have been true. The point is that there is no presumption that bribery raises or lowers social welfare in the abstract. However, in general corruption will reduce welfare when the benchmark is one where the government is purely benevolent. Thus, if our policy maker would have done the right thing in social welfare terms to begin with, then clearly bribery can only make things worse. On the other hand, in any model where there is already some potential imperfection in the resource allocation process to begin with, there is no a priori prediction about this. This “second best” theme is echoed in the analysis in chapter four.

\(^{11}\)This does depend, however, on the kinds of instruments that are available to compensate gainers and losers. Besley and Coate (2001) develops a model in which there are coordination failures between lobbyists that can result in a Pareto dominated policy outcome.
In terms of the Wicksellian definition, corruption only strengthens the tendency towards government failure by making it more likely that the project goes ahead. However, there is no need for things to go in this direction in general if we allowed for bribery to be among the group who wants smaller government. There may, however, be a tendency towards those who favor specific projects being more organized. Nonetheless, if the government were inclined towards implementing the project and the transfers were from those citizens who are against it, then this could reduce the chances that a Wicksellian political failure occurs. However, those who bribing to prevent the project from going ahead, would be worse off than if the government were constitutionally prevented from implementing the project while the policy makers would be better off. Hence, the possibility of government intervention can generate a set of transfers that constitute a government failure! This is true even if no intervention takes place. This further motivates the Public Choice pre-occupation with constitutional constraints – this being the only way to prevent this from happening.

Corruption here is only a label. Influence activities could equally well lead governments to do “good” as well as bad things. In cases where there are two sides to an issue, there will always be a group who feels that the policy process is stacked against them and will voice their concern about the way in which political influence is used.

Of course, there are good reasons to frown on corruption for other reasons than those emphasized here. First, we may not like per se the distributional effect of making transfers to politicians. It could also distort the allocation of talent if individuals enter public life to capture private benefits. Third, corruption could also induce a deadweight loss of its own if it is directed through inefficient means in order to keep it secret. Thus we could introduce a transactions cost on bribery whereby the citizens lose \( r \tau \) to deliver \( r \) to the policy maker where \( \tau > 1 \).

All of this notwithstanding, this analysis paints a somewhat more benign view of corruption than in a lot of other literature. That is not to say that the general opprobrium that greets corruption is not correct. First when the policy maker is elected in a legitimate democratic process, then we might expect bribery to undermine the democratic process. Moreover, one suspects that many of the policy favors granted through the process of corruption have no distributional merit according to any reasonable social welfare function. There are also hard to model systemic costs whereby corruption undermines norms of good behavior and the legitimacy of the state.\(^\text{12}\)

**Costly Rent-Seeking**

In our model of bribery, no resources were expended in influencing politicians’ decisions. In this section, we allow policy makers to be subject to costly influence activities – which can be thought of as rent-seeking or lobbying. In reality individuals expend resources in

\(^{12}\text{See Tirole (1996) for a model of collective reputations with multiple equilibria which he applies to the amount of corruption.}\)
order to secure political favors. To the extent that such activities involve hiring labour, people are drawn out of productive occupations.

The extensive literature on rent-seeking originating with Tullock (1967) and Krueger (1973) has studied how private actions influence policy. Following Tullock (1980), formal analysis has focused mostly on modelling competition among individuals or groups to obtain an indivisible policy favor, the aim being to characterize the aggregate expenditure on rent-seeking activities (see, for example, Baye, Kovenock and de Vries (1994) and the references therein). This fits very well the structure of the example that we are studying. It is frequently asserted that rent-seeking is a cause of government failure and hence it is important to see how it fits into our framework.

Suppose that citizen \( i \) can pay \( r_i \) to influence the policy maker either to go ahead with or desist from the project. We assume that \( r_i \) is a real resource cost, such as labour time, which cannot be appropriated by the policy maker. In fact, we assume that the policy maker gets no payoff from the influence (positive or negative). Suppose that each citizen commits resources \( r_{ia} (\geq 0) \) in favor of the project and \( r_{if} (\geq 0) \) against. Let total resources committed in favor of the project be \( R_f (\sum_{i=1}^{N} r_{if}) \) and the total against be \( R_a (\sum_{i=1}^{N} r_{ia}) \). Then assume that the probability that the project goes ahead is captured by the following simple function:

\[
\frac{R_f}{R_f + R_a}
\]

This kind of “contest” function is popular in the rent-seeking literature and will serve to make the main points of interest. We look for a Nash equilibrium in influence levels where all citizens have access to the influence technology and there is symmetric behavior within the two groups of citizens. It is easy to see that citizens who favor the project will not commit any resources lobbying against it and that the converse is true for those who oppose the project.

Consider the decision of citizen \( i \) who favors the project. His payoff if he contributes \( r_{if} \) is:

\[
\left( \frac{\sum_{k=1}^{N} r_{kf}}{\sum_{k=1}^{N} r_{kf}} \right) + R_a \left( b - \frac{c}{N} \right) - r_{if}.
\]

Citizen \( j \) who opposes the project has a payoff of:

\[
- \frac{R_f}{R_f + \left( \sum_{k=1}^{N} r_{ka} \right)} \frac{c}{N} - r_{ja}.
\]

We now solve for a Nash equilibrium where each citizen in favor and each against puts in the same effort level. Solving for the Nash equilibrium in the usual way, it is straight-
2.4. SOURCES OF GOVERNMENT FAILURE

forward to see that the equilibrium probability that the project goes ahead is:

\[ \left( 1 - \frac{c}{N b} \right) \]

The key magnitude here is \( c/bN \) – the ratio of the cost of construction per capita to the benefit to having the project for those who favor it. As the cost per capita becomes small (high \( N \) or low \( c \)), then the probability that the project is constructed goes to one. This is because only those who gain a benefit engage in influence activity – the cost of the project to those who are against is negligible.\(^{13}\)

The total expenditure on “rent-seeking” at a Nash equilibrium is:

\[ \frac{c}{N} \left( 1 - \frac{c}{N b} \right) \]

Using this, it is now straightforward to consider the welfare consequences of costly rent-seeking.

Evaluating the welfare consequences of costly rent-seeking requires considering two complications. First, the outcome needs to be evaluated both in terms of the policy that is chosen and the resources spent on influencing the decision which impose a social cost. The second issue is whether ex ante or ex post evaluation of the outcome is appropriate. For example, if the ex post policy outcome is \( e = 0 \) then everyone who committed positive resources to influence is worse off than they would have been without the possibility of political influence. But from an ex ante point of view, engaging in influence activities is still individually rational. The ex ante viewpoint therefore makes more sense.

The outcome with influence is ex post Pareto efficient. However, it can be Pareto dominated from an ex ante point of view by fixing the probability that the project is implemented at \( q = \left( 1 - \frac{c}{N b} \right) \). If citizens knew that this was fixed, they would not expend any resources on rent-seeking and would have the same probability distribution over the project outcome as in the Nash equilibrium with influence activities described above. In this sense, costly rent-seeking is always a source of government failure in the Pareto sense. However, to bring about the Pareto improvement would require having a policy technology in which the polity could commit to a (possibly random) policy allocation up front. Only by committing to this ex ante would influence activity be closed down.

Closing down costly rent seeking without using the same probability distribution over policy achieved in the Nash equilibrium of the influence game does not necessarily create a Pareto improvement. For example, an ex ante commitment of \( e = 0 \) is not Pareto superior to the Nash equilibrium with influence activity as the citizens in favor of the project going ahead are made worse off.

\(^{13}\)This is a feature of the project being a pure public good so that the cost of the project does not increase with the population size.
Turning now to ex ante total surplus (i.e. before the rent-seeking activity has been undertaken), we need to take into account the resources spent on influence. Aggregate (ex ante) surplus at the Nash equilibrium in influence is given by:

\[
\left(1 - \frac{c}{bN}\right) \left(N\gamma b - c\frac{(N + 1)}{N}\right).
\]

The first term is the equilibrium probability that the project goes ahead and the second is the surplus that it generates. Whether the latter is positive or negative depends on comparing the total benefit \((N\gamma b)\) with the total project cost \(c\) and the per capita project cost \(c/N\). Where \(c/N\) is small, it is clear that whether total surplus is positive or negative is not really affected by influence. However, the welfare economic calculus which compares \(N\gamma b\) with \(c\) clearly neglects the costs of rent-seeking activity and requires a stronger criterion for the project to go ahead.

From a surplus maximization point of view, the outcome with costly influence activities could be better or worse than what would happen without rent-seeking. If the outcome would be \(e = 1\) with probability one without influence, then influence makes things worse — recall that we have assumed that \(\gamma Nb > c\). But if it were \(e = 0\), the effect on social surplus is ambiguous.

All the same, (2.1) gives us a sense of how to compare market failure and government failure in the decision to go ahead with a public project from a surplus maximization perspective. Our “market failure” arose because we assumed that coordination failure lead to sub-optimal private provision of the public project. The additional influence costs now need to be weighed alongside the traditional welfare economic costs and benefits.\(^{14}\)

From a Wicksellian point of view, there is also an ambiguity as to whether a government failure has occurred. This depends on what the policy outcome would have been without the influence activity. If this were \(e = 0\), then allowing influence makes things worse in a Wicksellian sense as there is now a positive probability that the project goes ahead. However, if \(e = 1\) were to be the outcome in the absence of influence, then permitting influence could improve the situation from a Wicksellian point of view by empowering those who disapprove of the project.

We used a special and highly stylized model of how influence works. An important area of recent concern is influence activity in the form of campaign finance. A recent literature (see, for example, Grossman and Helpman (1996)) has studied how influence in the form of campaign contributions can distort policy in a model of electoral competition.\(^{15}\)

Coate (2004) develops a model where the process of political influence through campaign contributions is modeled explicitly. This boils down to something whose reduced form looks rather similar to rent-seeking models. He shows that, in some cases,

\(^{14}\)See Acemoglu and Verdier (2000) for a related analysis.

\(^{15}\)See also Persson and Tabellini (2000) and Besley and Coate (2001).
the probability distribution over policy is not affected at all by campaign finance and hence that banning campaign finance will result in a Pareto improvement. This provides an excellent illustration of the general principle that we have discussed and one that may have practical relevance.

However, as we have emphasized, dismissing of all influence activities as government failures is not very convincing. It depends on assuming that there is up front commitment to the ex post policy outcome. This does not seem the kind of scenario that most crude analyses have in mind. Without this, there are distributional effects as well as losses in resources that need to be weighed up in any careful welfare analysis.

There may of course be intrinsic concerns about the exercise of non-democratic influence. However, with any approach that emphasizes outcomes rather than processes, the trade-offs revealed here seem likely to prevail. As we observe in the case of bribery, once it is recognized that the policy mechanism that will prevail in the absence of influence is also flawed in some or another, there is really no reason to believe that the exercise of political influence is damaging in a wholesale way from a welfare economic point of view. That does not mean that empirically many cases of influence are indeed not damaging. Whether influence activities are socially costly has to be assessed on a case by case basis and by bringing empirical evidence to bear. Sweeping claims are far from self-evident.

2.4.3 The Quality of Leadership

The models that we studied above attach no weight at all to the quality of the leaders that hold policy authority.\textsuperscript{16} We now consider the possibility that the policy maker has an effect on policy. There are broadly two reasons why the type of the policy maker can matter. First, the quality of the policy could be embodied in policy makers. For example, some individuals can implement policies more cheaply or may even have more insight into what works. In this case, the only way to improve politics is to change the individuals who make policy. Second, some policy makers may be better at carrying out the citizens’ wishes. Whether this actually happens depends on the kinds of incentives that can be offered to policy makers for good behavior – good incentives might yield good policy regardless of the policy maker’s type. These issues will be discussed in greater detail in the models of chapters three and four. Here, we discuss – in a very simple way – how this can be a source of government failure.

Suppose that the policy maker in office is drawn from a pool of potential policy makers that are differentiated according to the cost of implementing the public project. Thus, we think of the task of policy formation depending on the human capital of the policy maker. Specifically, let $c_i \in \{c_L, c_H\}$ for $i$ in the set of potential policy makers. Thus $c_i$ is a measure of policy maker competence with cost type $c_L$ being a competent

\textsuperscript{16}This theme is discussed in Besley (2005).
policy maker. Suppose initially that

\[ b > \frac{c_H}{N} > \frac{c_L}{N}. \]

This implies that any policy maker who in office will desire to implement the project as long as he values it personally. Moreover all citizens who value the project also prefer that the project be implemented.\(^{17}\)

To think about government failure issues, the outcome has to include both a policy outcome (here whether \( e \) is zero or one) and a policy maker (type \( c_i \)) who implements that outcome. This generalized approach to think about policy making was suggested in Besley and Coate (1997) which used the citizen-candidate model of political competition to generate a policy outcome as well as a policy-maker’s identity.

With any type of citizen in office, the outcome where \( e = 1 \) is itself Pareto efficient. However, if a type \( H \) citizen is in power, then all citizens (including the policy maker!) is worse off than if a type \( L \) were choosing policy. Even those who don’t like having the project go forward would prefer that it was implemented by a citizen with cost type \( c_L \).

Thus, having a type \( H \) citizen in office constitutes a government failure since a Pareto improvement can be generated by replacing that citizen with a type \( L \). Thus, defining the outcome in terms of who is picked for office in addition to the policy outcome adds a further dimension to the set of possible government failures. This justifies a focus on issues on political selection in political economy models, i.e. worrying about who is selected to office. Clearly it would be in the interest of societies to generate systems of leadership selection that are sensitive to finding competent policy makers.

Turning to our other definitions of government failure, it is evident that having a type \( H \) policy maker in office may interfere with attaining a surplus maximizing policy outcome Moreover, if

\[ \frac{c_H}{N} > \frac{c_L}{N}, \]

a type \( H \) policy maker might implement the project even though it reduces social surplus.

Allowing for heterogeneous policy makers can work for or against the creation of Wicksellian government failures. To see this, consider what happens if

\[ \frac{c_H}{N} > b > \frac{c_L}{N}. \]

In this case, having a type \( H \) policy maker in office leads to the project (which was a government failure anyway) not being implemented. Type \( L \) policy makers in this world

\(^{17}\)By focusing on competence as a “common value”, policy makers do not get any benefit from their own incompetence. This differentiates competence and rent-seeking models of low quality leaders. In the latter, low quality leaders earn a rent from being in office. This makes it more difficult to create a Pareto improvement.
will tend to implement government projects which are harmful to citizens who have to pay for them without deriving any benefit. Thus, it is better to have incompetent policy makers who then decide that it is simply too costly to go ahead with them. This is a rather perverse logic – once government fails, it may be better to have more distortions. Again, it has to do with the nature of second best reasoning in these kinds of settings – the political system is already distorted, so adding a further distortion is not always welfare decreasing.

2.5 Sources of Political Failure

The issues that we discussed in the last section do not hinge on whether or not the government is democratic. In this section, we analyze two classic problems that arise in government in democratic settings. The first looks explicitly at the issue of when voting over policies and/or policy makers will result in a government failure. The second looks at decision making in legislatures and the problems of distributive politics. These are both areas where there has been extensive discussion of political failure. Thus a closer look at the basis of this merits some attention.

2.5.1 Voting

Voting is at the heart of democratic decision making and is the basis of political resource allocation in many instances. One important issue concerns whether voting is endemically linked to any form of political failure.

To explore this, suppose that the individual who has to make social decisions in this setting is chosen in an election and that there is a choice between two types of citizens – those in favor and those against the project. In this case, voting is purely a method for aggregating competing views about policy. The conflict of interest is purely between the citizens who are for and against the project. It is reasonable to suppose that the outcome will be to pick a citizen to choose policy whose policy preference coincides with majority opinion – the median voter outcome in this context.

Formally, this yields the following policy decision rule:

\[ e = 1 \text{ if } \gamma \geq 1/2 \text{ and } e = 0 \text{ otherwise.} \]

The outcome depends on which is the numerically largest group.

It is immediate from this that the median voter outcome is always Pareto efficient. However, there is some confusion about this point in some of the existing literature on political failure. For example, Bergstrom (1979) uses the Downsian model to analyze

\[18\]This could be modeled by means of a simple citizen-candidate game along the lines laid out in Besley and Coate (1997). In this simple setting, this will yield the same outcome as allowing direct voting over the policy outcome.
whether political competition will produce an efficient level of public goods. He shows that strong restrictions are needed for the median voter’s desired level of a public good to satisfy the Lindahl-Samuelson condition for the provision of a public good. The latter requires that the good be provided if and only if \( \gamma Nb > c \). Thus, the outcome with voting seems to coincide with the Lindahl Samuelson rule when \( \gamma \geq 1/2 \) implies \( N\gamma b \geq c \) which clearly does not hold in general. Indeed if

\[
\frac{b}{2} < \frac{c}{N}
\]

then there is always a value of \( \gamma \in \left(\frac{1}{2}, 1\right] \) which results in the project going ahead when the Lindahl-Samuelson rule rejects it.

But this disjunction between the Samuelson rule and the voting outcome is far from surprising. For the purposes of the voting analysis, we have fixed the tax system to be a uniform head tax. By contrast, the Lindahl-Samuelson rule as an efficiency criterion assumes that there are lump sum taxes and transfers, at least if the losers are to be compensated by the gainers. As we have seen from our example, holding the method of financing fixed, the majority voting outcome is trivially efficient. These issues are bound to arise in models which work with restrictions on the policy space. But it is important to consider the welfare comparison holding the policy instruments fixed.

This discussion makes plain why the outcome achieved under majority voting and based on social surplus will almost certainly diverge in general. Voting cannot be used to register preference intensity. However, it is essential to most distribution criteria that the intensity of preferences for and against – as well as the numerical strength of these groups – count.

The outcome achieved under voting is a Wicksellian political failure if \( \gamma \geq 1/2 \) since the group that is against the project going ahead would be better off in the no-government status quo. This observation that Wicksellian justice is inconsistent with majority rule is a core observation in Buchanan and Tullock (1962)’s critique of democratic policy making.

The example makes clear that there can be no general presumption of efficiency or inefficiency from the median voter outcome (except in the Paretian sense). In fact the result on Pareto efficiency holds true whenever the political equilibrium picks a Condorcet winner. When it comes to distributional criteria and the Wicksellian ideal, then it is a fair to say that there is a good reason to think that voting will lead to political failure.

We now consider a different argument due to Fernandez and Rodrik (1991). They apply the argument to economic reforms in general. However, our public project could easily be thought of as a program of economic reform as building a single public project. Suppose that \( \gamma < 1/2 \), i.e. a minority of the population would gain \( b \) for sure from the

---

19 This point was noted in an important article by Wittman (1989).
2.5. SOURCES OF POLITICAL FAILURE

project. However, unlike the baseline model, we suppose that the remaining \((1 - \gamma)\) of the population do not know whether they will receive \(b\) from the project. Specifically we suppose that, with probability \(\pi\), they receive \(b\) if the project goes ahead and with probability \((1 - \pi)\) they receive nothing. Suppose also that:

\[ \pi b < \frac{c}{N} \]

so from an ex ante point of view, those who are uncertain about whether they will gain are opposed to the project, i.e. would vote against it. Thus, with majority rule, \(e = 0\) will be the policy outcome. We know, however, that if \((\gamma + (1 - \gamma) \pi) b > c/N\), then this is a political failure in terms of ex ante social surplus. That said, this insight is not substantively different from the basic case without uncertainty.

More interesting, is the observation by Fernandez and Rodrik (1991) that if

\[ \gamma + (1 - \gamma) \pi > \frac{1}{2} \]

a majority would vote for the reform \textit{ex post}, i.e. if the identity of the gainers and losers were already known. Thus individual-specific uncertainty is responsible for the fact that \(e = 0\).\(^{20}\)

Whether this kind of individual-specific uncertainty is good or bad from a welfare point of view is unclear. It has no bearing on whether policy choices are Pareto efficient. It will, however, tend to reduce the probability of a Wicksellian political failure and hence could be viewed as welfare enhancing in models that value greater inertia in the policy process.\(^{21}\) In terms of social surplus, things could go either way. For example, if

\[ (\gamma + (1 - \gamma) \pi) b < c/N, \]

then individual specific uncertainty can prevent projects that fail the surplus maximizing criterion being implemented even though they would be implemented under certainty.

Thus, while introducing uncertainty does yield some interesting insights, it does not give a new source of political failure in public decision making.

2.5.2 Log-rolling and Legislative Behavior

So far, we have discussed highly simplistic models of the policy process. However, in practice policy processes work with interactions among a group of politicians collectively

\(^{20}\)Fernandez and Rodrik (1991) draw the conclusion that this will lead to status quo bias in reform. This point turns out not to be robust in their framework – see Ciccone (2004).

\(^{21}\)Observe that if \(\pi b > c/N\) and \((\gamma + (1 - \gamma) \pi) < 1/2\), then a project that would be rejected under majority with certainty would be implemented were there uncertainty about who gains and who loses. Thus, uncertainty need not be a device to minimize Wicksellian political failures. Note also that ex post a vote to abandon the project would be successful in this case. I am grateful to Sanjay Jain for discussion on this point.
CHAPTER 2. THE ANATOMY OF GOVERNMENT FAILURE

charged with making policy decisions. One important example of this is policy making in legislatures. The question is how agreements are structured according to the rules of the legislature and how this affects policy outcomes.

The seminal work of Tullock (1959) and Buchanan and Tullock (1962) emphasized how log-rolling in legislatures could lead to policy distortions. Tullock (1959) contrasts two systems of social decision making – straight majority decision making (what he calls a referendum) and some kind of bargaining between subgroups of voters which he calls logrolling. He gives the following insightful account of the main ideas:

“A township inhabited by one hundred farmers who have more or less similar farms is cut by a number of main roads maintained by the state. However, these roads are limited access roads, and the farmers are permitted to enter the primary network only at points where local roads intersect it. The local roads are built and maintained by the township. Maintenance is simple. Any farmer who wishes to have a specific road repaired puts up the issue to vote. If the repairing is approved, the cost is assessed to the farmers as part of the real property tax. The principal use of the local roads by the farmers is to get to and from the major state roads. Since these major roads cut through the district, generally there are only four or five farmers dependent on any particular bit of local road to reach the major road.

Under these circumstances, the referendum system would result in no local roads being repaired as an overwhelming majority would vote against repairing any given road. The logrolling system, however, permits the roads to be kept in repair through bargains among voters.” (page 573).

An instructive benchmark for legislative policy making is a Coasian view that supposes that a group of individuals will reach joint agreements which internalize any externalities between them due to differences in policy preferences. This kind of Coasian bargain will result in policies that are efficient for the legislatures and hence are Pareto efficient from the point of view of the economy as a whole. Hence, the presumption will be that policies will be efficient. Even in this case, there is no reason to think that legislatures will satisfy Wicksell’s concerns – there is plenty of scope for legislatures to pick policy outcomes which do not Pareto dominate the no-government status quo.

Whether efficient bargains can be generated in legislatures given the kinds of rules that obtain for reaching agreements is moot. Clearly it requires investigation by developing models of bargaining in legislatures and there are many important contributions that look at this. Here is not the place to discuss these models in general. However, it is useful to observe that these models can sometimes create sources of political failure in the sense of Pareto (in)efficiency.

\[22\] Since legislators are citizens, they will pick some kind of agreement that is at least as good as any other for the group of legislators.
One classic example of this is the well-known model due to Shepsle, Weingast and Johnsen (1981). They consider a legislature which allocates a number of public projects to various districts. They suppose that each project benefits one and only district. Each district is represented in a legislature by one elected representative. The notion they propose is that the legislature operates according to a “norm of universalism” in which each district gets a project provided that all the others do. But each district in deciding how far to fund a project will realize that it bears only a fraction of the cost – the remainder being born by residents of the other districts. The result is an inefficiently large number of projects being financed. The outcome can sometimes be Pareto dominated from the point of view of the members of the legislature – everyone could be made better off with a collective reduction in the levels of the public goods being funded. Thus, the outcome does constitute a political failure in the Paretian sense. However, this happens because the model simply excludes the possibility of Coasian bargains among the legislators without stating the reason. To tackle the problem of inefficiency would require developing some other rule of operation that can internalize policy externalities.

To explore this issue, we extend the model above in a very simple way. Suppose now that policy decisions are made in a legislature comprising representatives selected from geographic regions. The \( n \) districts that they represent are labeled \( j = 1, ..., n \). Each district is of equal size containing \( m \) citizens so \( m \times n = N \). A project can be undertaken in each district and is enjoyed solely by the residents of that district, i.e. there are no spillovers across districts. Thus the legislature can authorize the building of up to \( n \) projects (one for each district).

Let \( e_j \in \{0, 1\} \) denote whether a project goes ahead in district \( j \). We assume that there is common pool financing – the taxation levied on each citizen across the districts is equal to the total cost of projects that are financed divided equally by all citizens in the polity (regardless of residence). We also suppose (following Shepsle, Weingast, and Johnsen (1981)) that project allocation is governed by a rule in which the representative in each district can unilaterally decide whether to implement a project in its district as long as all other legislators enjoy that privilege.

For simplicity, we suppose that each representative maximizes the average utility of a district resident which is

\[
e_j \gamma b - \frac{\sum_{k=1}^{n} c_{k}}{N}.
\]

Note that this assumes that each district comprises an equal fraction (\( \gamma \)) of citizens who are in favor of the project.

It is apparent that the representative in any district will wish to have a project located in his district provided that

\[
N \gamma b > c.
\]

In effect, he compares the benefits as if they accrued to the whole population with the
cost – this is because the cost is shared with other districts. But, from a social surplus point of view, a project is desirable only if the surplus that it generates in the district that it is located in is positive, i.e., if \( m \gamma b > c \).

Thus the legislative process that we have posited along with common pool financing will yield excessive publicly financed spending (according to the social surplus criterion) if

\[
N \gamma b > c > m \gamma b.
\]

The policy outcome is Pareto inefficient if

\[
\frac{c}{\gamma N} < b < \frac{c}{m}.
\]

Then even the citizens in favor of the project would prefer not to have it. Hence, it corresponds to a political failure. The outcome can be Pareto dominated by a cooperative solution in the legislature, i.e. one where all projects are simultaneously agreed upon rather than delegating that decision to the representative within a district. This raises the issue of why the norm of universalism would never be seen in the first place. An efficient solution could also be found by using a tax system which tried to share the costs of project financing so that each district paid more taxes if it had a project located within it, moving away from common pool finance.

The outcome under the norm of universalism is a political failure according the to the Wicksellian criterion if \( N \gamma b > c \). Again the tendency for political resource allocation is to authorize too many projects. Thus, the posited rules of legislative decision making lead to an outcome which can fail according to all three definitions of political failure.

Since legislatures that are asked to allocate resources across space are a central feature of democratic policy making, it is clear that political failures generated this way merit close attention. However, it is also apparent that the policy outcome depends on the details of legislative processes, a key issue being why Coasian bargains among legislators cannot be struck. While the above kind of analysis has been influential among economists, the fact that the norm of universalism is maintained is puzzling and it is important to give persuasive micro-foundations before using it. That said, those who have looked at legislative decision making in more detail confirm that this can be an important source of political failure.\(^{23}\)

### 2.6 Dynamics

We now explore government failure in models where policy making takes place in more than one time period. This is important since it highlights the role of commitment

\(^{23}\)See Battaglini and Coate (2005) for an overview and development of a dynamic model of legislative bargaining applied to an economic policy model.
2.6. DYNAMICS

issues and re-election concerns in shaping policy. We will develop an example with two time periods. The main issues arise because there are linkages between decisions made in period one and period two. We identify three sources of linkages that are potentially important.

1. **Investment linkages**: This refers to cases where private investment decisions are affected by future economic policies. The classic time-consistency problem in government policy first identified by Kydland and Prescott (1977) falls into this category.

2. **Political linkages**: If policy choices in period one affect the type of policy maker who is in office in period two, then period one policies are tied to those that are undertaken in period two.

3. **Policy linkages**: Policies that are undertaken in period one affect the policy choices that incumbents make in period two.

With either political or policy linkages, period one policy choices acquire a *strategic* element if policy makers are forward looking.

### 2.6.1 Investment Linkages

Suppose now that there are two time periods labeled \( t \in \{1,2\} \). A project can be implemented in each period and the policy decision in period \( t \) is denoted by \( e_t \in \{0,1\} \).

As above, the cost of the project is \( c \) in each period and is financed equally by all \( N \) citizens. Citizens receive some period one payoff from the project denoted by \( b_i(e_1) \). But we allow citizens to make private investment decisions denoted by \( x_i \in \{0,1\} \) for citizen \( i \) which costs \( \kappa x_i \). The payoff to citizen \( i \) from the period two project is denoted by \( B_i(x_i, e_1, e_2) \) i.e., it depends on their investment decision in period one and the actions of government in both time periods.

We consider two cases:

1. **Commitment**: The government chooses \( (e_1, e_2) \) and then citizens choose whether or not to invest.

2. **No commitment**: The government first chooses \( e_1 \), then citizens invest. The government then chooses \( e_2 \).

We study each in turn.
Commitment

In this case, we will use \( * \) to denote the decisions made by government and citizens. Let \( \{e_1, e_2\} \) be a fixed pair of government policies. Then the optimal investment decisions by the citizens are described by:

\[
x_i^* (e_1, e_2) = \arg \max_{x_i \in \{0,1\}} \{ B_i (x_i, e_1, e_2) - \kappa x_i \}.
\]

(2.2)

Suppose that the government is benevolent, caring about social surplus. Then:

\[
\{e_1^*, e_2^*\} = \arg \max_{e_2 \in \{0,1\}} \left\{ \sum_i (b_i (e_1) + B_i (x_i^* (e_1, e_2), e_1, e_2) - \kappa x_i^* (e_1, e_2)) - c (e_1 + e_2) \right\}.
\]

We make two assumptions:

**Assumption 1:**

\[
(B_i (1, e_1, 1) - B_i (0, e_1, 0) - \kappa) - \frac{c}{N} > 0
\]

for all \( i \) and for all \( e_1 \in \{0,1\} \).

This says that any situation in which government chooses \( e_2 = 1 \) and citizen \( i \) chooses \( x_i = 1 \) for all \( i \), Pareto dominates an outcome where \( e_2 = 0 \) and \( x_i = 0 \) for all \( i \).

**Assumption 2:**

\[
B_i (1, e_1, 1) - B_i (0, e_1, 1) > \kappa > B_i (1, e_1, 0) - B_i (0, e_1, 0)
\]

for all \( i \) and for all \( e_1 \in \{0,1\} \).

This says that only if citizens predict that the government will choose \( e_2 = 1 \), then it is optimal for them to invest in period one, i.e. set \( x_i = 1 \). In other words, the marginal return to private investment is increased by the period two investment project. Together Assumptions 1 and 2 imply that it is surplus maximizing (and hence Pareto efficient) to have \( e_2^* = 1 \).

Assuming that the period two policy is optimal, then the socially optimal period one policy will be \( e_1^* = 1 \) if and only if:

\[
\sum_i (b_i (1) + B_i (1, 1, 1)) - c > \sum_i (b_i (0) + B_i (1, 0, 1))
\]

which simply compares the cost of the project with the benefit.

This provides a benchmark for studying the case where the government cannot commit up front.
2.6. DYNAMICS

No Commitment

Ever since the classic analysis of Kydland and Prescott (1977), it has been known that government’s inability to commit to a policy ahead of time can reduce welfare. We now study this in our framework. We will use a “hat” (\(^\) ) above a variable to denote the equilibrium outcome in this case.

In this case, we work backwards beginning with the government’s second period decision. The government will take \((x_1, ..., x_N, e_1)\) as given. Its optimal time-consistent policy will satisfy:

\[
\hat{e}_2 (x_1, ..., x_N, e_1) = \arg \max_{e_2 \in \{0, 1\}} \left\{ \sum_i B_i (x_i, e_1, e_2) - ce_2 \right\}.
\]

We next consider the investment decision by the citizens. They take the period one policy choice and form (rational) expectations about government policy in period two. Thus:

\[
\hat{x}_i (e_1) = \arg \max_{x \in \{0, 1\}} \left\{ B^i (x, x_1, e_1, e_2 (x_1, ..., x, ..., x_N, e_1)) - \kappa x - \hat{e}_2 (x_1, ..., x, ..., x_N, e_1) \frac{c}{N} \right\}.
\]

The fact that the period two policy outcome depends on the full vector of period one investment decisions implies that private investment decisions are now interdependent. We thus look for a Nash equilibrium in private investment decisions.

**Assumption 3:**

\[
\sum_i [B_i (x_i, e_1, 1) - B_i (x_i, e_1, 0)] - c < 0
\]

for all \(e_1 \in \{0, 1\}\) and all \(x_i\) for \(i = 1, ..., N\).

Then, if every citizen invests in period one, the government chooses \(\hat{e}_2 (1, ..., 1, e_1) = 0\), i.e., not to implement the project in period two.

This implies that the time-consistent policy without commitment, has \(x_i = 0\) for all \(i \in \{1, ..., N\}\) and \(\hat{e}_2 = 0\). The period one policy will be \(\hat{e}_1 = 1\) if and only if:

\[
\sum_i (b_i (1) + B_i (0, 1, 0)) - c > \sum_i (b_i (0) + B_i (0, 0, 0)).
\]

Assumption 1 implies that the policy achieved by a benevolent government is Pareto dominated. Thus, the failure to commit to \(e_2 = 1\), constitutes a government failure in the Pareto sense.
CHAPTER 2. THE ANATOMY OF GOVERNMENT FAILURE

This policy cum investment outcome will also constitute a failure according to any kind social welfare function that respects the Pareto criterion, including surplus maximization. Hence, it is also a distributional failure.

It is less clear whether there is a government failure in Wicksell’s sense. If

\[ b_i (1) + B_i (0, 1, 0) - c/N < b_i (0) + B_i (0, 0, 0) \]

for some \( i \), there will be a Wicksellian government failure if \( \hat{e}_1 = 1 \) since some citizens are made worse off by implementing the project. The fact that \( \hat{e}_2 = 0 \) means that a Wicksellian failure may be avoided in period two compared to the full commitment outcome as long as some citizens prefer the period two project not to be implemented.

This simple example illustrates the classic time-consistency problem as it afflicts a benevolent government. It arises here because the government’s policy preference changes once the private investment decisions have been made. However, this example of government failure has nothing to do with politics. There are no elections in the story – a benevolent dictator without commitment power would generate this kind of problem. However, the time consistency problem is due to an intertemporal linkage created from the fact that there is a durable private decision by the citizens – in this case the private investment decision – that affects government policy incentives. We now show how intertemporal political and policy linkages can induce government failures. Note also that the time-consistency problem as generated here is not due to inefficiencies in period one policy making. This will contrast with those that we demonstrate in the next section.

2.6.2 Political and Policy Linkages

As we have seen, commitment problems arise even when the identity of the (benevolent) policy maker remains fixed over time. But the essence of political economy models is the selection and turnover of policy makers through a political process. In this section, we explore the consequences of this for policy choice. To focus on the issue of political and policy linkages, we will now abstract from private investment decisions and hence suppress \( x_i \) from the analysis.

A key additional consideration in this section is to describe the process of survival and turnover which determines the probability that the policy maker remains in office as a function of his period one policy choice. Here, we will develop a “black box” approach to the process of turnover among policy makers. However, it could be given a microfoundation. For example, Besley and Coate (1998) develop an analysis of turnover based on a citizen-candidate model of politics. The political agency models that we study in chapters three and four can also provide a foundation for this. In these models, as we shall see, the key idea is that turnover reflects political accountability. They make sense of this using a model of imperfect information where either the policy outcome or the
politician’s type is uncertain ex ante. Considerations of political turnover could equally well be studied in a model of non-democratic politics as in Acemoglu and Robinson (2002).

We now suppose that there are two groups of citizens in the population – those who have a high level of demand for government projects and those who have low demand. We use \( \tau \in \{ L, H \} \) to denote the type of each citizen. Let \( \gamma_t \in [0, 1] \) be the proportion of citizens who are of type \( H \) in period \( t \). We allow for the possibility of turnover in the electorate which leads \( \gamma \) to change over time. There could also be some uncertainty about the future proportions of citizens of each type.

Let \( b_t(0) = 0 \) and
\[
 b_t(1) = \begin{cases} 
 \bar{b} & \text{if } \tau = H \\
 \bar{b} & \text{if } \tau = L.
\end{cases}
\]

be the period one payoff from implementing the period one government project with \( \bar{b} > \bar{b} \). The period two project payoff depends on whether the project was undertaken in period one. Let \( B_t(e_1, e_2) \) be the period two valuation of a type \( i \). We assume that \( B_t(e_1, 0) = 0 \) and:
\[
 B_t(e_1, 1) = \begin{cases} 
 \bar{B}(e_1) & \text{if } \tau = H \\
 \bar{B}(e_1) & \text{if } \tau = L.
\end{cases}
\]

where \( \bar{B}(e_1) > \bar{B}(e_1) \).

It is evident from this that we have a policy linkage between the two time periods since the demand for the period two project could be affected by whether it is implemented in period one. As above, the cost of the project is \( c \) in each period and is divided equally across all \( N \) citizens.

We suppose that a citizen holds office in each period. This means that he must either be of type \( L \) or type \( H \). Hence we consider only policy choices that are optimal for some kind of citizen as a policy maker. In addition to policy concerns, a policy may also care about the rent from holding office which we denote by \( E \geq 0 \). We suppose that a policy maker of a particular exogenously given type is in office in period one.

There are two aspects of political turnover to consider. First, there are concerns that the policy maker has about his own survival in office. This is particular important in the presence of rents. Second, there is the issue of what type of policy maker will be in office in future. So a policy maker could lose office, but be guaranteed that he would be succeeded by someone with the same policy preferences. Hence, only the personal rent to holding office is lost.

Let \( \sigma_t \in \{ L, H \} \) denote the type of the policy maker in office in period \( t \). We capture these two aspects of turnover as follows. Let \( \pi(\sigma_1, e_1) \in [0, 1] \) denote the probability that an incumbent of type \( \sigma_1 \) is re-elected as policy maker for period two as a function of

\[\text{We have not imposed the assumption that } \bar{b} = 0 \text{ since we want it to be possible for every citizen to favor implementing the project in period one.}\]
the period one project that he implements. There are three possibilities. The project is *politically neutral* if \( \pi(\sigma_1, 1) = \pi(\sigma_1, 0) \), *politically advantageous* if \( \pi(\sigma_1, 1) > \pi(\sigma_1, 0) \), and *politically damaging* if \( \pi(\sigma_1, 1) < \pi(\sigma_1, 0) \). Let \( q \in [0, 1] \) be the probability that the second period incumbent is of type \( L \). There are a number of ways to motivate \( q \neq 1 \). For example, we could suppose that \( \gamma \) is a random variable due to changes in the electorate or turnout so the median type could be of type \( H \) or \( L \). We let \( ^\ast e_t \) denote the equilibrium project choice in each period.

The timing of the model is as follows. In period one there is an (exogenously given) incumbent policy maker of type \( \sigma_1 \) in office who chooses \( e_1 \). Nature then determines whether the incumbent is replaced according to \( \pi(\sigma_1, e_1) \). The latter could represent either the outcome of an election or a power struggle in an autocratic system. If the incumbent is replaced, then nature determines the type of the new incumbent. The period two incumbent then chooses \( e_2 \). We study the outcome working backwards beginning with the period two policy choice.

It is straightforward to see that the period two policy choice depends solely on the type of the period two policy maker, i.e.

\[
\hat{e}_2(\sigma_2, e_1) = \begin{cases} 
1 & \text{if } B_{\sigma_2}(e_1, 1) \geq c/N \\
0 & \text{otherwise.}
\end{cases}
\]

Now let

\[
W^\tau(e_1, \sigma_2) = \left\{\hat{e}_2(\sigma_2, e_1) \left[ B_\tau(e_1, \hat{e}_2(\sigma_2, e_1)) - \frac{c}{N} \right] \right\}, \quad \tau \in \{L, H\}
\]

be the period two utility of citizen of type \( \tau \) as a function of the period one policy choice and type of the period two policy maker.

Now consider the period one policy maker’s choice. This has a forward looking element. The policy maker’s preference is:

\[
e_1 \left( b_{\sigma_1}(e_1) - \frac{c}{N} \right) + \pi(\sigma_1, e_1) \left( E + W^{\sigma_1}(e_1, \sigma_1) \right) + (1 - \pi(\sigma_1, e_1)) \tilde{W}^{\sigma_1}(e_1) \tag{2.3}
\]

where:

\[
\tilde{W}^\tau(e_1, q) = qW^\tau(e_1, L) + (1 - q)W^\tau(e_1, H)
\]

is his expected period two payoff if he is not re-elected. This reflects the uncertainty about the period policy maker’s type.

The optimal period one policy for an incumbent of type \( \sigma_1 \) is therefore:

\[
\hat{e}_1(\sigma_1) = \arg \max_{e_1 \in \{0, 1\}} \{e_1 \left( b_{\sigma_1}(e_1) - \frac{c}{N} \right) + \pi(\sigma_1, e_1) \left( E + W^{\sigma_1}(e_1, \sigma_1) \right) + (1 - \pi(\sigma_1, e_1)) \tilde{W}^{\sigma_1}(e_1, q)\}
\]

This illustrates the strategic aspects of policy making that arise in a dynamic political economy setting. This equation embodies the three main considerations that shape policy making in dynamic political settings:
• **Short term policy considerations** – these are represented by \( b_{\sigma_1}(e_1) - \frac{\xi}{N} \).
This depends on whether the policy could be worthwhile or otherwise in terms of its current costs and benefits.

• **Long-term policy considerations** – these are reflected in the dependence of \( W^{\sigma_1}(e_1, \sigma_1) \) and \( \bar{W}^{\sigma_1}(e_1) \) on \( e_1 \). This term arises since the the period one policy may affect payoffs from future policies. This is the source of policy linkages.

• **Survival considerations** – This is represented by the way in which \( \pi(\sigma_1, e) \) depends on \( e_1 \) and affects the probability that the incumbent will survive as a function of his period one policy choice. This is the source of political linkages.

We now study the implications of policy choices that maximize (2.3). We will illustrate both policy and political linkages. Before doing that, however, we develop a benchmark result based on Besley and Coate (1998). Suppose that

- (i) the policy is politically neutral, i.e. \( \pi(\sigma_1, 0) = \pi(\sigma_1, 1) \) for \( \sigma_1 \in \{L, H\} \)

- (ii) the payoff from the period two project choice is unaffected by the period one project choice, i.e. \( B_{\tau}(0, e_1) = B_{\tau}(1, e_1) \) for \( \tau \in \{L, H\} \).

Then the policy choices \((e_1, e_2)\) will be Pareto efficient.

To see why this is true, observe that if conditions (i) and (ii) hold, then the expected period two payoff of the current policy maker (and the citizens in the polity) do not depend on \( e_1 \) at all. Thus, the period one policy choice is determined solely by whether \( b_{\sigma_1}(e_1) \) is bigger or smaller than \( c/N \). But then the model effectively reduces to a one period model with all intertemporal links having been severed.

Even though the outcome in this case is guaranteed to be Pareto efficient, there is no reason to expect it to maximize social surplus or to avoid government failures of a Wicksellsian variety. The reasoning here follows exactly that in the static model developed above.

We now consider what happens when either conditions (i) or (ii) above fail. The first is the case of a political linkage and the second that of a policy linkage. We now show that in each case, we can construct an example where the policy choice is Pareto inefficient and hence a government failure in this sense.

**Political Linkages**

Our first example illustrates a case where the desire of an incumbent to survive in office leads to a Pareto inefficient policy choice. Suppose that all citizens wish to implement

---

25In principle, we could also make \( q \) depend on \( e_1 \).
the first period project, i.e., $h > c/N$. The citizens disagree, however, about period two policy. Specifically, we assume that

$$B(0) > \frac{c}{N} > B(1)$$

and

$$\bar{B}(1) \geq \bar{B}(0) > \frac{c}{N}.$$  

The first of these says that the low types want the project in period two only if it was not implemented in period one. The second says that the high valuation types want the project to be implemented regardless of whether the first period project is implemented.

Given these assumptions on preferences, any situation where the first period project is not implemented constitutes a government failure in the Pareto sense – all citizens are better off for any fixed policy decision in period two. We are interested, therefore, in studying cases where this does not happen. The starkest case of this is where it is politically costly to set $e_1 = 1$. Suppose that:

$$\pi(\sigma_1, 1) = 0 < \pi(\sigma_1, 0) = 1.$$  

This says that the period one incumbent will survive in office only if he/she fails to implement the period one project.\(^{26}\)

It is clear that for large enough $E$, the period project will not be implemented, i.e. $\hat{e}(\sigma_1) = 0$ for $\sigma_1 \in \{L, H\}$. This makes sense. The incumbent has desire to hold on to office since the rents are large, but he will lose office if he chooses to implement it. However, a government failure can also occur if $E = 0$. To see this, let $h - c/N$ become small so that the direct benefits of the project are small to a type $L$. Then if $\sigma_1 = L$, the incumbent will suffer a utility loss of $B(1) - c/N$ if a type $H$ policy maker holds office in period two as the latter will set $e_2 = 1$. Thus as long as $q$ (the probability of a type $L$ policy maker) is closer enough to zero, it will be optimal for a type $L$ policy maker to set $e_1 = 0$. In this case, it is the decision rent from controlling the policy process rather than the direct rent from holding office which induces a government failure.

This example illustrates how the desire to survive in office affects the efficiency of policy choice and leads to a Pareto inefficient outcome. This arguments extends easily and unsurprisingly to other social objectives such as surplus maximization. It is more subtle in Wicksellian terms. From a Wicksellian point of view, the right outcome is $e_1 = 1$ and $e_2 = 0$. By failing to implement the project in period one, the type $L$ incumbent prevents a period two government failure.

\(^{26}\)A micro-foundation for this could be generated in an agency model of the type studied in the next chapter. This approach was used in the key contribution of Coate and Morris (1995). Besley and Coate (1998) also develops a model which is consistent with this kind of outcome using a citizen-candidate approach.
2.6. DYNAMICS

Even though the welfare consideration is not always transparent, the idea that political survival can play an important role in affecting economic policy choice has been widely studied.\textsuperscript{27} For example, Aghion and Bolton (1990) and Milesi-Ferretti and Spolaore (1994) develop models in which strategic policy choice can also lead to changes in who is elected. For example an incumbent may realize that if he runs a deficit then this can make election of the challenger less attractive. These ideas are applied to privatization policy in Biais and Perrotti (2002). Many governments have encouraged privatizations because they create a class of share-holders who then show favors towards right wing governments. This will encourage governments to underprice privatizations to create a class of stakeholders. Besley and Coate (1998) pulls these ideas together and show that these strategic effects can be sources of real inefficiencies.

A variety of papers look at how political incentives for re-election shape the kind of public projects that are chosen. Glazer (1989) discusses how this can lead to projects which have long-run effects being preferred to those that payoff only in the short run since these have effects on political equilibria. Coate and Morris (1995) use an agency model of the kind discussed in chapters 3 and 4. They show how the fact that policy choice affects re-election chances leads to inefficient policies being chosen. This is driven by the desire of politicians to capture rents by being re-elected. Robinson and Torvik (2005) discuss how public project choices can provide a commitment device since only certain kinds of politicians will continue with these projects in future. Jain and Mukand (2003) extend the model of Fernandez and Rodrik (1991) to allow for an explicit dynamic political economy model where the government can compensate the losers in period two for the consequences of a project choice in period one. They also find that whether the leader survives is important to understanding period one policy choice.

These insights have also been useful in understanding how politics affects growth and development. There are now a number of models that develop this theme using the insight that government efforts to promote growth can have adverse political consequences for incumbents. For example, Acemoglu and Robinson (2002) build a theory of underdevelopment based on the possibility that governments are unwilling to invest in productive things because they affect the politician’s tenure. In similar vein, Krusell and Rios-Rull (1996) develop a model where politicians can affect the future political equilibrium via today’s policy choices leading to stagnation.

Policy Linkages

We turn now to policy linkages. To focus on these, suppose that the period one policy is politically neutral, i.e., $\pi(\sigma_1, 1) = \pi(\sigma_1, 0) = 0$. This could describe a case where all politicians are term-limited so that considerations of rents ($E$) cannot distort their policy choices. Hence, political rents play no role in this example.

\textsuperscript{27}Besley and Coate (1998) studies the implications of political equilibrium in a dynamic model using the Pareto criterion.
We suppose that type of policy maker who controls period two policy is uncertain with an equal chance of any kind of period two policy maker \( q = \frac{1}{2} \).\textsuperscript{28} As above, we assume that there is universal demand for the period one project, i.e. \( b > c/N \). However, we now suppose that type \( H \) citizen demand the project in period two only if the period one project was implemented, i.e.,

\[
\bar{B}(1) > \frac{c}{N} > \bar{B}(0).
\]

However, type \( L \) citizens never desire the project in period two:

\[
\frac{c}{N} > \bar{B}(0) = \bar{B}(1).
\]

As above, this implies that any situation in which the period one project is not implemented is Pareto inferior.

Now consider what happens if the first period policy maker is of type \( L \). He/she will be replaced by a type \( H \) with probability \( \frac{1}{2} \). The latter will implement the project only if \( e_1 = 1 \). Thus if:

\[
b - \frac{c}{N} + \frac{1}{2} \left[ \bar{B}(1) - \frac{c}{N} \right] < 0
\]

then a type \( L \) incumbent will set \( e_1 = 0 \).

The logic here differs from the case of political linkages – the period one project choice has no effect on whether the incumbent survives. The effect is driven from the fact that:

\[
1 = \hat{e}_2 (H, 1) > \hat{e}_2 (H, 0) = 0.
\]

Again, it is clear that this kind of logic can underpin government failures using a wide set of welfare functions. The Wicksellian implications are similar to the example with political linkages. The period one policy maker of type \( L \) prevents a period two government failure by setting \( e_1 = 1 \). Hence, this does not constitute a government failure in the Wicksellian sense.

A number of papers in the literature explore policy linkages and their implications. For example, Tabellini and Alesina (1990) and Persson and Svensson (1989) all explore why political equilibria can lead to excessive fiscal deficits if debt is used strategically to influence policy outcomes beyond a politician’s current term. They show that a larger deficit prevents the future incumbent from spending on his own preferred policies. This is more likely to happen if there is greater ideological conflict between the incumbent and his potential challenger.

\textsuperscript{28}This could easily be underpinned by a political model where the two groups are of similar size and the policy maker is elected from a contest between one representative from each group.
2.6.3 Investment and Politics

To round off our discussions, we now consider the possibility that private investment decisions affect policy choices – an issue that we abstracted from in the last sub-section. However, we study this using a somewhat different policy example than that which has run through the rest of this chapter – one where investments increase private productivity and government policy is in the form of tax and transfer policy.

Suppose that the citizens in a polity are divided into three groups. A fraction $\gamma_H$ of high income citizens earn $y_H$, a fraction $\gamma_L$ low income citizens earn income $y_L (< y_H)$ and a fraction $\gamma_M$ of mobile citizens earn income $y_L$ unless they make a private investment that costs $\kappa$ in which case they earn $y_H$. We assume that $\kappa < (y_H - y_L)$ so that the investment is worthwhile. Let $x_i \in \{0, 1\}$ be the investment decision of the $i^{th}$ mobile citizen. We assume that the utility of each citizen depends solely on their consumption which equals their post-tax and transfer income.

The two period structure of the last section is maintained with these productivity enhancing investments being undertaken in period one and realized in period two. As above, we suppose that a policy maker is a citizen and hence corresponds to one of the three types of citizen that we have described. As a policy maker, we assume that the government can set the rate of a tax on income that we denote by $t \in [0, \bar{t}]$ where $\bar{t} < 1$. The latter could be due to the ability of citizens to retreat into some kind of informal (non-taxable) activity. We assume that any tax proceeds are distributed back to the citizens in lump-sum fashion with the transfer being denoted by $T$. Let $\mu_s$ be the mean income in the society in period $s \in \{1, 2\}$. Then the government budget constraint is:

$$T_s = t_s \mu_s.$$ 

The timing as follows. We begin with a period one policy maker in place. In period one, the incumbent sets $(t_1, T_1)$. Then the mobile citizens choose whether or not to invest. The type of the period two politician is then determined in a manner to be described below. This policy maker then sets $(t_2, T_2)$.

We will consider an explicitly democratic political process where the type of policy maker is determined according to majority rule. We assume that:

$$\gamma_M + \min \{\gamma_L, \gamma_H\} > \frac{1}{2} > \max \{\gamma_H, \gamma_L\}.$$ 

This says that neither the high or low income citizens are a majority. However, each becomes a majority if they combine with the mobile citizens.

As a benchmark, suppose that there is no government so that $t_s = T_s = 0$ for $s \in \{1, 2\}$. In this case, all the mobile citizens find it optimal to invest.

We now study what happens with voting. As above let $\{\sigma_1, \sigma_2\}$ be the type of citizen in office in each period. First consider period two policy making. It is easy to
see that the optimal tax rate is:

\[ t_2 (\sigma_2, \mu_2) = \arg \max_{t \in [0, \bar{t}]} \{ (1 - t) y_{\sigma_2} + t \mu_2 \}. \]

This implies that:

\[ t_2 (\sigma_2, \mu_2) = \begin{cases} \bar{t} & \text{if } y_{\sigma_2} < \mu_2 \\ 0 & \text{otherwise.} \end{cases} \]

A citizen with income \( y_H \) always prefers to set \( t_2 = 0 \) while a citizen of type \( y_L \) sets \( t_2 = \bar{t} \).

It is now straightforward to see how policy in period two depends on period one investment decisions by mobile citizens. If all mobile citizens invest, then citizens with income of \( y_H \) are in the majority and hence \( t_2 = 0 \) while if all the mobile citizens choose not invest, then citizens with income level \( y_L \) are in a majority and the tax rate is \( t_2 = \bar{t} \). This becomes interesting if:

\[ (1 - \bar{t}) (y_H - y_L) < \kappa, \]

since investment by mobile citizens then occurs only if these citizens anticipate that the other mobile citizens will also choose to invest. In other words the political system creates a strategic complementarity between the mobile citizens.

The political model now has multiple equilibria. In one equilibrium no mobile citizen invests so that the majority of citizens earn low income in period two and \( t_2 = \bar{t} \). In the other, all mobile citizens invest in period one and there is no redistributive taxation in period 2 since a majority of citizens have high incomes, i.e., \( t_2 = 0 \). National income is different in each case with the income level being higher when taxes are low.

To complete the model, we need only to determine the first period policy choice. However, this does not have any baring on whether the mobile citizens choose to invest.\(^{29}\) Hence:

\[ t_1 (\sigma_1, \mu_1) = \arg \max_{t \in [0, \bar{t}]} \{ (1 - t) y_{\sigma_1} + t \mu_1 \}. \]

It is the fact that period one policy does not affect private investment that differentiates this example from those above.

We now consider whether this example constitutes a government failure and in what sense. Since government policy in this case is purely redistributive, any \( t_s > 0 \) constitutes a government failure in Wicksell’s sense. This could be achieved either by passing a constitutional restriction that \( t_s = 0 \) or providing some form of guarantee that a high income citizen will always be in office. Social surplus will also be maximized in this example where redistributive taxation is zero in period two.

Turning now to Pareto efficiency. We will show that the outcome described may constitute a political failure. Consider the outcome where \( t_2 = \bar{t} \) so that mobile citizens

\(^{29}\)This assumes that either \( \kappa \) is a purely non-financial (i.e. utility) cost or else that \( \kappa < y_L \). Otherwise government redistributive policy could lead to an increase in investment.
2.7. IMPLICATIONS

do not invest. Suppose that the tax rate is cut to:

\[ \hat{t} = \frac{(y_H - y_L) - \kappa}{(y_H - y_L)}. \]

This is the highest tax rate at which mobile citizens are willing to invest (they are indifferent between investing and not investing). We need to show that the low income citizens are better off at \( \hat{t} \) than at \( \tilde{t} \). This is true if:

\[ \hat{t} \frac{\gamma_H}{(1 - \gamma_L)} < (y_H - y_L) - \kappa. \]

This requires that the gain from investing be large enough and that be sufficiently many mobile citizens. This implies that transfers to the low income groups are small when the mobile do not invest and increase substantially when they do. This shows that the example that we are studying here is a political failure.

This example is an instance of a general phenomenon whereby private decision making affects policy outcomes. This general class of policy distortions includes the example studied by Coate and Morris (1999) who emphasize the fact that policies can persist due to individuals making private investment decisions to benefit from them. The endogeneity of groups affecting political support is also at the heart of the mechanism in Acemoglu and Robinson (2001) sustaining inefficient transfer policies.\textsuperscript{30}

2.7 Implications

It seems harsh to call something a failure unless it can be shown that there is something better. The theory of market failure provides a framework for shaping the case for government to improve what the market can achieve. Indeed, many of the textbook functions of government – to provide public goods, regulate externalities and regulate monopolies are rooted in the theory of market failure.

How to react to government failure is less clear-cut. One widespread view is that identifying government failures should primarily provide a basis for constitution design. This view is often identified with the public choice approach, especially the work on Buchanan.\textsuperscript{31} The idea is that the rules of the political game are codified ex ante, possibly behind a veil of ignorance, and that political resource allocation is shaped by these rules. The rules take on normative significance in responding to the kinds of difficulties that might arise in the operation of politics.

\textsuperscript{30}Krusell and Rios-Rull (1996) is also related although in their model current policy (rather than just anticipated future policy) influences investment and hence future policy. Thus, it also has elements of the policy linkages studied in the previous section.

\textsuperscript{31}See, for example, Frey (1983) and Mueller (1996) for development of these ideas.
Buchanan (1967) divides constitutions into two parts. One of them is a set of rules laid down for the conduct of the policy process. He refers to this as a *procedural constitution*. This might include designation of specific policy authority, separation of powers, the electoral system and rules government who may vote or hold political office. The other part of the constitution may refer to policies directly. Buchanan calls this a *fiscal constitution*. However, it is clear that the slightly broader term *policy constitution* might be more apt.

There are opportunities, particularly when new nations are founded, to specify the political architecture from the ground up. However, more commonly the focus is on more modest, piecemeal reforms. There are three main categories of democratic constitutional reform that are debated: Democratic Structure, Government Architecture and Policy Rules. We discuss the main elements of each in brief.

1. Democratic Structure: There are many facets of a constitution which shape the conduct of democratic politics. These include

   (a) Voting Rules: First, there are rules about who is eligible to vote and how they can register. The nineteenth and twentieth centuries saw a significant extension of the franchise, notably with women securing the right to vote. Second, there are the rules for aggregating votes and determining the spatial pattern of representation. There are countless alternatives which vary between a straight majoritarian system with single-member districts through to list systems with proportional representation.

   (b) Electoral Conduct: Along with opening of the franchise came a decline in rules restricting who could hold public office. That said, most democratic systems still limit access to political office more than they limit the right to vote. Also important are rules about the conduct of campaigns which affect how elections are fought.

   (c) Legislative Structure: There are many aspects of rules within a legislature that can affect how politics is conducted post-election. This category also includes the possibility of directly elected authority as in the case of a president. Also important is the choice between unicameral and bi-cameral systems.

   (d) Direct democracy: How and whether there should remain a direct voice for citizens remains open to debate. Some countries, notably Switzerland, allow a significant role for citizens’ initiatives while most others rely very little on this.

2. Government Architecture: This mainly includes the rules that determine who has policy authority and on what basis.
2.7. IMPLICATIONS

(a) Independent agencies: Even in democracies, there are many agencies that have direct policy authority. These include independent central banks to determine monetary policy which operate alongside democratic legislatures. The role of the judiciary is also important in many polities. In common law countries, in particular, the judiciary plays a key role in shaping many policies through establishing precedent. The judiciary in the form of constitutional courts may also have the authority to limit the power of elected representatives.

(b) Decentralization: Most countries have multi-tiered governments and the degree of policy authority delegated below the central government level is an important policy parameter. This has been an extremely active policy area of late and includes some non-democratic polities such as China.\(^{32}\)

(c) Structure of the Executive: The way in which government agencies are structured within government and how this is accountable to legislative authorities varies greatly across democratic systems. Parliamentary systems typically have ministers who are members of the legislature whereas some systems have appointed cabinet members.

3. Policy rules: There are also many legal rules that have a direct impact on policy outcomes. It is unclear, in some cases, how far these should be viewed as constitutional provisions of merely policies that are subject to change. Nonetheless, they clearly have great practical force. They include:

(a) Fiscal deficit regulations: Many countries use means to limit the ability of governments to run deficits. This is interesting in view of the large literature – which we referred to above – which looks at why deficit finance can be subject to political failure. We discuss this further in chapter four.

(b) Private property: Many countries have evolved systems for the protection of property which place limits on the power of government to tax and regulate its citizens.

(c) Civil liberties: Policy rules also have to function in the context of evolved rules for the protection of individual freedom. Policy choices that conflict with human rights provisions often require amendment to prevent inconsistencies between policies. Freedom of information provisions may also have a direct impact on the kinds of policies that can be implemented if their provisions are deemed to transcend those of particular policy provisions. These are areas where the role of courts is crucial.

\(^{32}\)See, for example, Qian and Roland (1998).
This book will not be able to do justice to this rich array of constitutional issues. However, the welfare economic framework that we have developed here around the notion of government failure should be helpful for thinking about some of these issues. A general intellectual approach suggests itself.

- **Step 1:** Develop a theoretical and empirical model of the effect of a particular constitutional rule.\(^{33}\)
- **Step 2:** Use Step 1 to identify who gains and who loses.
- **Step 3:** Use Step 2 to inform a process of discussion/aggregation to make a social decisions.

The ideas of government failure that we have discussed are implicit in the step from 2 to 3 where our different ideas of government failure corresponded to ways of weighting gainers and losers.

This procedure is, of course, a highly idealized process and real debates about constitutional reform are messy and may appear to be influenced by idiosyncratic events rather than reasoned debate. Just as special interests may influence policy so they frequently try to influence constitutional reform and the idea of founding fathers reasoning behind a veil ignorance is far fetched. But this makes the influence of social scientists even more crucial. It is imperative to have a space for structured and scientific reasoning, however slight that impact may sometimes appear to be.

But modesty too is needed. Economists have a comparative advantage in discussing many policies – especially those that involve the commitment of public resources. However, our competence in the field of human rights and calculating the value of freedom is more limited.\(^{34}\) The Wicksellian approach that we discussed above is the only approach to government failure which gives a direct role for such concerns. But in discussing policies such as abortion rights, the decision to go to war or the incarceration of terror suspects, we should acknowledge that our welfare economic framework is limited. But that still leaves a vast domain in which the ideas discussed here are relevant and one that may be at the heart of many constitutional reform discussions.

The subsequent chapters will discuss some of the constitutional choice issues raised above. However, this will mainly be done through the lens of a specific approach to politics – political agency models. This is a class of dynamic models which take imperfect information issues in politics seriously. They also provide a vehicle for weighing the selection and incentives issues emphasized in the last chapter. We will use these models to explore issues concerning the quality of economic policy choices and politicians.

\(^{33}\)See Besley and Case (2003), and Persson and Tabellini (2003) for progress in this area. Acemoglu (2005) discusses more broadly the idea of political equilibrium inducing preferences over institutions.

\(^{34}\)See Cooter (2000) for an attempt to tackle issues of rights in constitutional choices more directly.
Reflecting on institution design in this way, makes it clear that the reasoning of this chapter could also apply to institutions more broadly, i.e. to have a theory of institutional failure. This project is anticipated in Buchanan (1967) when he says:

"Theoretical welfare economics enables us to define the necessary marginal conditions that must be satisfied for an allocation of economic resources to be efficient. Straightforward extension of this analysis to "theoretical institutional economics" should enable us to define a similar set of conditions that would have to be met if an institutional arrangement or rule is to be classified as "efficient." It now seems quite possible that future developments will in fact allow for general statements of such conditions."


Nearly forty years on, the future developments that Buchanan refers are not yet a reality. However, much progress is being made.35

2.8 Concluding Comments

The main purpose of this chapter has been to identify sources of government failure. We have introduced three criteria by which such failures can be defined. We have also shown that they are distinct. Defining government failure in terms of Pareto efficiency – to parallel the classic definition of market failure – has little bite in static settings with unitary political actors. It can, however, be important when multiple actors are engaged in policy making. It is also non-trivial in dynamic models of government resource allocation.

Bator (1958) argued that non-appropriability, non-convexity and public goods are the main sources of market failure defined as a Pareto efficiency. The parallel list suggested by the discussion here is non-Coasian legislative institutions, poor selection of policy makers, costly rent seeking, and intertemporal investment, policy and political linkages.

This discussion of government failure serves as a useful background to the specific discussion that follows, but also to more general debates about the achievement of government. In criticizing government intervention it pays to be specific about the sense in which government is failing and the remedy that is needed. To do this in a rich economic policy model is by no means straightforward. But the desiderata as laid out above are useful as an intellectual structure. Economists have benefited enormously from having a rigorous notion of market failure and there is every reason to think that putting government failure on a similarly firm intellectual footing should also pay off.

35See Acemoglu (2005) for further discussion.
Figure 1: Pareto Inefficiency and Distributional Failures
Utility Possibility Frontier

Figure 2: Wicksellian Failures and Pareto Inefficiency
Chapter 3
Political Agency and Accountability

“(I)n the large the electorate behaves about as rationally and responsibly as we should expect, given the clarity and the alternatives presented to it and the character of the information available to it.” Key (1966, page 7).

“The problem is not how 180 million Aristotles can run a democracy, but how we can organize a political community of 180 million ordinary people so that it remains sensitive to their needs.” Schattschneider (1960, page 138).

3.1 Introduction

In the last chapter, we painted a broad picture of the issues at stake in thinking about whether government is likely to be run in the public interest. In this chapter, we explore a specific class of models which are ideally suited to pushing this debate forward – political agency models. Specifically, we look at the structure of political agency models with a focus on applications. We also lay out a canonical approach which should give the reader who is unfamiliar with the approach some idea of the model’s potential. The simplicity of the model also provides a building block to look at many of the issues that have been studied in the literature.

At the heart of political agency models is the principal-agent relationship between citizens and government – the principals are the citizens/voters while the agents are the politicians or bureaucrats. This flips around the standard principal agent model that has been widely used in public economics and is at the core of the mechanism design problem.¹ In that approach, a welfare maximizing government tries to discern the privately observed characteristics of citizens which it will use to make good social decisions. For example, the standard optimal income tax problem assumes that each individual privately knows his own ability to earn. This implies that individualized

¹See Atkinson and Stiglitz (1980).
redistributive lump-sum taxation may not be incentive compatible. Similarly, efficient provision of a public good requires that a government has sufficient information about the values that citizens attach to the good to discern the optimal level at which the good should be provided. The problem then lies in designing an optimal mechanism for public good provision.\(^2\)

In the political agency approach, the incentive problem arises because the citizens have delegated authority to policy makers who enjoy an informational advantage. There are two main problems:

- **Monitoring**: the policy maker may act opportunistically. There is a need to establish whether this has happened and to reward/punish behavior accordingly in order to minimize opportunism.

- **Selection**: there is a need to select the most competent policy makers and/or those whose motivations are most likely in tune with the public interest.

As we discussed in chapter one, elections are the core mechanism for solving these problems. For the purposes of this book, this will simply be assumed. However, clearly this begs a host of important questions of constitution design. A full-blown attack on the issues would begin with the primitives of information constraints and limits on contracting and would seek to design the optimal set of constitutional rules.\(^3\) It is less than clear whether the weight attached to the importance of elections in modern representative democracies would emerge from this approach. Moreover many of the assumptions of the political agency approach limit the incentive contracts that can be offered in apparently arbitrary ways. Clearly, in terms of theoretical purity, an approach rooted in mechanism design is on more solid ground. However, it is only as good as the available theoretical framework for capturing constraints on the set of mechanisms that are permissible. While this approach is strong when it comes to information constraints, it is hard to know how to capture other ideas like simplicity and fairness which may also be important in shaping institutional choices.\(^4\)

Our main aim here is to analyze the properties of a stylized model of representative democracy with imperfect information. We are not asking whether it is the best possible institution. It could even be Pareto inefficient in some contexts. However, by studying a fairly simple model some insights might be gained about what goes into making democracy work and the kinds of piece-meal reforms that can make things work better.

\(^2\)See, for example, Laffont and Maskin (1980).

\(^3\)See Laffont (2000) for an effort in this direction.

\(^4\)This may explain why, even after two decades of research, the mechanism design approach has yet to give a persuasive account of the attractiveness of markets for allocating resources except in the most stylized of situations.
3.1. INTRODUCTION

The problem of monitoring (moral hazard) is at the heart of the classical statement of principal agent problems (see, for example, Holmstrom (1979)). In this problem, an agent is completing a task on behalf of a principal which yields a publicly observable payoff. The problem is that the action which generates the payoff cannot be observed by the principal. The principal’s problem is then to design an incentive scheme for the agent which rewards him/her according to the payoff achieved. This model has been developed in a wide variety of ways to consider different problems.\(^5\)

In the standard problem, agents are homogeneous. However, there are extensions to allow for unobserved types (in addition to unobserved actions) in particular applications (see, for example, Laffont and Tirole (1986)). This can mean offering the agent a menu of contracts which achieves self-selection. The literature on the dynamics of regulation has emphasized the importance of “ratchet effects” in such settings in which information revealed by an agent is used in future periods, making information revelation harder to achieve in earlier periods.\(^6\) Such models assume that firms differ in some core level of competence that is fixed over time. In a political agency setting with re-election incentives and unobserved types, such dynamic issues are central – incompetent incumbents may aim to disguise their type to get re-elected. In modeling political agency, issues beyond competency may be important. For example, the issue of selecting an agent with an appropriate motivation may be an issue.\(^7\)

In this chapter, we study a variety of agency models tailored to look at concrete issues. Each author has tended to take quite a specific approach depending on their ends. Even then, the development of political agency models owes much to the more general understanding of agency problems that has developed within economics over the past thirty years.

As we discussed above, achieving accountability is one of the key roles of elections. In the loosest sense, this means tying the performance of politicians to what they do while they are in office. It is important to distinguish in this regard between formal (de jure) and real (de facto) accountability.\(^8\) It is in the nature of the institution of representative democracy that politicians are periodically subject to re-election. The formal rules – the term length and the rules governing electoral conduct – constitute the formal accountability rules. But they do not typically attempt to make any direct link between politician’s performance and his/her re-election chances. This depends on what happens in equilibrium, i.e. as a function of politicians’ behavior and voters’ strategies. The latter determines whether there is real accountability. In a poorly functioning

\(^5\)See Prendergast (1999) for a survey.

\(^6\)See, for example, Laffont and Tirole (1993) for discussion.

\(^7\)See Besley and Ghatak (2005) for a model of bureaucracy along these lines.

\(^8\)A useful definition of accountability in the political science literature can be found in Fearon (1999) who suggested that an agent A is accountable to principal B, if (i) there is an understanding that A is obliged to act in some way on behalf of B and (ii) B is empowered by some formal institution or perhaps informal rules to sanction or reward A for her activities or performance in this capacity.
democracy, politicians can perform acts which systematically displease voters without facing sanctions. This might be because voters are poorly informed or politicians intimidate voters.

The extent of real accountability might be thought of as a pretty good index of a well-functioning democratic system of government. It is through being accountable that politics is most likely to be responsive to voters’ wishes. Responsiveness and accountability are frequently held up as self-evident values in political systems. However, there is no necessary link between accountability and the welfare of society. This is likely to be true when the issues at stake are mostly common values (valence issues in the language of political science), i.e. the electorate broadly agree on what should be done. As we discussed in chapter 1, obvious examples include reducing corruption, improving public service efficiency and collecting taxes at the lowest possible cost.

If there is a large measure of disagreement on the right policy, there is no obvious reason to believe that making sure that elections outcomes are responsive to incumbent behavior increases any reasonable measure of social welfare. If the incumbent is accountable to the median voter then policies will be skewed towards these preferences even if (as argued in section 5.1 of chapter 2) these have no normative significance. To reflect this, we will work mainly here with models that assume that voters do have a common interest in achieving some outcome and discuss whether we would expect the political system to deliver it. We will return to this issue in our concluding remarks.

Our whole framework of analysis presupposes the existence of a legal system and a body of law governing the conduct of elections. For example, incumbents are supposed to leave office if they lose the election and to court voters fairly, i.e. avoiding bribery and intimidation. Similarly, there are legal limits on what politicians may do to serve their own interests and many politicians are threatened with legal sanction for using their offices corruptly. Many of the failures of democratic systems are legal. For example, Robert Mugabe’s unwillingness to leave his office in Zimbabwe makes real accountability impossible. We focus here on accountability within a well-functioning legal system. However, it is clear that the extent to which the formal institutions are upheld might be an important difference in the way democracies function.

The remainder of the chapter is organized as follows. In the next section, we discuss political agency models in general – the assumptions that go into them and general modeling issues. Section three puts forward a canonical agency model which serves to illustrate a number of the issues. Within that section (section 3.3), we discuss a specific empirical application to data on U.S. Governors. This will offer an opportunity to get into some more general empirical concerns. Section four explores a number of extensions of the basic model. In section five, we discuss concrete policy driven applications from the literature. Section six concludes.

---

9 Though in practice prosecutions for serving politicians are rare.
3.2 Elements of Political Agency Models

Common to all of the agency models discussed here is the importance of re-election incentives in shaping the behavior of incumbent politicians. To model this requires a dynamic structure—the simplest approach being a two-period setting. However, some contributions have looked at longer time horizons and we will do so in sections 4.6 and 4.7.

The agency model can be thought of as a game in which the players are voters and politicians. At any point in time, one or more politicians have the power to make policy. If they are re-elected, they retain this right. If they lose it, then another politician replaces them. The politicians could, in principle, be drawn from the pool of voters. However, typically, this process of selection is not modeled. The simplest (and most popular setting) is one where voters are homogeneous while politicians differ in their type.

The key modeling issues are:

- The nature of the uncertainty.
- The motives for holding office.
- The formal structure of accountability.
- The nature of voting.

We will look at each of these in turn.

3.2.1 The Nature of the Uncertainty

Models vary in what they specify as being unobservable to voters. At the core of many models are unobservable types of politicians. These could be their competence levels. For example some politicians may be skilled in choosing policies whereas others make poor choices. There could also be uncertainty about the motivations of politicians. Some may be honest while others are not. Some may have a strong sense of duty while others do not. Politicians may also vary in the utility they gain from being in office—due to the thrill of being powerful or whether they like or dislike media attention.

One key issue is whether the politicians know their own types. In the case of competence, it seems less plausible to suppose that politicians know their own capacities completely and may be learning about this along with the voters. A model of uninformed politicians is more in tune with the celebrated Holmstrom (1999) “career concerns” model in which individuals put in effort which reveals their type to the “market”. Persson and Tabellini (2000) develop a political agency model along these lines, i.e., in which voters and politicians are symmetrically informed.
Voters may also be poorly informed about the best policy. There are good reasons to think that governments have access to a wide variety of policy advice which should enable them to make better policy decisions than voters left to their own devices. Indeed, the notion of asymmetric information is at the centre of the notion that voters are rationally ignorant as suggested in Downs (1957). In reality there are ways of informing voters about policy such as think tanks, pressure groups and the media.

A further source of uncertainty is over the exact policy implemented. An important example is the quality of policy choices which may be uncertain for a considerable time after a policy is chosen. A good example is the recent war with Iraq – it will be some time before the wisdom of the decision and the competence of its execution is known. In some cases, the welfare consequences of a policy may be uncertain for many years into the future such as in the case of policies towards global warming.

In cases where politicians have a private interest in adopting some policy or other, voters may also find it hard to discern whether they have earned a rent from any particular policy. There is also a large literature which says that lack of transparency makes it difficult for voters to know exactly what is the true state of public finances. This may make it attractive for politicians to raise debt in order to postpone paying for certain policies in the short-run.

3.2.2 The Motives for Holding Office

It seems to be accepted that a cast iron law of politics is that politicians like to be re-elected. However, the sources of re-election incentives differ between models. One possibility is simply to posit an “ego rent” from holding office as in Rogoff (1990). This rent could be the intoxication effect from power or some sense of pride at having been approved in a ballot by one’s fellow citizens.

Another source of benefits from holding onto office may come in the form of material gain. This could be because politics provides an opportunity to reward cronies or an opportunity for corruption. It could even be because politicians’ wages and perks are attractive.

The desire for re-election may also stem from public goods concerns – this would be true, for example, if incumbents have strong policy preferences. We would expect these forces to be most important when there is a large degree of polarization in the political system. In such cases, the incumbent will be keen not to have another policy maker take over. Related to this, Maskin and Tirole (2004) discuss the possibility of a “legacy effect” associated with holding office whereby politicians care about the value created while they are in office.

In all three cases, incumbent politicians will desire re-election either instrumentally or for its own sake.
3.2.3 The Nature of Accountability

The agency model works best when applied to individual directly elected representatives such as mayors, presidents and governors. In this case, the incumbent is an individual who has certain responsibilities and spheres of discretion. The basis of accountability is then defined relative to these responsibilities.

In some contexts, this model has only limited applicability. For example, the individuals being held to account may be part of a collective such as legislature or a committee. In this case, there is the problem of untangling responsibilities for outcomes. In general, this may be expected to weaken the mapping from outcomes to re-election decisions on individuals.

In almost all democratic settings, parties play an important role in the workings of political competition. How they affect the process of political accountability is not altogether clear. Parties may have longer time horizons and may therefore make individuals who hold power take a longer view, leading to less opportunistic behavior by incumbents. Obviously the degree of attachment of the incumbents to parties and the kinds of sanctions that parties have are crucial here.

When thinking in terms of parties, it is less clear how to interpret the politicians’ types. For example, it could reflect some underlying collective reputation or competence of the party members. Voters may then cast their votes based on party rather than individual characteristics. This view is intriguing. However, modeling it clearly requires some kind of model of the internal workings of the party.

3.2.4 Retrospective Voting

A key idea in the models is that voters hold incumbents to account for their actions while in office. This constitutes an optimistic view of the political process in which sufficient voters are informed about policy outcomes and use this information when deciding for whom they will vote. This is still an issue which is hotly debated by political scientists. But influential contributions by Key (1966) and Fiorina (1981) have significantly shifted the presumption towards the view that voters are rational and make their assessments based on the records of candidates while in office.

Fiorina (1981) identifies three distinct theories of retrospective voting: The first, which he identifies with Downs (1957), is that voters use retrospective information because such information is to hand and is cheaper than scrutinizing the records of candidates. The second he associates with Key (1966) is that voters are result orientated. Fiorina’s view attaches more weight to party identities in shaping retrospective voting decisions.

Both Fiorina (1981) and Key (1966) argue that there is evidence of retrospective voting in the context of U.S. presidential elections. Their work has also been a key force behind accepting a model of political behavior based on standard postulates of
rational choice applied to voters. That said, they did not have a framework of the kind that political agency models provide which permits an analysis of the equilibrium implications of retrospective voting. This requires some notion of why past records provide information about future performance. This is captured in the political agency model by assuming that voters are learning from past actions and use Bayes rule to update their beliefs. One key implication of this approach is that there really is no meaningful distinction between prospective and retrospective voting. It is precisely because there is information content in past actions about future behavior that retrospective voting is rational.

3.2.5 Model Types

The models break down into three basic types. The first generation political agency models focused on hidden actions – specifically shirking by incumbents. Barro (1973) and Ferejohn (1986) both focused predominantly on moral hazard problems in government. In this framework, all politicians are alike in their desire to use political office as a means for pursuing their own agendas. The question is how well a re-election mechanism based on retrospective voting can discipline incumbents. The basic idea is that voters will choose a threshold that gives the politician an incentive for restraint. The latter always faces a choice between extracting maximal rents in the current period or showing restraint and continuing to enjoy rents in future. The main modeling assumptions concern the extent to which voters can observe the actions of incumbents and whether there is an unobservable state of nature. Since such models avoid different kinds of politicians, voters are always indifferent between re-electing the incumbent and the challenger.\footnote{This is quite a severe restriction. As we shall see below, even a little bit of heterogeneity can have a big effect on the predictions.}

In a pure adverse selection model, the only issue is selecting the right kind of incumbent for the job. There is nothing that an incumbent can do to affect the electorate’s opinion of them. The key question is how the observations that voters can make allow them to detect the type of the incumbent. This is the core model of Besley and Prat (2004). In reality, these models have fairly limited appeal as most politicians are likely to have some ability to disguise or reveal their type via their actions.

The most interesting and challenging types of models come from combining hidden action and different types in the same model. Among the earliest efforts in this direction are Banks and Sundaram (1993) and Austen-Smith and Banks (1989).\footnote{See also Banks and Sundaram (1998).} They assume that politicians differ in their competence and that the actions that they take are not observed by voters. Politicians that differ in their motivations are central to the analyses of Besley and Case (1995a), Coate and Morris (1995), Fearon (1999) and Rogoff (1990). One of the key issues that arises in these models is the use of policy choices as a signaling
3.3. THE BASELINE MODEL

The baseline model as different types of politicians try to differentiate themselves from one another. Two types of distortions are possible. Good politicians may try to exaggerate their behavior in a way that cannot be mimicked by bad ones. Alternatively, bad politicians may try to pool with good ones to enhance their re-election chances. Such pooling behavior may or may not be optimal in political equilibrium.

Persson and Tabellini (2000) develop political agency models that are more in tune with the original career concerns model of Holmstrom (1999) in two senses. First, neither public servants nor the public know the competence level of their public servant. Second, the decisions made by incumbents are chosen prior to the election taking place. They take the form of an investment that will payoff after the election. This differs from the main thrust of the literature which focuses on retrospective voting. Models based on the classic Holmstrom career concerns model are also developed by Ashworth (2005) and Ashworth and Bueno de Mesquita (2005a,b). This kind of approach also gets away from some of the signaling issues as the politicians do not know their own types. For some applications (for example where a type is some kind of competence issue) this may be natural. However, when preferences are part of the type, this is a less plausible assumption. Obviously, what is reasonable will depend largely on the application in question.

For the most part agency models work with situations where re-election incentives are a force for good either by improving selection of politicians or by creating greater discipline. Essentially, this works by giving incentives for politicians to build reputations. However, there are circumstances in which such reputation building can be counter-productive. This happens when incumbents are deterred from doing the right thing because they know that voters will not reward them for doing so. In section 4.3, we discuss a specific example of this.

3.3 The Baseline Model

This section develops a very simple approach to political agency problems. It provides the simplest agency framework which puts adverse selection and moral hazard together. At the heart of the agency model is the notion that some politicians are more able/willing to give voters what they want. We begin by setting out a very simple model that looks at behavior of politicians in such a world. We then consider a variety of modeling choices and assumptions.

---

12 More precisely, the model in section 4.5.1 (and section 9.1) of Persson and Tabellini (2000) assumes that an incumbent chooses rents in period one and consumes these whether or not he is re-elected.

13 Political agency models have much in common with models where experts are hired to make decisions. They typically assume that payoffs cannot be made contingent on the agent’s performance and hence being re-appointed is the main incentive – see, for example, Prendergast and Stole (1996) and Ottaviani and Sorensen (2001).

14 The closest cousins to the model used here are Barganza (2000) and Smart and Sturm (2005).
3.3.1 The Environment

There are two time periods denoted by $t \in \{1, 2\}$. In each period, a politician is elected to make a single political decision, denoted by $e_t \in \{0, 1\}$. The payoff to voters and politicians depends on a state of the world $s_t \in \{0, 1\}$ which is only observed by the incumbent. Each state occurs with equal probability. Voters receive a payoff $\Delta$ if $e_t = s_t$ and zero otherwise. Voters and politicians discount the future with common discount factor $\beta < 1$.

There are two types of politicians – congruent and dissonant – the type is denoted by $i \in \{c, d\}$. Let $\pi$ be the probability that a randomly picked politician from the pool is congruent. All politicians (good or bad) get a payoff of $E$ from holding office. This can be thought of as comprising any pure ego rents plus any wages from holding office. It could also depend on office perks such as pensions and free housing. We normalize the politician’s outside option to zero.

Congruent politicians’ payoffs are $E + \Delta$ in every period because they always choose $e_t = s_t$, i.e. they share voters’ objectives exactly. Dissonant politicians get a private benefit (dissonance rent) of $r \in [0, R]$ from picking $e_t \neq s_t$, where $r$ is drawn independently from a distribution whose cdf is $G(r)$. The mean of $r$ is $\mu$ and we assume that $R > \beta (\mu + E)$.

Let the realized value of “dissonance rents” in period $t$ be denoted by $r_t$. The action of a politician at time $t$ is denoted by $e_t(s, i)$ with $s \in \{0, 1\}$ and $i \in \{c, d\}$.

The interpretation of dissonance could be quite broad. It could include politicians whose competence is limited and who find it costly, therefore, to do what voters want. It could also capture the case of a politician who gives into a special interest group which wants the opposite of what voters want. It could also include a politician with a private agenda of his own (say an ideological disposition) which he wishes to pursue. In the case of a legislator, it could also capture the idea of delivering “constituency service” to the voters. The exact interpretation is not important for most of the discussion.

Timing is as follows. Each period, nature determines the state of the world and the type of politician (if that politician is newly elected) neither of which is observable to voter. Nature draws $r_1$ from the distribution $G(r)$. The incumbent politician then picks his preferred action. Voters observe their payoff and then decide whether or not to re-elect the incumbent or select a challenger at random from the pool of potential politicians. After the election, a dissonant politician receives a fresh draw from $G(r)$, denoted by $r_2$. The period two action then follows after which period two payoffs are realized. Thereafter, the game ends.

---

15This assumption guarantees that dissonant politicians pick the wrong action from voters’ point of view some of the time.
3.3. **THE BASELINE MODEL**

3.3.2 Equilibrium

The equilibrium concept most commonly used in solving a model like this is perfect Bayesian equilibrium. This requires that in every period each type of politician behaves optimally given the re-election rule that the voters put in place. Voters use Bayes rule to update their beliefs.

It is very easy to work out the equilibrium behavior of politicians in this model. In period two, every kind of politician takes his short term optimal action. Thus $e_2(s,c) = s_2$ and $e_2(s,d) = (1 - s_2)$. Anything of interest in political behavior must, therefore, happen in period one.

To explore this, it is useful to introduce two key notations. First, let $\lambda$ be the probability that a dissonant politician does what voters want in period one. It is therefore an index of political discipline. Congruent politicians always do what voters want provided that they are re-elected for doing so.\(^{16}\) If a dissonant politician does the right thing for voters with probability $\lambda$ and voters use Bayes rule, their belief that the politician is congruent conditional on having received a payoff of $\Delta$ is:

$$\Pi = \frac{\pi}{\pi + (1 - \pi) \lambda} \geq \pi.$$  

Thus good behavior always improves a politician’s reputation (now measured by $\Pi$). This implies that there is always an equilibrium in which any politician who produces $\Delta$ for voters is re-elected, assuming voters are retrospective, i.e. use the incumbent’s performance during period one as their basis for voting. A politician who fails to produce $\Delta$ for voters is not re-elected since such an agent is dissonant for sure and will yield voters a zero payoff in period two.

Based on this, it is straightforward to determine the optimal period one action of dissonant politicians. By now the value of their dissonance rent $r_1$ has been revealed. This must be compared with the longer-term benefits of doing what voters want and being re-elected which is $\beta(\mu + E)$. Thus, the probability that a dissonant politician takes the action that voters want is:

$$\lambda = G(\beta(\mu + E)).$$

Under our assumptions on the distribution of the dissonance rent, $\lambda \in (0, 1)$. Thus we have:\(^{17}\)

**Proposition 1** Congruent politicians always set $e = s$. Dissonant politicians choose $e = (1 - s)$ in period two and will choose $e = s$ in period one if they earn sufficiently

\(^{16}\)Below, we explore a variant of the model in which this is not the case.

\(^{17}\)If $(1 - \beta(1 - \pi)) \Delta < \beta E$, there is also an equilibrium where all politicians pool on the action $e_1(s,i) = (1 - s_1)$ and voters believe that any politician who generates $\Delta$ is dissonant. However, this equilibrium is ruled out by applying Cho and Kreps (1987) intuitive criterion. Hence, we focus on this equilibrium throughout.
small rents from being dissonant. All politicians who choose \( e = s \) in period one are re-elected.

Thus the model makes clear how elections can motivate politicians – the non-congruent politicians are now willing to mimic the congruent ones some of the time. Moreover, voters hold politicians to account for not delivering a return. At the same time, the re-election mechanism is imperfect. Some of the time, dissonant politicians survive to the second period. This provides a starting point for thinking about how political accountability works.

### 3.3.3 Implications

We now consider some normative and positive implications of the models before proceeding to some discussion of empirical testing.

#### The Quality of Government

The model provides a useful way of thinking about determinants of the quality of government. In doing so we will abstract from the payoffs of politicians and measure quality of government in terms of whether voters get what they want.\(^{18}\) To this end, we consider ex ante voter welfare in both time periods. In period one, it is:

\[
V_1(\lambda) = (\pi + (1 - \pi) \lambda) \Delta
\]

while in period two, it is:

\[
V_2(\lambda) = \pi[1 + (1 - \pi) (1 - \lambda)] \Delta.
\]

It is clear that period one welfare is increasing in \( \lambda \) (incumbent discipline) while period two welfare is decreasing in \( \lambda \). The latter is due to the fact that fewer dissonant incumbents are caught when there is more period one discipline.

Discounted welfare, which is

\[
W(\lambda) = V_1(\lambda) + \beta V_2(\lambda),
\]

is increasing in \( \lambda \). This is because the loss in period two welfare is of the order \( \beta \pi (1 - \pi) \) while the gain in period one is of the order \( (1 - \pi) \). Voters gain from ill-discipline only

---

\(^{18}\)The time honored tradition of ignoring the payoffs of politicians means that we cannot assess whether the outcome is efficient. For that, we would have to consider whether a politician could be given a transfer by voters which makes everyone better off. But that, in turn, raises issues of whether we should consider different kinds of contracts for rewarding politicians. The tradition in political economy of ignoring the payoffs of politicians in the welfare calculus is strange and contrasts with analyses of efficiency in a market context which seem always to look at the payoffs of all market participants.
3.3. THE BASELINE MODEL

when a dissonant politician is replaced by a congruent one in period two whereas they
gain from improved discipline whenever there is a dissonant politician in office in period
one.

Voter welfare is also increasing in $\pi$ – the quality of the pool of politicians. This
is obvious as it increases voter payoffs in both periods and does not affect the quality
of dissonant politicians’ actions. This captures very simply the idea that a polity with
more civic virtue (a higher quality political class) is better for the voters.

The model also predicts that voter welfare and political turnover are negatively
related. To see this, observe that turnover is equal to $(1 - \pi) (1 - \lambda)$ which is decreasing
in $\lambda$ and $\pi$, while as we saw above, voter welfare is increasing in $\lambda$ and $\pi$. This is due
to the fact that a lack of political turnover is symptomatic of voter contentment – only
politicians who do what voters want get re-elected. This is not always true as we shall
see below.

The Incumbency Advantage

As discussed in Ashworth (2005), the agency model can be used to make sense of an
incumbency advantage in politics. To see this simply, suppose that there is a ran-
dom popularity shock $\delta$ to the incumbent which is uniformly distributed on the interval
$[-\frac{1}{2}, \frac{1}{2}]$ and determines the election outcome along with the reputation of the incum-
bent.\footnote{This is equivalent to the challenger and the incumbent experiencing a shock with the \textit{difference} in
the shocks being represented by $\delta$.} Then, the probability that the incumbent is re-elected is:

$$\max \left\{ \frac{1}{2} + \xi \pi (1 - \pi) (1 - \lambda), 1 \right\} > \frac{1}{2}. $$

This incumbency advantage is natural because of the asymmetry in opportunities be-
tween challengers and incumbents – the challenger has no way of signaling his congruence
to the voters.

The Term Limit Effect

Political agency models predict that the behavior of politicians depends on how far their
political careers extend into the future. This is the reason that, in this two period
setting, voter welfare and politicians’ behavior differs across the two time periods even
though the policy environment is the same. Comparing politicians who can and cannot
run for re-election, the model predicts a term limit effect – politicians behave differently
when they can and cannot run for re-election. Specifically, discipline from a dissonant
politician is $\lambda$ in period one and zero in period two. While this is quite stark in the
current setting, the existence of a last period effect is quite general in agency models.
CHAPTER 3. POLITICAL AGENCY AND ACCOUNTABILITY

However, even though period two dissonant politicians’ behavior deteriorates, it is still a good idea to re-elect politicians on the basis of their first period behavior. Thus, expected performance among the pre-selected group of surviving period two politicians is higher than a randomly selected politician.

This observation from agency models motivates empirical tests of last-period effects in political life. For example, one strand of the political science literature focuses on data on representatives in the U.S. Congress. It aims to test whether representatives who are not going to run behave differently from those who plan to. Lott and Bronars (1993) analyze Congressional voting data from 1975-90, and find no significant change in voting patterns in a representative’s last term in office. However, it is far from clear that Congressional representatives who announce they are stepping aside provide an adequate picture of the behavior of state governors who are bound by law not to run again for reelection. A provocative paper on potential end-games in the U.S. Congress is provided by McArthur and Marks (1988), who observe Congressional behavior in a lame duck session of Congress: in post-election sessions, members who have not been reelected are at times called upon to vote on legislation before the swearing in of the new Congress. McArthur and Marks observe that lame duck representatives were significantly more likely in 1982 to vote against automobile domestic content legislation than were members who were returning.

The U.S. house of representatives does not (yet) have term limits. Hence, whether these kinds of studies can pick up exogenous variation in the time horizon of politicians is moot. In the next section, we look at the evidence from U.S. Governors. Here, the term limits imposed are genuine and there is the possibility of comparing the behavior of a given politician in their term-limited and non-term-limited period in office. While the model predicts that there is a selection bias into second terms, this provides a much more convincing basis for testing whether the time-horizon matters.

Empirical Evidence from U.S. Governors

Models of political agency have had many applications. But one of the most compelling is to U.S. governors. This is for two main reasons. First, U.S. governors correspond fairly well to the standard agency model – a single agent being held accountable for their actions with well-defined election dates and rules. Second, there is plenty of data available for U.S. governors making this context a rich testing ground for the models. The contrast with U.S. Presidents who may also be important examples of political agents here is clear – there is roughly fifty times more data available at the Gubernatorial level. Our look at the data here is sketchy. However, it will still serve as a useful means of breathing life into the empirical model.

On the downside, the case of governors is made less clear-cut since they must work together with state legislatures to determine policy outcomes. Hence, on some specific policy issues, they may have only limited input into the policy process. However, this
is hard to know a priori. The institutional rules affecting the powers of Governors also varies across states. For example, only a sub-set of Governors have a line-item veto which allows them to strike out specific budgetary provisions. Governors also receive very different levels of remuneration and they may also face different degrees of media scrutiny. While in principle, these many aspects of state-to-state heterogeneity could be brought into the empirical study, we will stick here to a straightforward empirical analysis. The role of this sub-section is to illustrate the potential of the political agency framework as a way of explaining patterns in the data.

The period for our data is 1950-2000. We will study two things – (i) the determination of re-election chances of Governors who are eligible to stand again, and (ii) the effect that not being able to stand (being term-limited) has on policy choices. In the first case, we will look for evidence that Governors are being held to account by voters for their actions while in office. In the second, we will look first at whether governors who are unable to run appear to behave differently from those who cannot.

Accountability The idea that voters respond to changes in the economy is widely documented in a variety of empirical contexts – see, for example, Nannestad and Paldam (1994) for a review of the literature. For the U.S. states, this means linking voting outcomes to general indicators of state-level economic health. Chubb (1988) considers the determinants of state elections as a function of the performance of the state economy and other factors using data from 1940-82. He finds very little evidence that changes in state income levels affect election outcomes. These results are consistent with the findings of Adams and Kenny (1986) and Peltzman (1987) who both show that whether a government is re-elected is not correlated with economic growth in the state in question.20

We begin by looking at whether voters hold incumbents to account as follows. Let $r_{sgt}$ be an indicator variable that denotes whether the incumbent $g$ is re-elected in state $s$ at time $t$. Suppose that this can be modeled as a linear probability model:

$$r_{sgt} = \alpha_s + \beta_t + \rho x_{st} + \gamma z_{gt} + \theta \Delta_{st} + \varepsilon_{st}$$

where $x_{st}$ is a vector of state level controls in part proxying for differences in the electorate, $z_{gt}$ are personal characteristics of the governor, and $\Delta_{st}$ are policy (and other economic) changes during the term that the incumbent has been in office. The model also includes state fixed effects $\alpha_s$ and year effects $\beta_t$.

We will control for state population and the incumbent’s lagged vote share in $x_{st}$. Our incumbent specific characteristics are age, years of education, experience in both political and non-political careers and the proportion of that experience that has been in politics. The means for these variables are in Table 3.1. For the vector $\Delta_{st}$ we include the growth in state taxes per capita in the past two years, the growth in state spending.

\[\text{20Lowry, Alt and Ferree (1998) find that state income growth relative to the national average affects the incumbent’s vote share in Gubernatorial elections.}\]
per capita in the past two years and the change in (log) income per capita in the same window.

Table 3.1 about here

The results are in Table 3.2. We present two sets of results which vary slightly in the sample of governors that we look at. In the first instance (columns 1 and 3 of Table 3.2), we look at all governors who are eligible to stand for re-election regardless of whether they did. In effect, this is treating a decision not to run as a defeat. The second sample considers only those governors who actually ran again for office (columns 2 and 4). In columns (1) and (2), we do not control for the governor’s personal characteristics – these are introduced in columns (3) and (4). One finding from the data is that governors who put up taxes are significantly less likely to win in all specifications. There is also evidence that the economic growth rate within the state is related to the election chances of the governor (contrary to earlier results by Adams and Kenny (1986) and Peltzman (1987)). By contrast, spending growth is not correlated with re-election chances.

Table 3.2 about here

Across the board, having an older governor is associated with a lower re-election probability. Years of experience before governorship is positively correlated with election success and this is particular true if this is years of political experience. The vote share at the last election is positively correlated with being re-elected for governors who choose to stand again.

In Table 3.3, we look at the votes that the Governor received if he was re-elected denoted by $v_{gst}$ as a function of the same set of right hand side variables. Specifically, we run:

$$v_{gst} = \rho x_{st} + \gamma z_{gt} + \theta \Delta_{st} + \lambda v_{gst-1} + \eta_{st}.$$  

These results parallel those from the previous Table – we find that the governor has fewer voter if he puts up taxes. Put together, this evidence confirms the relevance of retrospective voting based on policy and economic concerns for U.S. governors.

The Term-Limit Effect  There are a number of previous contributions that have looked at term limit effects. One way to examine term limits is cross-sectionally, looking at the permanent differences in policies in those states that have such limits. This is the approach in early work on the topic by Crain and Tollison (1977, 1993), and Crain and Oakley (1995). Crain and Tollison (1977) make the interesting and important point that if political office is a productive asset, one used to produce political outcomes, then candidates for the office should be willing to pay more for the opportunity to serve in states with longer terms, and in states without term limits. They find, using cross-sectional data for races run in 1970, that challengers spend less money when running
for two two-year terms than do those running for one four-year term. In addition, challenges spend less in states with term limits than in states without term limits. Crain and Oakley (1995) examine whether states that allow governors to succeed themselves indefinitely have different public capital stocks and flows than do states where governors are restricted by some sort of term limit. They find, using data from the 1980s, and controlling for a number of state institutions, that the stock of state government capital per capita, the change in the stock, and the percentage change in the stock, are all lower in states with term limits. Bails and Tieslau (2000) argue that term limits should lower the rate of growth of spending, by making policy makers more responsive to citizens’ preferences (the premise being that voters want smaller government). They test for this using a random effects model for the period 1969-94, and confirm a negative coefficient on state expenditures. All of these results raise the usual issue of whether such limits are merely proxying for omitted state level characteristics. State income per capita and state population are significantly lower in states with term limits, to name but two important differences between states with and without term limits (Besley and Case 1995b, Table III, p. 778).

Besley and Case (1995b) identify the effect of a term limit from the difference between first and second terms in office for incumbents who face term limits. Controlling for state fixed effects and year effects, and using annual data from the 48 continental U.S. states from 1950-86, they find that a variety of policy measures are affected by term limits. Specifically, state taxes and spending are higher in the second term when term limits bind in states that have them. Such limits tend to induce a fiscal cycle with states having lower taxes and spending in the first gubernatorial term compared to the second.

List and Sturm (2001) apply a similar methodology to cross-state variation in environmental policy. They find that governors in their last term in office are significantly more likely to spend resources on environmental protection using data for the period 1960-1999. However, this term limit effect is muted in states where a larger fraction of citizens belong to environmental organizations. They also show that their term limit effect varies according to the margin of victory in the gubernatorial race going into the last term in office – with term limit effects being attenuated when the margin of victory is larger.

The empirical framework that we use here is as follows. Suppose that there is a policy that can be chosen in state \( s \) at time \( t \) denoted by \( p_{st} \). Then, we model this as

\[
p_{st} = \alpha_s + \beta_t + \gamma t_{st} + \theta y_{st} + \varepsilon_{st}
\]

where \( \alpha_s \) is a state fixed effect and \( \beta_t \) is year dummy variable. The variable \( t_{st} \) is one in years in which there is a binding term limit and zero otherwise. The term limit effect is then estimated from \( \gamma \). The control variables \( y_{st} \) that we use are log of real per capita income, log of state population, proportion aged over 65, proportion under the age of 18, whether the governor is a Democrat, whether the Democrats hold a majority of the seats.
in the house and senate of the state government and whether there is divided control of the house/senate and the governor’s chair. To interpret this equation, note that all governors who are facing a term-limit are being compared to all those governors who are unconstrained.

The means of the raw data are in Table 3.4.

Table 3.4 about here

Results from running a specification of this form are given in Table 3.5. In the first column, we look at real government spending per capita. These show that there is a term limit effect with states that have a term-limited governor spending significantly more. There is a weak effect of a similar kind for total taxes (column (2)). When taxes are disaggregated, we find a significant effect for both personal income and corporate income taxes (columns (4) and (5)). The consistent effect of term-limits can be contrasted with the patchy significance of the other political variables. There is some evidence that Democrat control in legislatures and the governor’s chair raises some kinds of taxes and spending. The evidence of divided government is of reduced taxation, especially of corporate income, but no effect on spending. This is consistent with earlier results by Alt and Lowry (1994) suggesting that a divided government tends to increase state deficits.

Table 3.5 about here

Table 3.6 looks at a different aspect of term limits. Instead of using a policy variable on the left hand side, we use a measure of policy congruence between citizens and governors generated by Berry et al (1998). Specifically, we use the absolute difference between the ADA and COPE scores of the governor and the citizens in state s at time t as a measure of congruence. These are ratings of the governors and voters ideology by two different groups – Americans for Democratic Action (ADA) and the AFL/CIO’s Committee on Political Education (COPE). They are placed on a scale between -100 and 0 where 0 denotes “perfect” congruence. The empirical results are striking. Governors who cannot run are significantly more congruent than those who can. This is suggestive of a strong selection effect into last periods in office where voters get to find incumbents who are more in tune with their preferences.

Table 3.6 about here

---

21In both cases, we condition on the state having a personal income tax and a corporate income tax, i.e. only consider states who permit such a tax.

22See Berry et al (1998) for more detail on these variables.

23As documented in Besley and Case (1995b) Table II there were some states that switched from a one term limit where every new incumbent is term limited. If these term-limited governors are separated out from term limited governors in general, then we find that most of the term limit effects are coming from them. This is in line with existence of selection effects and suggests that the selection effect for incumbents who survive beyond a single term is strong. For further discussion of one-term term limits and their interpretation, see Alt, de Mesquita and Rose, (2006).
3.4 Extensions

This section discusses a variety of extensions to the baseline model. As well as increasing the range of issues that can be discussed, it will also provide a way of locating many of the contributions to the literature that have used a political agency framework.

3.4.1 Polarization and Competition

In this subsection, we discuss three key ideas which affect the way in which elections achieve accountability. First, we consider the implications of randomness in the voting process. This could be interpreted in one of three ways – (i) non-policy relevant factors which shape the popularity of candidates, (ii) lack of information which leads to citizens to vote randomly for one or other candidate (iii) voter irrationality. This adds “noise” to the voting process which weakens the ability of the voting mechanism to hold politicians to account.

Second, we consider the implications of polarization among voters and politicians, i.e. competing views about policy. This could, for example, be based on ideology or ethnicity. Hence, we have in mind a world with two sets of policy issues – partisan issues on which voters disagree and non-partisan (or valence) issues on which they agree. The baseline model considered only a single valence issue. Here, we continue to study accountability on the valence dimension. When there is polarization on partisan issues, voters have a reason to like or dislike a candidate on the basis of some (other) reason other than their performance on non-partisan issues. This weakens the extent to which some voters hold politicians to account on the valence dimension.

In this more polarized world, the candidates may differ naturally in their electoral strength due to their partisan position – some jurisdictional boundaries may be drawn in favor of one of the candidates. The strength of that advantage is a crude way of thinking about the degree of political competition. Thus, our third extension is to consider how political competition impacts on accountability. It is sometimes suggested that polarized politics is a problem. However, we show that it is the lack of competition, coupled with polarization, rather than the polarization per se that matters to the quality of performance on valence issues.

To examine these issues, we extend the basic model above, by supposing that the incumbent and challenger have some kind of party affiliation. We label the parties $A$ and $B$ and assume that each party represents a fixed policy position and is associated with a policy agenda which is separable from the valence issues studied in the previous section which we shall now refer to as the non-partisan issue. It is clear that this separable approach is artificial. We will relax this when we consider the implications of multiple policy issues below.

\textsuperscript{24}The empirical analysis is similar in spirit to Svensson (2000).
CHAPTER 3. POLITICAL AGENCY AND ACCOUNTABILITY

We now consider heterogeneous voters. Specifically, voters are of three types – supporters of each party and non-committed voters who vote solely on the basis of the non-partisan issue according to which incumbents are held to account. Let $\omega$ be the fraction of partisan voters and $(1 - \omega)$ be the fraction of non-partisan voters. Partisan voters are assumed to get some utility $\phi > 0$ from having their preferred ideology in office. With $\phi = 0$, we are back to the model above. For low levels of $\phi$, the non-partisan issue will dominate in voting decisions. However, for large values of $\phi$, only the non-partisan voters will vote on the non-partisan issue. This is the case where there is political polarization and will be the case studied for the remainder of this section.

Even if there is polarization, there is no necessary reason to expect it to weaken accountability if the two sets of partisan voters voting intentions simply cancel each other out as will tend to be the case when they are equal in numbers. Hence suppose that $\frac{1}{2} + \eta$ of the partisan voters prefer party $A$.

We assume that there is some noise in the votes of non-partisan voters. We will model this within a simple probabilistic voting framework of the kind used extensively by Persson and Tabellini (2000). There is an idiosyncratic shock $\iota$ which is uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$ and an aggregate popularity shock $\delta$ which is uniformly distributed on the interval $[-\frac{1}{2\sigma}, \frac{1}{2\sigma}]$.

Suppose that the incumbent is of type $A$. It is now straightforward to see that the incumbent beats the randomly selected challenger if and only if $\omega\eta + (1 - \omega) [\Delta [\Pi - \pi] + \delta] > 0$

where $\Pi$ is the posterior probability that the incumbent is congruent. The term $\Delta [\Pi - \pi]$ is the reputational advantage of the incumbent while $\eta$ is the natural advantage and $\delta$ the advantage from the popularity shock.

So now the probability that the incumbent wins if he takes the congruent action is:

$$\sigma (\theta + \Delta [\Pi - \pi]) = \begin{cases} 
\frac{1}{2} + \xi [\theta + \Delta [\Pi - \pi]] & \text{if } \theta + \Delta [\Pi - \pi] > \frac{1}{2\sigma} \\
0 & \text{otherwise} 
\end{cases}$$

where $\theta = \frac{\omega}{1 - \omega} \eta$. This equation makes transparent how noise, polarization and competition drive accountability. We can interpret the parameter $\theta$ as a measure of pro-incumbent bias driven by whether the incumbent or challenger enjoys an advantage in terms of partisan supporters and how prevalent are partisan supporters in the population. On average we would expect incumbents to have a positive $\theta$ given that they have already been elected once in order to be incumbents.

The base-line model is a special case of this model where $\xi \to \infty$, $\omega \to 0$ and/or $\eta \to 0$. If there are sufficiently many partisan voters ($\omega \to 1$), then there is no effective accountability since $\sigma (\theta + \Delta [\Pi - \pi]) \to 1$ or 0 depending on whether $\eta$ is positive.
or negative. Lack of competition (high $\eta$) is only important when there is sufficient polarization (i.e., when $\phi$ is above $\Delta [\Pi - \pi]$). Thus, low levels of political competition require sufficient polarization if they are to affect the extent of electoral accountability.

Suppose we are in the case where the probability of winning lies between zero and one. Then, it is straightforward to see that he will choose the congruent action if and only if:

$$r_1 \leq \frac{[\sigma (\theta + \Delta [\Pi - \pi]) - \sigma (\theta)] \beta (\mu + E)}{\xi [\Delta [\Pi - \pi]] \beta (\mu + E)}.$$  

The term $\sigma (\theta + \Delta [\Pi - \pi]) - \sigma (\theta)$ represents the gain in the probability of winning if he produces $\Delta$ for the voters. Then the probability that the dissonant politician takes the congruent action is characterized by

$$\lambda = G \left( \xi \left[ \Delta \pi \left( \frac{(1 - \pi)(1 - \lambda)}{\pi + (1 - \pi) \lambda} \right) \right] \beta (\mu + E) \right).$$

One key observation with this is that an increase in electoral uncertainty (lower $\xi$) reduces discipline by dissonant politicians. This makes sense – voters are no longer able to credibly commit to vote for the incumbent even if he provides the congruent action. Thus, the greater the noise in the voting mechanism the less likely it is that a dissonant politician will behave in a congruent fashion in the first period.

The assumption of uniformity in the popularity shocks means that $\theta$ does not affect the rent seeking decision of the incumbent on the margin. This is clearly quite special. Suppose instead that the shock $\delta$ has a general distribution function $H (\delta)$ which is symmetric around zero (its mean). Then (3.4) needs to be modified to:

$$\sigma (\theta + \Delta [\Pi - \pi]) = \begin{cases} H (\theta + \Delta [\Pi - \pi]) & \text{if } \theta + \Delta [\Pi - \pi] > \frac{1}{\eta} \\ 0 & \text{if } \theta + \Delta [\Pi - \pi] < -\frac{1}{\eta} \end{cases}.$$  

(3.5)

Clearly this implies that the level of congruent behavior depends on $\theta$. Specifically,

$$\lambda = G (\sigma (\theta + \Delta [\Pi - \pi]) - \sigma (\theta)) \beta (\mu + E)$$

We show that in the most natural case, where $h (\delta)$ is unimodal and $\theta > 0$, this results in an incumbent having a smaller incentive to take the congruent action. Then,

$$\frac{\partial \sigma (\theta + \Delta [\Pi - \pi]) - \sigma (\theta)}{\partial \theta} = h (\theta + \Delta [\Pi - \pi]) - h (\theta) < 0$$

if $\theta > 0$. Thus, an increase in pro-incumbent bias as measured by $\theta$ reduces the incentive of the incumbent to produce the congruent action.$^{25}$ In the case where $0 >$

---

$^{25}$Clearly, in the limiting case where $H (\theta + \Delta [\Pi - \pi]) = H (\theta) = 1$ and there is no incentive for dissonant politicians to undertake congruent actions.
\[ \theta + \Delta [\Pi - \pi] > -\frac{1}{2\pi}, \] the comparative static on the effect of \( \theta \) goes the other way. This is the case where an incumbent has a very low chance of being re-elected. However, we do expect incumbents on average to have a positive \( \theta \) given that they were elected in the first place.

Thus, we have:

**Proposition 2** Suppose that the incumbent enjoys an advantage. Then, an increase in the incumbent’s advantage will reduce the congruence of first period actions among dissonant incumbents. This reduces the welfare of non-partisan voters.

In summary, this analysis gives reasons for believing that electoral accountability will be more effective in shaping incumbent’s actions when (i) there is less noise voting, (ii) there is less polarization among voters and (iii) there is evenly balanced political competition.\(^{26}\)

This theoretical finding echoes a widespread sentiment linking political competition and policy outcomes. This is often argued by political scientists in quite broad terms. For example, in his classic book, Schattschneider (1960) suggests that “the people are powerless if the political enterprise is not competitive. It is the competition of political organizations that provides people with opportunity to make a choice. Without this opportunity popular sovereignty amounts to nothing.” (Schattschneider (1960, page 140)).

There is a certain amount of suggestive empirical support for this proposition. List and Sturm (2002) observe that the Gubernatorial term limit effect for environmental policy depends on the margin with which the governor was originally elected which could be thought of as a measure of \( \theta \). Besley and Preston (2004) use data from U.K. local governments to argue that local authorities where the pattern of districting favors one party over another tend to have higher taxes and spending as well as large public employment. In similar vein, Besley and Burgess (2002) find that India states where the legislature has a tighter margin (in terms of the seats difference between the two major party coalitions) are more responsive in term of providing food aid when there are droughts and floods.

Having introduced heterogeneous voters, it becomes somewhat more difficult to make unequivocal statements about the quality of government. While it is clear that increases in \( \lambda \) still have favorable consequences for the non-partisan voters, whether the incumbent survives or not has a different welfare impact on the partisan voters who are happy to trade reduced accountability for a more favorable partisan bias in policy in their own direction. A welfare calculus would now require taking a stance on which of these groups is more deserving.

---

\(^{26}\)These results are similar to those derived in the more standard career concerns setting by Ashworth and Bueno de Mesquita (2005b).
3.4. EXTENSIONS

3.4.2 Information and Accountability

The political agency model is a natural vehicle for thinking about the role of information in achieving political accountability. This is because it focuses on the use that voters make of information that acquire about politicians and policy. In turn, we have seen that incumbents respond to this information when choosing their actions. This means that the political equilibrium is likely to be affected by the information that voters have. In this section, we explore how information structures affect political accountability through the lens of the agency model.\textsuperscript{27}

In practice, voters have many different sources of information provision such as the media, think tanks and academics. Information may also be part and parcel of political campaigns. Below, we will discuss how these specific sources of information can feed into the model.

The basic model supposed that the consequences of the policy choices of incumbents become known immediately since the voters observe $\Delta$. However, this is very strong and unrealistic. The political agency model frequently assumes that information about the quality of policy becomes known only after the election, even if the incumbent’s action is observed. This seems natural in the case of building a highway whose quality and location is difficult to gauge or fighting a war, the consequences of which may not be felt until far into the future. Obviously the world comprises a whole range of actions by politicians some of which may have immediate consequences and others for which it will take a while for information to be revealed.

Suppose then that with probability $\chi$ voters observe $\Delta$ after the politician has chosen $e$. We assume that this probability is the same whether or not the politician is congruent. Hence, there is no information from observing nothing. In addition to the possibility of observing $\Delta$, we also suppose that there is some other source of information about the incumbent’s type which does not come from their policy choice. Specifically, the incumbent’s type becomes known to the voters with probability $\tau$ some time after they have made their policy decision ($e$) but before the election is held. Everything else about the structure of the basic model is maintained. If no information is revealed, we assume that voters re-elect the incumbent.

Nothing is changed in period two. In period one, we need to modify the condition for the dissonant incumbent to be willing to do what voters want in period one. If the bad politician does not do this, voters will observe this with probability $\tau + (1 - \tau) \chi$. Thus, he will be re-elected with probability $(1 - \tau)(1 - \chi)$. However, if he takes the action that voters prefer, then he will be re-elected unless the voters get direct information about his type which occurs with probability $\tau$. Thus, weighing up, these re-election benefits against the forgone rent from taking the dissonant action, the probability that

\textsuperscript{27}In a related analysis, Ferejohn (1999) considers the possibility that politicians choose an information structure for voters which will act as a restraint on their own behavior.
a dissonant politician will be disciplined is:

$$\lambda = G\left((1 - \tau) \chi [\beta (\mu + E)]\right).$$

Inspecting this equation, it is clear that increasing $\chi$, the policy relevant information, increases the quality of the incumbent’s period one action and hence voter welfare in period one. However, increasing the probability that a dissonant incumbent will be found out for non-policy reasons reduces period one discipline. This is because the gain from taking the action that voters like is lower if he is not re-elected. In the limiting case where $\tau$ equals one, the incumbent has no advantage from doing what voters want in period one.

We now ask how increasing information (measured either by higher $\chi$ or higher $\tau$) affects voter welfare. It is straightforward to see that voter welfare is:

$$W(\lambda; \tau, \chi) = \{\pi + (1 - \pi) \lambda + \beta \pi [1 + (1 - \pi) [1 - (1 - \pi) [1 - \chi (1 - \lambda)]]]\} \Delta.$$

Observe that $\tau$ and $\chi$ enter this formula directly while also affecting the incumbent’s actions (as measured by $\lambda$). It is straightforward to see that the direct effect of an increase in $\tau$ or $\chi$ is positive since there is an improvement in the selection of politicians.

As in the base-line case, an increase in $\lambda$ raises voter welfare. Although the sign of the effect on period two welfare is the opposite of that on period one welfare, the order of the period two effect is smaller. Since $\lambda$ is increasing in $\chi$ an improvement in the observability of policy actions unambiguously raises voter welfare.

The effect of an increase in $\tau$ is, however, ambiguous. There is a positive effect on voter welfare in period two, since the improved information increases the chances of sorting in congruent politicians in period two. However, as we observed above, this comes at the price of reduced period one discipline — lower $\lambda$.

Differentiating our expression for voter welfare with respect to $\tau$, the overall effect is:

$$(1 - \pi) \left[1 - \beta \pi (1 - \tau) \chi \right] \frac{\partial \lambda}{\partial \tau} + \beta \pi [1 - \chi (1 - \lambda)] \lambda.$$

This expression is clearly negative for small enough $\pi$. In other words better (non-policy based) information about incumbents reduces welfare when the fraction of dissonant incumbents is sufficiently large. This is true because the selection benefits of better information are negligible in this case — it is highly likely that one dissonant incumbent will be replaced by another — and yet there is a negative discipline effect.\(^{28}\) We summarize this as:

\(^{28}\)We shall see other applications of this result in the next section. There, the information structure that achieves this is slightly different — there is some action that only a congruent incumbent is willing to take in equilibrium. The key thing is that there be something that can happen which reveals for sure that an incumbent is good. I thank Michael Smart for helpful discussions on this point.
3.4. EXTENSIONS

**Proposition 3** Improving information about which policy choice the incumbent has taken (i.e., higher $\chi$) raises voter welfare. Improving information that is not directly related to the policy action but which reveals the incumbent’s type (higher $\tau$) need not raise voter welfare and will reduce it when there are sufficiently few congruent types in the pool of potential politicians.

These results show that claims about the effect of information in politics must be treated carefully. In the next chapter we consider a number of applications of these ideas in a public finance setting. Underlying this observation is the fact that we are in a second best setting where voters imperfectly discipline politicians with highly incomplete incentives.

Whether these results should lead us to question the wisdom of providing better information about political behavior to voters is moot. On the whole we might expect this to yield an improvement in accountability. But there is clearly no formal result to support this unequivocally. The specific applications in a public finance model in the next section underline this.

What does appear general is that to evaluate the effects of informational improvements, both discipline and selection effects should be weighed up. In cases where the political class is of low quality (as represented here by low $\pi$), the impact on discipline will be key. However, high $\pi$ polities may yield different conclusions given the importance of selection. The models certainly cast light on the importance of the quality of the political class in thinking about these issues.

The ideas in this section are particularly relevant in thinking about the role of the media in achieving accountability. Besley and Burgess (2002) and Besley and Prat (2004) have used these models to look at the role of the media. Besley and Prat (2004) use a model where improving the role of the media is very similar to an increase in $\chi$ in this simple model. They use the model to derive observable implications for the effect of the media on turnover of incumbents. To see how this plays out, note that the probability that the incumbent is defeated in period one is:

$$(1 - \pi) [1 - (1 - \tau) [1 - \chi (1 - \lambda)]] .$$

First observe that, for fixed $\lambda$, turnover is increasing in both $\tau$ and $\chi$. Thus, we would expect this effect of increased media activism to be associated with more short-lived politicians. As is observed by Besley and Prat (2004), the effect operating through $\lambda$ does lend some ambiguity to this proposition – if media activism raises $\lambda$ then this will tend to dampen (and may even offset) the direct effect. Note, however, that if media activism raises $\tau$ then both of its effect (direct and operating through $\lambda$) will increase political turnover.

---

29 In a related contribution, Prat (2005) considers a model where the action taken by an agent may not be observed and shows that this may sometimes be a good thing.
In Besley and Prat (2004), the information provided by the media is endogenous. This requires a model of the media sector and the possible implications of attempts by incumbent politicians to silence the media. They assume that the media are motivated to reach the largest possible audience. Thus, competition for audience pushes the media to look for interesting news and to establish a reputation for reliability. For simplicity, we focus in on the implications of $\chi$ and set $\tau = 0$. Suppose that in the absence of media, $\chi = 0$. In this case, $\lambda = 0$. Voters receive no information and re-elect all politicians as $\tau = 0$.

The media industry is made up of $n$ identical outlets. Assume that they all receive verifiable news on the voter’s payoff with probability $\chi$ and that this can be reported to the public (for simplicity we assume that they are either all informed or all uninformed). We assume that a media outlet cannot fabricate news and that an outlet that reports informative news has a higher audience than one that reports no news. Moreover, the audience share of an outlet that reports news is decreasing in the number of other outlets that report news. The best case for an outlet is to be the only one to break news. In this world, news can only be bad for a dissonant incumbent. The important assumption is that news cannot be fabricated. Then, the only situation in which the voters’ behavior is affected by news is when the news media report that $e \neq s$.

Suppose that the media market comprises a potential audience of $a$ which is divided up among the number of media outlets who print news. The incumbent is assumed to be willing/able to bribe the media in exchange for silence if the media have news that the voter’s payoff is zero. Assume that there is a transactions cost so that out of any bribe to the media, they receive only $1/\nu$ of the payment. The timing of the game is as follows: (1) The media outlets receive or do not receive verifiable information about the incumbent; (2) The incumbent knows what information the media got and makes them transfer offers; (3) Each outlet chooses whether to accept or reject the offer; (4) The outlets that accept the offer suppress their information; the ones that reject it report their information to voters; (5) Voters re-elect the incumbent or replace her with a challenger.

In searching for the equilibrium of this game, the main question is whether the incumbent finds it profitable to buy off the media industry or not. If an outlet thinks that all the other outlets are going to be quiet, then its incentive to reject the incumbent’s offer goes up because it would be the only one to break news to voters and it would gain a large audience. This means that, in an equilibrium in which all media sell out, the incumbent must pay each outlet as if it were the only one who could break news. Even if we keep the total industry potential revenues constant, increasing the number of media outlets makes it more expensive for the incumbent to buy their silence. This is the sense in which media pluralism is good for media independence.

Adapting the main result of Besley and Prat (2004) to our setting, whether or not
the media is free depends on a comparison of \((\nu \times a \times n)^{30}\) - the cost of silencing the media – with \(\beta (\mu + E)\) the benefits to a dissonant politician of being re-elected. From this, it is clear that besides the number of outlets, the other parameters that determine whether media are captured or not in equilibrium are transaction costs and the amount of audience-related revenues. Both \(n\) and \(a\) decrease the likelihood that the incumbent manages to silence media. Instead, the probability that media are informed does not affect media capture but of course it increases the probability that voters are informed.

The Besley-Prat model links three key endogenous variables – whether the media operates freely, the turnover among politicians (higher with media freedom) and the quality of government policy (here measured by the probability that voters receive \(\Delta\)). While it is hard to look at this empirically, it is interesting to look at the core correlations between these three variables. To measure quality of government we use cross-country data on corruption. Clearly this is only suggestive and there are many issues with such data. However, it is notable that in Table 3.7 that these correlations come out exactly in the way in which the agency model suggests. Turnover is higher with more print freedom and lower where there is more corruption. Corruption and print freedom are also negatively related. While only suggestive, it is hard to think of other simple models that would so readily provide a parsimonious account of these basic facts.\(^{31}\)

<table>
<thead>
<tr>
<th>Table 3.7 about here</th>
</tr>
</thead>
</table>

Besley and Burgess (2002) which also invokes a political agency model in studying the media has a slightly different focus. Here the issue is whether a politician is responsive to a particular vulnerable group of citizens that has been hit by a shock. The context they have in mind for this is India where state governments have to decide whether or not to respond to droughts or floods through public food distribution or organizing public works. They consider a two period model where the responsiveness decision of the politician must be made in each period. Vulnerable voters care about electing responsive politicians – those who will act to protect the vulnerable apart from the electoral gain that it generates. They must infer responsiveness from the observed policy choices. In this model the media can expand the scope for reputation formation by informing a larger group of voters about the politician’s period one actions. Besley and Burgess (2002) assumes that the vulnerable are a geographically dispersed group and hence they can only see whether or not the incumbent has helped their group if they are subject to a shock. However, if some other vulnerable group experiences a

\(^{30}\)This is equal to the number of media times the amount that has to be paid to them to prevent them from running the story times the size of the transaction cost.

\(^{31}\)Besley and Prat (2004) argue that the degree of media freedom could be reflected in the structure of media ownership within a country using data from Djankov et al (2003). They find a robust correlation between state and foreign media ownership, concentration in ownership and corruption levels. They argue that these variables capture the underlying parameters that the agency model suggests are important.
shock, they may not see. Whether the incumbent responds to that group is relevant to learning the politician’s type. Media provide that information and this can enhance the incentive that an incumbent has to respond. Thus, the presence of active media can enhance the role of re-election incentives and hence increase government responsiveness.

Besley and Burgess (2002) look at the association between public food distribution and food production in Indian states over the period 1958-92. They find that states that have more newspapers per capita tend to have higher levels of food distribution and are more responsive on the margin to food production. They find similar evidence for the responsiveness of calamity relief expenditures to floods.

Media is clearly only one source of information provided in a democracy. The model could also be used to think about whether better information as provided via an electoral campaign yields welfare improving discipline and selection. The results that we have found do suggest that all information may not be equal in this regard.

When looking cross-sectionally at the quality of government with relatively similar electoral institutions, the provision of information could be an important source of heterogeneity in government performance.\footnote{For example, Ferras and Finan (2005) use an agency model to interpret their empirical results on the effect of audits on electoral outcomes in Brazil. They find that such audits have a significant effect on election rates of incumbents, reducing them by around 27%.

33 Alt and Dreyer Lassen (2002) and Alt, Dreyer Lassen and Skilling (2002) argue that fiscal transparency has an impact of fiscal policy both in the OECD and in U.S. states respectively.} We will discuss application of these ideas to fiscal transparency in greater detail in the next chapter.\footnote{For example, Ferras and Finan (2005) use an agency model to interpret their empirical results on the effect of audits on electoral outcomes in Brazil. They find that such audits have a significant effect on election rates of incumbents, reducing them by around 27%.}

The agency model is a natural vehicle for thinking about this.

\subsection*{3.4.3 The Nature of the Distortion}

In the first generation models of Barro (1973) and Ferejohn (1986), the focus of the agency problem is on politicians who “shirk”. There are two main varieties of shirking – reduced effort below what voters would like and diverting resources towards private ends. The latter could be interpreted as straightforward corruption or just rewarding cronies for political favors. What is key in both cases is that there is under-supply of something which is valuable to voters. In a broad sense, the non-congruent action of the incumbent in the base-line model developed here could be interpreted as a form of shirking. In this setting, the role of election incentives is to promote greater levels of “good behavior” and reduce shirking. The analysis then examines how far the threat of not being re-elected can sustain outcomes that voters desire.

In this world faithful politicians continue to deliver goods even in the face of re-election. Elections improve the behavior of the wayward without damaging the behavior of the well-intentioned. A small emerging literature is concerned with the possibility that agency can lead to poorer quality social decisions because politicians tend to choose outcomes that are too close to what voters want. This is most relevant when politicians
3.4. EXTENSIONS

have better information than voters. A conflict arises when this information goes against what voters would most likely think to be optimal. Re-election incentives may then lead to politicians to choose excessively popular policies.

To introduce these ideas into our setting, we change the model slightly to create incentives for a congruent politician to pick an action in period one that is not in line with what voters want. We assume now that dissonant politicians are biased towards picking one particular action by supposing that the rent is attached solely to the policy choice $e = 1$ regardless of the state of the world. By choosing $e = 0$, regardless of the state, he earns no rent. We also assume that the payoff from the politician’s action is only observed after the election (equivalent to $\chi = 0$) in the information model of the last section. We continue to assume that each state occurs with equal probability.

The analysis of period two behavior is now changed somewhat even though each type of politician still picks their preferred action. Now a dissonant politician will pick $e = 1$ regardless of the state while a congruent politician continues to do what voters would like. Hence $e_2(s, d) = 1$ and $e_2(s, c) = s_2$.

Turning to period one, we are interested in characterizing the behavior of both kinds of incumbents. There are a number of possibilities. However, we focus on a situation where (i) a congruent politician picks $e = 0$ in period one and is re-elected for sure, (ii) a dissonant politician also picks $e = 0$ when his rent from choosing $e = 1$ is low (less than $\beta(E + \mu)$) and is re-elected for doing so, (iii) a dissonant politician picks $e = 1$ when his rent is above $\beta(E + \mu)$ and is removed from office by the voters.

To see when an equilibrium of this form exists, we need to check that voters and politicians are optimizing. Observe that the probability that a dissonant incumbent picks $e = 0$ is now:

$$\lambda = G(\beta(\mu + E)).$$

The logic for a dissonant incumbent being willing to do this is exactly that encountered in the baseline model – weighing up his period one rent from $e = 1$ against his future payoff from being re-elected. We now show that any incumbent picking $e = 0$ will be re-elected. By Bayes rule, the probability that an incumbent is congruent conditional on having picked $e = 0$ in period one is:

$$\frac{\pi}{\pi + (1 - \pi)\lambda} > \pi.$$ 

Hence, voter’s re-elect for sure. Finally, we need to consider the incentives of congruent politicians. So far, we have not been very specific about their payoffs except to say that they are inclined towards doing what voters want. We suppose now that they receive a payoff of $E$ from holding office and care directly about voter welfare too. Then in each period that they are in office they receive:

$$\Delta(es + (1 - e)(1 - s)) + E.$$
These payoffs would lead them to (optimally) behave as we have assumed so far. We also suppose that they care about voters whether or not they are actually in office. Thus, they might sometimes be willing to sacrifice their position in public life “for the good of the voters”.

In state $s = 0$, a congruent politician faces no dilemma since he can pick $e = 0$ and still be re-elected – his private payoff and voters’ payoffs move in the same way. In $s = 1$, he weighs up picking $e = 1$ and being voted out of office in which case his payoff is:

$$\Delta + \beta \pi \Delta$$

or picking $e = 0$ and getting an expected payoff of:

$$\beta [E + \Delta].$$

The equilibrium where the congruent incumbent always chooses $e = 0$ in period one is possible when:

$$\frac{\Delta}{E + \Delta} < \frac{\beta}{1 + \beta \pi}.$$

It is clear that a necessary condition for this is that $E > 0$, there is some private utility from holding office. Moreover, this condition will always hold when the private rents from holding office are large enough. The condition is also more likely to hold when $\pi$ is close to zero. This is because then a congruent incumbent knows it is very likely that he would be replace by a dissonant incumbent in period two and he is making a bad policy choice today in order to protect the voters in period two. For this same reason, the equilibrium where $e = 0$ is always chosen in period one is also more likely when the incumbent is patient. Relating this back to the discussion in chapter 2, this constitutes another example of political failure (in the Pareto sense) due to a political linkage.

In this equilibrium, the congruent incumbent is “timid” to use a phrase from Smart and Sturm (2004). While he knows that he is picking something that is bad for voters, he is not courageous enough to do so if the price is to lose the election. The key thing here is the fact that dissonant politicians are biased towards one particular action. Congruent politicians cannot take that action without being mistaken for being dissonant themselves.³⁴

This equilibrium is also related to the idea that politicians sometimes pander to the voters. This idea is developed in Morris (2001) and Maskin and Tirole (2004) using a somewhat different, but related, model.³⁵ There are two time periods with an election

³⁴ Leon (2004) applies this kind of logic to the issue of why governments may be deterred from pursuing long-term projects in the development process when dissonant incumbents are biased towards picking long-term projects.

³⁵ Pandering is also modeled by Canes-Wrone, Herron and Shotts (2001). They have a slightly different set-up from Maskin and Tirole (2004) – their politicians are not differentiated by their motivation, but by whether or not they know which is the optimal action. However, the definition of pandering
between them. Suppose that there is a single social decision to be taken in each period which can take on one of two values denoted by \( x \in \{a, b\} \). There are two kinds of politicians – those whose preferences are congruent with the electorate and those that are not – this is not observed by the voters. The incumbent politician knows the true state of the world. Voters do not observe this, but have a prior belief that action \( a \) is more likely optimal than \( b \). A good example would be a politician who sees that a tax increase is necessary to stave off a fiscal crisis. However, the voters’ prior is that tax increases are more likely to enhance opportunities for rent seeking and hence do not look favorably on it.

While good politicians care about making good decisions they also get some kind of personal return to holding office. It is the latter that creates the crucial agency problem. If he cares sufficiently about staying in office, a good politician can be induced to take the action that the voters believe to be optimal even if his information says that action \( b \) is optimal for voters. This is the pandering effect of political agency modeling. Voters now learn nothing from seeing the action \( a \) being taken and re-elect. By contrast, when re-election incentives are weak, good politicians choose the unpopular action even when it is known that they will be fired for doing so – Maskin and Tirole (2004) refer to this as the courageous equilibrium. Voters fire them because they believe that such actions are more likely to come from a bad politician than a good one. This is very similar to what we have described above. As observed by Canes-Wrone et al (2001), there are good examples of pandering in real political behavior.

As with many of the results discussed here, pandering hinges critically on the ultimate success of the action taken not being observed before the election. To the extent that information is revealed throughout a term in office, we would expect political courage to be shown earlier in an election term while pandering is more likely when an election looms. This would give rise to a pandering cycle which is parallel to Rogoff (1990)’s equilibrium political budget cycle which we discuss further below.

When distortions affect the behavior of congruent politicians, there is a question of whether subjecting individuals to re-election is a good idea at all. While periodic election is a cornerstone idea in representative democracy, there are limits to the kind of offices that we require periodic election for. In the U.K., the House of Lords comprises individuals who are not subject to re-election concerns. Many systems of government do not use elected judges or elected regulators. One argument for this is that even good politicians are induced to distort their behavior. In similar vein Smart and Sturm (2004) argue that timidity among politicians could yield an argument in favor of term-limits.

Here, let’s illustrate this by contrasting voter welfare over the two period horizon with and without an election between.\textsuperscript{36} In the case studied so far, it will typically be is the same – a politician fails to follow information about the optimal action for fear of its electoral consequences.

\textsuperscript{36} Maskin and Tirole (2004) contrast three possibilities in their model: representative democracy, direct democracy and judicial decisions. The latter differs from representative democracy since the
optimal to hold an election. However, these need no longer be true. Observe that the payoff with a “life term” for the incumbent, i.e. no re-election, is:

$$\pi (1 + \beta) \Delta$$

while with an intermediate election it is:

$$\left[ \pi \left( \frac{1}{2} + \beta \right) + (1 - \pi) \left( \frac{1 + \lambda}{2} + \pi (1 - \lambda) \right) \right] \Delta.$$ 

Thus an election is optimal if

$$(1 - \pi) \left[ \frac{1 + \lambda}{2} + \pi (1 - \lambda) \right] > \frac{1}{2} \pi.$$ 

This is always the case as $\pi$ tends to zero and never the case as $\pi$ tends to one. This makes sense – if you can guarantee a congruent politician from the start, then an electoral mechanism is now purely a source of distortion. Equally, if the politician is certain to be bad some discipline is achieved by having an election.

This section has reinforced the point that political agency is not solely about the problem of under-achievement by dissonant politicians. It may also have to deal with electoral distortions among broadly well-meaning politicians. However, this does require that congruent politicians have some reward from holding office. The model suggests that crude efficiency wage arguments need not apply in politics since they may induce distortions of the kind discussed here.

### 3.4.4 Within-Term Cycles

A stock idea in political economy is that politicians behave differently when elections loom. The agency model provides a natural way of thinking about such issues in general and it is useful to see how the baseline model can be amended to handle them. One simple way to do this is by supposing that the quality of policy choices is most visible the longer the voters have had to assess their consequences and hence there is a difference between actions taken close to an election and those that are further from one.

Suppose then that we augment our model by having three time periods – two before the election and one afterwards. This is artificial, but serves to illustrate our points. We label these three periods 1, 2, and 3. We assume that the rent from taking the dissonant action is drawn afresh each period and draws from the distribution $G(r)$ are independent. Suppose also that the payoff of voters is only observable with a one period
lag. Thus, the payoff from the period one action of incumbents is observed by voters only in period two and the payoff from period two’s action is only observed in period three after the election has been held. We show that this will lead to different incentives for dissonant politicians to do what voters want in periods 1 and 2. Let $\lambda_1$ and $\lambda_2$ denote the probability that the dissonant incumbent will take the action that voters desire in periods one and two respectively.

The voters now conditions their voting decisions on the action taken in both periods and any payoff information that he has. It is easiest to begin with period two which follows the case from the model of section 4.2 with $\chi = 0$ and $\tau = 0$. Hence, $\lambda_2 = 0$ – the dissonant politician will take whatever rent is on offer and voters get no information on the action taken by the incumbent.

The period one action adds in the fact that the payoff information will be available to evaluate the incumbent by the time that the election comes around. Hence, when the incumbent weighs up the benefit of taking the rent in period one and waiting to collect the full rent in period three, the probability that he gives up the rent in period one is

$$\lambda_1 = G(\beta^2(\mu + E)).$$

This is basically the same as the baseline model but for the fact that the agent now waits for two periods and hence the discount factor is $\beta^2$.

It is now clear that the model generates different payoffs to voters depending on whether an election is close or not. Moreover, this is fully consistent with voter and incumbent rationality. The expected payoff to a voter in period one is $(\pi + (1 - \pi) \lambda_1) \Delta$ while in period two it is $\pi \Delta$. In this model, pre-election government behavior (period two) is more ill-disciplined than that at the beginning of the term (period one) which seems to parallel the conventional wisdom on political cycles.

The classic instance of election cycles is the macroeconomic political business cycle. Early interest in this idea was sparked by contributions by Nordhaus (1975) and Hibbs (1977). However suspicion was cast over their claims since they seemed inconsistent with rational behavior by voters. However, Rogoff and Siebert (1988) and Rogoff (1990) developed these ideas using an agency model in which both politicians and voters are rational.

Rogoff (1990) reformulated the model in a political agency model as follows. The model is an infinite horizon where an incumbent politician must choose levels of public goods, taxes and public investment. The level of public investment is observable with a lag while taxes and public goods are observable immediately. An election takes place every other period, but in each period, there is a shock to competence observed only by the incumbent. High competence means that an incumbent can produce a higher level of public goods for a given tax revenue. The competence level in any given period is a moving average of the shocks at dates $t$ and $t - 1$. The value of this shock becomes known to voters with a one period lag. Hence when an election comes, they observe the
“off-year” competence shock but not the election year value.

Rogoff assumes that politicians share voters’ concerns about taxing and spending. However, they get some independent level of utility from holding office (ego rents). These drive the agency problem as incompetent politicians will not be willing to voluntarily step down from office even though doing so would do a favor to voters. Faced with this, the onus on competent politicians is to find some way of revealing to voters that they are good. This signaling is what generates the equilibrium budget cycle. Moreover, there is greater uncertainty about more recent shocks than those that occurred earlier in the politician’s term.

The signaling process works as follows. A competent incumbent finds it less costly to choose higher levels of public goods provision than an incompetent one as the latter has to give up more public investment to finance a given level of public goods provision than the former (holding taxes fixed). For some level of public goods provision, therefore, an incompetent incumbent would rather choose a lower level of such spending and be defeated for sure.

In this model, the competent incumbents are responsible for the electoral cycle in taxes and spending. If there were no re-election possible, then no cycle would occur. Moreover, in this kind of model, a higher level of public goods spending and a lower level of taxes are associated with improved electoral chances. Le Borgne and Lockwood (2001b) consider the possibility that politicians are chosen endogenously. They find that the argument for cycles is muted by this.

Robust evidence of the political business cycle has been remarkably hard to find. For example, the survey by Alesina and Roubini (1992) is decidedly equivocal on whether such cycles are in the data. However, most have not made direct use of the agency framework in looking at the data. An important exception is Pettersson-Lidbom (2003) which tests Rogoff’s electoral cycle model on Swedish data. He finds that spending is raised and taxes are cut in the election year. Moreover, in the election year, spending is higher for a government that will be re-elected as compared to those that will not be re-elected. He also finds that, in the post-election year, spending is higher and taxes are lower for re-elected governments compared to newly elected ones. Reelected governments appear to spend less and tax more in the post-election year as compared to the election year. These findings are all broadly consistent with Rogoff’s equilibrium budget cycle model. The political agency model has also been used in the empirical tests of Shi and Svensson (2002) which looks at cross-country data and finds evidence of significant spending increases before elections which are consistent with the Rogoff model.

3.4.5 Multiple Issues

The baseline model assumes that there is a single valence issue. In our discussion of polarization, we introduced some separable partisan issues. However, they were not
directly linked to the agency problem facing voters. In this section, we will introduce multiple issues in a way that interacts with agency problems. This serves to ask questions about the distortion of policy priorities in the face of agency problems.

There are a number of papers in the literature that have studied the way in which agency affects the choices across policy instruments. Harrington (1993) develops an agency model in which the extent of voter uncertainty about the optimal policy reduces the incentives of politicians to manipulate policies for electoral purposes. This is because the value of the signal created is less useful to voters. Mukand and Majumdar (2003) look at the incentives of politicians to experiment in the face of re-election concerns and investigate how this may lead to either excessive or insufficient experimentation. In a more practical setting, Besley and Burgess (2002) argue that the media can change the salience of issues and lead to more resources being targeted to disadvantaged groups in India. List and Sturm (2002) consider a model where the issue is whether U.S. governors will respond to environmental groups. The importance of environmental issues is endogenous and affected by the extent to which these issues are electorally salient compared to others. Greater political competition will tend to yield greater attention to issues in which agency is important.

Here, we consider two applications of multiple issues in a little more detail. First, we consider what happens when politicians can target particular groups of voters. Second, we consider when the incumbent will pick the most efficient policies from the view point of the voters.

**Clientelism and Targeted Policies**

We supposed in the model above that all voters receive $\Delta$, i.e. there is no way of targeting specific groups of voters. Suppose instead that there are three groups of voters of equal size (each comprising one third of the population). Each group is labelled $j \in \{A, B, C\}$. The incumbent can now decide whether to pick $e_j^t \in \{0, 1\}$ for each group $j$ at each date $t$. We assume that the state of the world is common for all three groups. If a dissonant politician picks $e_j^t = (1 - s_t)$ then he gets $\frac{1}{3}$ of the rent on offer. Hence, a politician can now decided to please only a sub-set of voters. We begin by assuming that the voters only observe their own $\Delta$ – they know whether the politician has taken the action that they like, but nothing about what the politician has done for other voters.

In this simple example, a dissonant incumbent need now only do what a majority of voters want which here means targeting two of the groups. Hence, he can consume the policy rent for one of the groups and still get re-elected. The probability that he chooses to do what a majority (here two thirds) of the voters wants is:

$$\lambda = G\left( \frac{3}{2} \beta [\mu + E] \right)$$
which means that the probability of some discipline is increased in the targeted case.\footnote{It is now possible to have $\lambda = 1$.} If we assume that the incumbent picks two of three groups to favor at random then the probability that an incumbent who produces $\Delta$ is congruent is:

$$\frac{\pi}{\pi + (1 - \pi) \lambda^2}$$

which exceeds $\pi$. Whether targeting increases discipline from an ex ante point of view depends on whether:

$$\frac{2}{3} G \left( \frac{3}{2} \beta [\mu + E] \right) > G (\beta [\mu + E])$$

In the case of a uniform distribution $G(R)$, these probabilities are equal.

If the incumbent has even a small preference towards pleasing a particular group of voters, then the model delivers a kind of clientelistic politics in which the politician always serves that group of voters ahead of any others when assembling the majority that he needs to be re-elected. In that case, dissonant politicians (but not congruent ones) will use clientelism to get re-elected.

The result shown here depends crucially on voters only seeing what the politician has done for them rather than voters at large. If the entire vector of decisions made by the incumbent could be observed then it would be possible to spot the difference between a dissonant politician and congruent politician if the politician fails to produce $\Delta$ for any group of voters. If we thought that media was a vehicle for doing this, then this again provides a possible role for how information provision changes politics and policies. However, whether this increases or decreases the likelihood that all voters get what they want is unclear ex ante.

The Choice Between Efficient and Inefficient Policies

One abiding issue in political economy concerns whether incumbent politicians pursue their ends efficiently. Conditional on having a rent seeking incumbent voters would prefer that they take their tributes in the way that imposes least burden on themselves. The notion that rent seeking will be efficient is a key component of the Chicago School of political economy associated with Becker, Stigler and Peltzman.

In a world of full information, there is little reason for incumbents to behave inefficiently – even if voters do not like what politicians are doing, they would still prefer that they pursue their agenda efficiently. However, in a political agency model where information is imperfect, Coate and Morris (1995) show why governments may choose to hand out political favors in an inefficient way. This is because politicians may wish to disguise the fact that they like making such transfers.
3.4. EXTENSIONS

They use a two period agency model with an election at the end of period one. The main social decision to be taken is whether to construct a project. Incumbents also have the technology to make a cash transfer to a special interest group. For concreteness sake, let us describe the project as building a road. There are four key assumptions about road building. First, it benefits the special interest. Second, it may or may nor benefit voters. Third, politicians know better than voters whether the road is worthwhile. Fourth, citizens do not learn ex post (at least before the period one election) whether the road is worthwhile even if it is built.

The decision to build the road is made in period one and voters observe whether it has been built. There are two kinds of politicians. Good ones care solely about voter welfare whereas bad ones also care about the income received by the special interest. The voters care about having a good politician in office in period two since a bad incumbent will make a transfer to the special interest group in that period.

The initial electoral mechanism does not screen between the two types perfectly and there is some chance that a bad incumbent holds office in period one. If a bad incumbent is in office, then voters would prefer that he makes a cash transfer to the contractor rather than allowing him to build a road that is not worthwhile to the voters. However, such cash transfers would be observable to the voters and would lead them not to re-elect the incumbent. This leaves a bad incumbent an incentive to hide his type by mimicking a good incumbent. However, by building the road even if it is not worthwhile for the voters, this creates a disguised transfer to the special interest.

There are both pooling and separating equilibria. In the separating case, the bad incumbent makes a transfer to the special interest and is not re-elected for sure. In the pooling equilibrium, he makes the transfer via building the road. For this to be worthwhile to him, it must be the case that he is re-elected. This will happen if the voters think that it is sufficiently likely that the building decision is generated by a good incumbent who has seen that its construction is a good idea. Coate and Morris (1995) show that there are parameter values of their model which make this true.

Thus, a political agency model can explain why an inefficient project is chosen in a way that is consistent with rational forward looking behavior by voters and incumbents. The model speaks to debates about whether democracy produces efficient results in general (see Wittman (1989)). It emphasizes the importance of intertemporal distortions created by the re-election mechanism (see also Besley and Coate (1998)).

An important feature of this result is that both moral hazard and adverse selection are necessary to explain such behavior. If incumbents were known to be bad, then the

38 This differentiates this policy example from something like a tariff policy which harms voters relative to a combination of a production subsidy and consumption tax for sure and hence can never be in the voters’ interest.

39 Unlike much of the literature, Coate and Morris (1995) also consider whether a good incumbent will have an incentive to distort his behavior to try to separate from a bad incumbent. They show that, under reasonable assumptions, this will not happen.
voters would prefer that any transfer that they make to the special interest be made efficiently. There would certainly be no electoral gain from disguising transfers. The only issue would be (as in moral hazard models) whether the electoral mechanism could be used to curb bad incumbent’s inclinations to make transfers to special interests.

A nice application of the Coate and Morris framework is Sturm (2001) who tackles the case of “green” protectionism, invoking trade sanctions against products that are deemed to be damaging to the environment. He argues that this is best explained by the fact that voters are very poorly informed about whether or not such protection is justified. This gives extensive lee way to politicians to engage in such protection in order to favor some domestic firms.

The main idea of the Coate-Morris model can be seen using the set-up from section 4.3. Suppose that \( E = 0 \) so that the congruent politician always picks the correct policy. We look for an equilibrium in which the dissonant politician always picks \( e = 1 \) (i.e. building a road) regardless of the state when the rent from this action is high and picks \( e = 0 \) regardless of the state when the rent is low. Incumbents who pick \( e = 0 \) are re-elected and incumbents who pick \( e = 1 \) are voted out of office. The probability that \( e = 0 \) is now:

\[
\lambda = G (\beta \mu) .
\]

Using Bayes rule, an incumbent who picks \( e = 0 \) will be re-elected for sure if:

\[
\frac{\frac{1}{2}}{\frac{1}{2} + (1 - \pi) G (\beta \mu)} > \pi
\]

or \( G (\beta \mu) < \frac{1}{2} \).

The inefficiency arises because the dissonant incumbent does not index his action to the state of the world. If the state is \( s = 1 \) and \( r < \beta \mu \), then voters would prefer to write a contract with the incumbent which would allow him to keep the rent \( r \) in period one and \( \mu \) in period two in exchange for the action \( e = 1 \) in period one. This would make everyone (including the politician) better off. The informational structure precludes this since \( s \) and \( r \) are not observable to voters and hence not contractible. Moreover, clearly it would require contracts of a kind that we do not see in representative democracies. The incumbent is picking an action which could come from a congruent incumbent even though it is inferior for voters. They do so to enhance their reputation and hence to earn rents in period two.

### 3.4.6 Multiple Two-Period Terms

The baseline model assumed that time ends after the second period. In this section, we investigate what aspects of this are artificial by extending the model to consider a potentially infinite number of two-term incumbents. We show how this model can be used to gain insights into what happens when there is an endogenous pool of politicians.
This model also eliminates a very artificial feature of the two period model – in the two period model, any politician who is elected will be a lame duck in period two. It would be less artificial to allow the voters to choose between a fresh politician with a two-period time horizon and re-electing the incumbent for a final term. Intuitively, we would expect this possibility to make voters more choosy when deciding whether to re-elect an incumbent since first period politicians tend to behave in a more disciplined manner.

Suppose now that time is infinite, but that politicians can serve at most two terms in office with each term lasting one period. Thus, let \( t = 1, \ldots \) and let \( j \in \{1, 2\} \) denote the term in which the politician is currently serving. There is an infinite pool of potential politicians. A politician can serve only once after which he returns to the pool. We now index politician behavior by time and by the term that they are serving: \( e_t(s, i, j) \). We will work with stationary solutions, i.e. those in which these strategies are not time-dependent.

Congruent politicians always choose the outcome that voters prefer in both terms in office, i.e. \( e(s, c, j) = s \) for \( j \in \{1, 2\} \). The decision of the dissonant politician is exactly the same as in the baseline case, provided that it is optimal for the voters to re-elect a politician for delivering \( \Delta \). In order to check this, let \( V^N(\pi) \) be the value to the voter of starting over with a new politician and let \( \Pi \) be the probability that the incumbent is congruent conditional on him generating a benefit of \( \Delta \) in his first term. Using Bayes rule, \( \Pi \) is the same as in the baseline model and is equal to \( \frac{\pi}{\pi + (1 - \pi) \lambda} \). Let \( \phi(\pi) = \pi + (1 - \pi) \lambda \) be the probability that a first term incumbent generates the good action for voters. Then, the value of a newly elected politician under this strategy is:

\[
V^N(\pi) = \phi(\pi) \left[ \Delta + \beta \Pi \Delta + \beta^2 V^N(\pi) \right] + (1 - \phi(\pi)) \beta V^N(\pi).
\]

This has two parts. The first is the case in which the politician in period one does what voters want. In this case he is re-elected for one period after which a new politician comes in. The second term is for the case where the politician does not produce a period one benefit in which case, there is a random selection from the pool. Thus voter welfare viewed from period one is

\[
V^N(\pi) = \frac{\Delta}{(1 - \beta)} \cdot \frac{\phi(\pi)(1 + \beta \Pi)}{[1 + \beta \phi(\pi)]} = \frac{\Delta}{(1 - \beta)} \cdot \frac{\phi(\pi) + \pi \beta}{[1 + \beta \phi(\pi)]}.
\]

This has a nice interpretation. The first term is the discounted value of voter welfare if everyone behaved like congruent politicians. The second term (multiplying this) is less than one and represents the reduction in voter welfare due to the selection and discipline problems in political life.

Using this expression, we now check whether re-electing an incumbent who has produced \( \Delta \) is optimal for the voters. This will be true if and only if:

\[
\Pi \Delta + \beta V^N(\pi) \geq V^N(\pi)
\]
which is equivalent to $\Pi \geq \phi(\pi)$ or

$$\pi \geq \left(\frac{\lambda}{(1-\lambda)}\right)^2.$$

Intuitively, this condition says that re-election is worthwhile if and only if the conditional probability that the incumbent will make the right choice for voters exceeds the probability that a new incumbent will.

In the case where $\pi < \left(\frac{\lambda}{(1-\lambda)}\right)^2$, there is no equilibrium in which voters re-elect politicians who yield a payoff of $\Delta$ with probability one. This is because this strategy yields too much restraint by dissonant incumbents, reducing the importance of selection to a point where it would always be optimal to fire the incumbent.\(^{40}\) In this situation, there is a mixed strategy by voters. Let $\sigma$ be the probability that an incumbent is re-elected when he generates $\Delta$ for the voters. Then in a mixed strategy this must make the voters indifferent between re-electing and not re-electing the incumbent. This will happen when:

$$\lambda = G(\sigma\beta(\mu + E)) = \frac{\sqrt{\pi}}{1 + \sqrt{\pi}}.$$ 

This can be summarized as:

**Proposition 4** Congruent politicians always set $e = s$. Dissonant politicians always set $e = (1 - s)$ in their second term in office. First period behavior among dissonant politicians falls in one of two cases:

1. If $\pi \geq (G(\beta(\mu + E))/(1 - G(\beta(\mu + E))))^2$, then dissonant politicians deliver what voters want in period one with probability $G(\beta(\mu + E))$ and are re-elected for doing so.

2. If $\pi < (G(\beta(\mu + E))/(1 - G(\beta(\mu + E))))^2$, then voters re-elect politicians who deliver $\Delta$ with probability $\sigma$ where $\sigma$ solves

$$G(\sigma\beta(\mu + E)) = \frac{\sqrt{\pi}}{1 + \sqrt{\pi}},$$

and dissonant politicians deliver what voters want in period one with probability $G(\sigma\beta(\mu + E))$.

\(^{40}\)It is clear that there is no equilibrium in which voters always fire politicians for generating $\Delta$. All bad incumbents would then take the dissonant action. But then, observing a payoff of $\Delta$ would (using Bayes rule) lead the voters to update that the incumbent was good with probability one and hence re-elect them – a contradiction.
3.4. EXTENSIONS

The first of these outcomes replicates the behavioral pattern of the baseline model. This will happen only if \( \lambda \) is sufficiently low. While this may seem paradoxical, it reflects the fact that seeing \( \Delta \) must make voters sufficiently confident that the incumbent is congruent, otherwise, they would prefer to return to the pool and pick a fresh incumbent. This condition has the best chance of being satisfied when \( \pi \) is high since then it is mostly likely that \( \Delta \) has been generated by a congruent incumbent rather than a dissonant behaving in a disciplined fashion.

The second case occurs when \( \pi \) is very low. In this case, voters prefer to try their chances with a new politician when the strategy of the baseline model is followed. To make their signal more credible, dissonant politicians have to choose what voters want less of the time. Voters create this incentive by re-electing only some fraction of the time when the incumbent delivers \( \Delta \). Thus, even some congruent incumbents are removed from office as part of an optimal voting strategy.

Extending the time horizon of the model without changing the time horizon of politicians actually reduces the probability that a dissonant incumbent does what voters want. Hence, it worsens discipline and improves selection.

This model has a striking implication for the sign of the term-limit effect. Since voters only re-elect incumbents if they are better than a first term incumbent would be, the politicians must perform better than average in their second term than in the first. This reveals the power of the selection effect in agency models with adverse selection.

To summarize, this section has extended the two-period model with adverse selection and moral hazard to a multi-period setting while retaining the assumption that politicians can serve only two terms in office. This yields some qualitatively similar findings to the two-period model. However, there are some important differences in the logic due to the fact that voters can always pick a fresh incumbent at any time. This implies that any politician going into his last term in office must be better on average than a randomly selected politician from the pool.

Standing for Political Office

The model with an indefinite number of finite terms is a natural vehicle for looking at some factors that make the choice of the pool of politicians endogenous. This is because it lends itself to a steady state analysis where the value of any two period experience as a politician can be computed. The baseline set-up would give a different value to being a politician in the first and second period for highly artificial reasons.

Here, we look at this issue very crudely – the winning politician is a random selection from among those willing to serve. This abstracts from many institutions that affect candidate selection such as campaigning, party selection and voting.\(^{41}\) We suppose that there is a continuum of potential politicians with outside per period wages \( w \in [0, W_i] \)

---

\(^{41}\)See chapter 5 for a preliminary effort in this direction.
with \( i \in \{c, d\} \). We assume that this is uniformly distributed, with the support of the distributions differing for the two groups of politicians. Suppose that the probability that a candidate is of the congruent type in the entire pool of possible politicians is \( \gamma \).

For congruent politicians, recruitment implies retention. A politician will stay in office for a second term if \( E \geq w \). Since he is re-elected for sure, he will also enter in period one if \( E \geq w \). Thus, the fraction of congruent citizens who will be willing to enter politics is \( \frac{E}{w} \).

We now turn to the dissonant citizens and their decision to run as candidates. Adding in the outside option will now lead to a slight modification of the model above as the optimal action of dissonant politicians now depends upon their private sector wage rate. Specifically,

\[
e(s, d, 1, w) = \begin{cases} 
  s & \text{if } r_1 \leq \beta ([E + \mu] - w) \\
  (1 - s) & \text{otherwise}.
\end{cases}
\]

Thus, politicians with more generous outside opportunities are less likely to take the congruent action. This is because their rent from holding political office compared to the private sector is smaller. Let \( \bar{w} \) be the maximal private sector wage that a dissonant politician will forego in order to become a politician. Then the probability of congruent behavior by a randomly selected dissonant politician who is willing to serve is now

\[
\Lambda(E, \bar{w}) = \int_0^{\bar{w}} G(\beta ([E + \mu] - w)) \frac{dw}{w}.
\]

Dissonant politicians are willing to serve if they are re-elected if \( \mu + E \geq w \). We now look at recruitment on the assumption that this holds. If a politician stays in the private sector, he will make a stream of utility \( V(w) = \frac{w}{1 - \beta} \). Then, the value from entering politics when the private sector option is \( w \) is:

\[
P(E + \mu, w) = E + \left( \int_0^{\beta \beta (E + \mu) - w} (r + \beta V(w)) dG(r) \right) + G(\beta [E + \mu - w]) \left( \beta (E + \mu) + \beta^2 V(w) \right).
\]

The dissonant politicians who participate in politics are those for which \( V(w) \leq P(E + \mu, w) \). Solving this yields:

\[
w \leq (E + \mu) + \psi(E + \mu - w)
\]

where \( \psi(x) = -\frac{\int_0^{\beta \beta (E + \mu) - w} r dG(r)}{1 + G(\beta x) \beta} < 0 \) and \( \psi'(x) < 0 \). The last term on the right hand side of this expression is the expected forgone rent (i.e. below \( \mu \)) due to behaving well in

\[42\text{Two interesting possibilities are seniority premia for being in office (in the private sector) and a risk of not being re-elected in which case retention can be a problem.}\]
the first period. Thus, the critical wage below which dissonant politicians will put
themselves forward for office is defined by
$$\tilde{w}(E, \mu) \equiv (E + \mu) + \psi(E + \mu - \tilde{w}((E, \mu))).$$

We assume that there is a unique solution to this equation.\(^{43}\) It is clear that \(\tilde{w}(E, \mu) < E + \mu\). Hence, consistent with what we assumed above, dissonant politicians who are
willing to enter are also willing to serve a second term if they are re-elected. This
implies that the fraction of dissonant politicians who are willing to become politicians is:
$$\frac{\tilde{w}(E, \mu)}{W_d}.$$ Using this, the fraction of congruent politicians in the pool of candidates available for
public office is:
$$\pi(E) = \frac{\gamma}{\gamma + (1 - \gamma)\left[\frac{\tilde{w}(E, \mu)}{E} \cdot \frac{W_c}{W_d}\right]}.$$ This depends upon the relative returns to politics of dissonant and congruent types and their relative private sector options. The additional rents (over and above \(E\)) earned by dissonant politicians make them more eager to enter politics than congruent ones and raising wages redresses the balance. Thus, the model predicts that, in addition to
discipline effects, raising the rents earned by politicians (as measured by \(E\)) improves
the pool of politicians who are willing to serve.\(^{44}\) The size of this effect depends on
the ratio \((1 - \gamma) / \gamma\). Thus, the effect is larger if there is a relative dearth of congruent
individuals.

While simple, this does provide a way of closing the model in a way that does not
require us to treat \(\pi\) as exogenous. However, the model basically predicts that the
underlying character of the population (as represented by \(\gamma\)) is important. But so too
are the relative rewards of public and private life \(\tilde{w}(E, \mu) / E\) to congruent and dissonant
politicians.

### 3.4.7 Indefinite Terms

While the previous model relaxed the two period structure, it continued to assume that
each politician has only a two period horizon. With some investment in notation,

\(^{43}\)The possibility of multiple equilibria is intriguing. The mechanism is as follows. There can be
a high reservation wage option where a significant fraction of dissonant politicians enter and expected
rents are high since these politicians are less likely to be tempted to pick the congruent action. In the
low reservation wage equilibrium, expected rents are low since dissonant politicians who enter are more
likely to pick the dissonant action.

\(^{44}\)Observe that
$$\frac{\partial \tilde{w}(E + \mu)}{\partial E} = 1.$$
the model could straightforwardly be extended to consider finite terms of any length. However, any such model would have the feature that eventually some term would be known to be the last. In this section, we extend the model to lose this feature and consider a model where politicians do not face a known terminal date. Hence, we consider what would happen if politicians faced no finite horizon, i.e. could stay in office forever.\textsuperscript{45}\textsuperscript{46} We will characterize a stationary strategy for a dissonant politician in which voters re-elect politicians who produce $\Delta$ in any period. In this equilibrium, any politician who is fired will be replaced by another politician who also has an infinite horizon. Arguably this kind of model makes more sense for considering reputation formation by a long-lived entity like a political party rather than a single individual.

A dissonant politician will use a cut-off rule such that he picks what voters want if and only if the return to the dissonant action is less than or equal to $r$.\textsuperscript{46} He will be prepared to do so provided that this rent is less than what he will look forward to getting in the next period. To characterize this, it is necessary to compare the value of earning a rent today and being removed from office with the continuation value of being re-elected. The continuation value is equal to:

$$W (\hat{r}) = E + \left[ \int_{\hat{r}}^{R} r dG (r) + \beta G (\hat{r}) W (\hat{r}) \right].$$

Solving yields:

$$W (\hat{r}) = \frac{E + \int_{\hat{r}}^{R} r dG (r)}{1 - \beta G (\hat{r})}.$$

The critical value is now characterized by $\hat{r} = \beta W (\hat{r})$. This yields:

$$\hat{r} = \frac{\beta \left[ E + \int_{\hat{r}}^{R} r dG (r) \right]}{1 - \beta G (\hat{r})}.$$

Let

$$\hat{\lambda} = G (\hat{r})$$

be the probability that a dissonant incumbent does what voters want in any time period. It is clear that compared to the term limited case, there is less congruence along the equilibrium path than in the first period of a term-limited spell. Let $\hat{\phi} (\pi) = \pi + (1 - \pi) \hat{\lambda}$ be the extent of congruence along the stationary equilibrium path.

\textsuperscript{45}This section is inspired by Smart and Sturm (2004) who pose this question in a more sophisticated setting where there is no stationary equilibrium in the infinite horizon model.

\textsuperscript{46}The solution in this simple setting is stationary. This is because the politician’s reputation does not affect his/her incentive to do what the voters want. In general this will not be true and it is this feature of the infinite horizon model that makes it tricky to characterize equilibria in this setting.
Proposition 5 Suppose that a politician faces an infinite horizon model, then dissonant politicians are more disciplined than in period one when they face a binding term limit i.e., $\hat{\phi}(\pi) > \phi(\pi)$.

Proof: Define:

$$
\kappa(x) = \frac{\beta \left[ E + \int_{x}^{R} rdG(r) \right]}{[1 - \beta G(x)]}.
$$

It is straightforward to check that this is a strictly quasi-concave function which reaches a maximum at $\kappa(x) = x$. To see this, observe that:

$$
\kappa'(x) = (1 - G(x) \beta)^{-1} \beta g(x) [\kappa(x) - x].
$$

Moreover, $\kappa(0) = \beta (E + \mu)$. Thus, $\kappa(\hat{\pi}) > \kappa(0)$ implying that $\lambda < \hat{\lambda}$. QED.

This makes sense. Dissonant politicians now face the possibility of earning rents in many future periods in office and hence behave better. As the dissonant politician becomes more patient, then he may be willing to take the congruent action in every period, sustained only by the legitimate rent from holding public office. We record this as:

Proposition 6 Suppose that politicians are sufficiently patient and/or office holding rents are sufficiently strong, (specifically $\beta E/ (1 - \beta) > R$) then $\hat{\pi} > R$, all politicians do what voters want.

Thus with sufficiently patient politicians, a high enough legitimate office holding rent and a long enough time horizon, the political agency problem can be solved completely.

It is now straightforward to work out the welfare of voters along this path. First, observe that electing a congruent politician is an absorbing state – he always does what voters want and hence is re-elected forever. Let

$$
V^* = \frac{\Delta}{(1 - \beta)}
$$

be the value to the voter from this. Second, observe that any incumbent who produces $\Delta$ will be re-elected as the probability that he is congruent is strictly higher than a randomly selected incumbent from the pool. Now consider, the value to the voter at time zero. This is:

$$
\hat{V}_0 = \hat{\phi}(\pi) \Delta + \beta \pi V^* + \left( \hat{\phi}(\pi) - \pi \right) \beta \hat{V}_0 + \left( 1 - \hat{\phi}(\pi) \right) \beta \hat{V}_0
$$

$$
= \hat{\phi}(\pi) \Delta + \beta \pi V^* + (1 - \pi) \beta \hat{V}_0.
$$
Thus:

\[
\hat{V}_0 = \frac{\Delta}{1 - \beta} \left[ \frac{(1 - \beta) \phi(\pi) + \pi \beta}{1 - \beta (1 - \pi)} \right].
\]

Observe that \( \hat{V}_0 \to V^* \) as \( \hat{\phi}(\pi) \to 1 \). It is clear from this that politicians behave better when the value of holding office is greater.

This logic suggests that it will never be a good idea to impose terms limits on politicians since it worsens their behavior and means that periodically good politicians must be removed from office. This can be seen here by comparing \( \hat{V}_0 \) with \( V^N(\pi) \) in section 4.6. A term limit is a good idea if

\[
\frac{\hat{\phi}(\pi) - \beta(\hat{\phi}(\pi) - \pi)}{1 - \beta (1 - \pi)} < \frac{\phi(\pi) + \pi \beta}{1 + \beta \phi(\pi)}
\]

It is clear that in the case where the dissonant politician always takes the congruent action, then voters are strictly better off without term limits.

However, the conclusion that term limits are bad is not general. Smart and Sturm (2004) focus on the case of timid equilibrium along the lines discussed in section 4.3. This, as we saw, leads to distortions in behavior by congruent politicians. Moreover, this gets more likely when \( E \) is high. In this case, a term limit provides a way of inducing congruent politicians to behave in a congruent way – since they know that they cannot be re-elected. Smart and Sturm (2004) show that a two-term term-limit can sometimes be an optimal institution in that setting.

### 3.4.8 Multiple Agents

So far, we have considered a situation where there is a single agent responsible for all aspects of policy. This is far from the setting in which many policy decisions are made where many agents have some input into the policy process. We now consider the simplest extension which allows for multiple agents. These could be legislators who jointly pass legislation in a legislature or they could be thought of as the acts of a bicameral legislature, with agents located in each chamber. The latter provides a way into thinking about the welfare consequences of the separation of powers in this setting. It is this example which motivates the work of Persson, Roland and Tabellini (1997, 2000) which uses the agency model to think about these issues. We discuss their work in detail below.

From a modeling point of view, introducing multiple agents is interesting since the policy process now has to resolve externalities between the agents if they face a conflict of interest in their policy decisions. For example, when a congruent and a dissonant agent have to make policy together, this may affect rent extraction possibilities open to the latter. The interesting issue is whether these externalities between policy makers
work for or against the voters. The analysis here confirms the idea that a multiple agents setting reduces rent extraction possibilities. However, this may also reduce the period one discipline of dissonant incumbents, creating a trade-off for voters.

Suppose then that there are now two politicians labelled $\ell \in \{1, 2\}$. Each politician is directly elected and accountable to the same set of voters. We work with the case where $E = 0$. This washes out the possibility that having two politicians is beneficial because that generates twice the ego rent to discipline politicians. Politicians’ types are independent draws from the same distribution comprising dissonant and congruent politicians exactly as in the baseline model – so each is congruent with probability $\pi$.

As in the baseline model, there is a single policy decision to be made in each period. We denote this by $e_t \in \{0, 1\}$. Voters receive $r$ if $e_t = s_t$ and zero otherwise. The policy outcome is now dependent on the actions of the two agents. We denote these actions by $a'_t (\in \{0, 1\})$ for politician $\ell$ at time $t$ where the action can be interpreted as a vote in favor of $e_t$ being zero or one. These actions are observed by voters in each time period.

We suppose that there is an initial policy $e_0 \in \{0, 1\}$ in place which will remain the same unless the agents who control the policy process decide to change it. There is a fixed pot of rent, denoted by $r_t$, that is available to politicians who choose the policy $e_t \neq s_t$. This is shared equally between any agents who pick the action $a'_t = (1 - s_t)$, but only if the outcome is $e_t = (1 - s_t)$. In the event that both politicians pick $a'_t = (1 - s_t)$ and the outcome is $e_t = (1 - s_t)$, then he captures the full rent. The best way to think about this is as bribery by a special interest which is paid only if the policy turns out as the special interest would like. Thus, rents attach to policies not actions.

The two agents determine policy in the shadow of constitutional rules that determine the policy outcome as a function of their individual actions. Suppose that these rules can be described by a mapping from the actions of the two politicians and the status quo policy into a new policy outcome denoted by:

$$e_t = \Gamma (a^1_t, a^2_t, e_{t-1}) .$$

We assume that this constitutional rule satisfies unanimity, i.e.:

$$\Gamma (a^1_t, a^2_t, e_{t-1}) = \begin{cases} 
1 &\text{if } a^1_t = a^2_t = 1 \\
0 &\text{if } a^1_t = a^2_t = 0.
\end{cases}$$

Thus, both policy makers must agree to change the policy outcome from the status quo. Otherwise, we suppose that the constitutional rule favors the status quo in the sense that:

$$\Gamma (1, 0, e_{t-1}) = \Gamma (0, 1, e_{t-1}) = e_{t-1} .$$

\[47\text{Thus this differs from the previous sections where observing the action is equivalent to observing the policy outcome.}\]
Thus, the policy does not change if the two agents cannot agree to do so.

To model the game played between the politicians, we suppose that the politicians do not know each other’s type. We assume that the agents behave non-cooperatively and hence look for a Bayes-Nash equilibrium.

**Period two Policy** In period two, congruent politicians will do what voters want and dissonant politicians pick $a_1^t = (1 - s_2)$. The policy outcome is $e_2 = (1 - s_2)$ if there are two dissonant incumbents. With two congruent incumbents, the outcome is $e_2 = s_2$ in which case voters get $\Delta$ in period two.

With one dissonant and one congruent incumbent, the outcome depends on whether or not the state matches the status quo policy. If $s_2 = e_1$, then the congruent incumbent can ensure that the status quo prevails and $e_2 = s_2 = e_1$. Voters then receive $\Delta$. If the status quo policy does not agree with the state, i.e., $(1 - s_2) = e_1$, then the dissonant incumbent can preserve the status quo and earn a rent. A key observation is that, unlike the single agent model, voters get $\Delta$ some of the time when only one of the incumbents is dissonant. Thus, congruent incumbents can impose an externality on dissonant ones in period two which works to the voters’ advantage. Recalling that each of the states is equally likely, the voters get $\Delta$ with probability $1/2$ in this case.

It is clear from this that voters always gain from re-electing congruent incumbents in period one. Hence, as in the baseline model, incumbents who acquire a reputation for being congruent in period one may be re-elected.

Using this, we can compute the expected rent of a dissonant politician if he is re-elected. Let $\pi_2$ be the probability that the other politician is congruent. Then he obtains:

$$
\left[ \frac{1}{2} \left( \pi_2 + \frac{1 - \pi_2}{2} \right) + \frac{1}{2} \left( \frac{1 - \pi_2}{2} \right) \right] \beta \mu = \frac{1}{2} \beta \mu.
$$

To see this, observe that the first term is what happens when $e_1 \neq s_2$. In this case, the dissonant incumbent gets all the rent when he is with a congruent incumbent and half the rent when he is with a dissonant one. The second term is for when $e_1 = s_2$ in which case he gets half the rent if the other politician is dissonant.

**Period one policy** Suppose that voters use the following re-elected strategies. Re-elect both politicians if $a_1^1 = a_2^1$ and they receive a payoff of $\Delta$. Otherwise fire any politician whose action is not in favor of a policy that yields $\Delta$. Below we will show that these strategies are consistent with Bayes rule.

We now consider the behavior of the incumbents. The congruent politicians have a dominant strategy to pick $a_1^t = s_1$. Thus, the only interesting issue is how a dissonant politician behaves. This will depend on whether or not $e_0 = s_1$. We consider each case in turn.
3.4. EXTENSIONS

**Good Status Quo Policy:** $e_0 = s_1$  
In this case a dissonant politician knows that he can earn a rent only if he is with a dissonant politician who decides to take the rent so his expected payoff is $\frac{r}{2} (1 - \pi)$ which he compares to $\frac{1}{2} \beta \mu$. Thus, in this case he sets:

$$\bar{\lambda} = G \left( \frac{\beta \mu}{(1 - \pi)} \right).$$

So in this case, there is strictly more discipline than in the single agent case. As $\pi \to 1$, the dissonant agent knows that he cannot extract any rents and will do what voters want in the hope of capturing some rent in period two.

**Bad Status Quo Policy:** $e_0 = (1 - s_1)$  
In this case, a dissonant incumbent has more chance to earn a short-term rent. However, he will have to share it if the other incumbent is also dissonant. Hence he is actually better off with a congruent incumbent in this case. Thus expected rent from taking the dissonant action is now $\frac{r}{2} (1 + \pi)$. This yields discipline of:

$$\underline{\lambda} = G \left( \frac{\beta \mu}{(1 + \pi)} \right)$$

which is strictly less discipline than in the case of a single agent.

Thus, $\bar{\lambda} > \lambda > \underline{\lambda}$ if $\pi > 0$. If $\underline{\lambda} = \bar{\lambda} = \lambda$ if $\pi = 0$, i.e. all politicians are dissonant. In this case, there is no effect on discipline in this model. In this sense it raises rather different issues from the analysis of Persson, Roland and Tabellini (1997, 2000) whose contributions are discussed below. Here, the entire issue is how congruent and dissonant incumbents collectively create externalities for one another in the policy process.

**Proposition 7** Congruent politicians always set $e = s$. Dissonant politicians choose $e = (1 - s)$ in period two and will choose $e = s$ in period one if they earn sufficiently small rents from being dissonant. Dissonant politicians exercise more discipline in period one if $e_0 = s_1$ than when $e_0 = (1 - s_1)$.

Thus, having two agents creates negative and positive externalities. In the case where $e_0 = s_1$, a dissonant agent must rely on the other agent to be dissonant if he is to earn a rent in period one and hence it reduces the marginal benefit from taking the dissonant action in period one. However, if $e_0 = (1 - s_1)$, he can unilaterally earn the rent and would prefer that the other politician were congruent as that way, he earns a larger share of the rent. This encourages ill-discipline on the part of a dissonant politician.

**Welfare Consequences**  
To explore the welfare consequences of the two agent setting in more detail we derive the expressions for voter welfare in each period. In period one, the expression is:

$$\hat{V}_1 (\lambda^*) = [\pi^2 + (1 - \pi^2) \lambda^*] \Delta$$

(3.6)
where $\lambda^* = \frac{\lambda + \Delta}{2}$. This is exactly as if there was a single agent who is good with probability $\pi^2$ and the dissonant agent exercises discipline in period one with probability $\lambda^*$. Period two welfare is equal to:

$$\hat{V}_2(\lambda^*) = \pi [1 + (1 - \pi)(1 - \lambda^*)] \Delta. \quad (3.7)$$

Let

$$\hat{W}(\lambda^*) = \hat{V}_1(\lambda^*) + \beta \hat{V}_2(\lambda^*)$$

be overall voter welfare. As with (3.3), an increase in discipline (as measured by $\lambda^*$) is good for voters even though it lowers welfare in period two since selection is worse. Comparing this with (3.3), we find that:

$$\hat{W}(\lambda^*) - W(\lambda) = [-\pi (1 - \pi) + \pi \lambda - \pi^2 \lambda^* + [1 - \beta (1 - \pi) \pi] (\lambda^* - \lambda)] \Delta.$$

The first term represents the fact that in period one, it is now necessary to have two good incumbents to have a guarantee that the right policy is enacted for the voters. (In general, this a cost of our two agent set-up.) Thus, a necessary for a two person decision making process to be better for voter welfare is that $\lambda^* > \lambda$, i.e. it improves discipline among good agents.

As we have seen from this example, when there are two agents making policy then they need to bargain somehow to bring about a policy outcome. This outcome can change political incentives and either raise or lower discipline. This insight underpins the analysis of Persson, Roland and Tabellini (1997, 2000) who model the separation of powers in a political agency framework with moral hazard by considering bargaining between two agents. However, in their setting all agents care only about rents – there is no equivalent of the congruent agents modeled here. They suppose that all incumbents desire to extract rent from voters with their actions being only imperfectly observable. They then compare a situation in which one politician (or a single chamber of a legislature) is given authority to make all policy decisions versus one in which there are two. In the latter case, both authorities are required to reach agreement before a policy is implemented. In that sense, their set-up is quite similar to what we have proposed here.

Persson et al [1997] lay down an infinite horizon model with each period corresponding to a term in office where the incumbents may choose how much to divert to private use in each period and how much to spend on public goods valued by voters. Each period also yields an unobservable shock to the productivity of providing public services, after which an election occurs. There is an infinite number of potential identical challengers, one of whom may be selected to face the incumbent in an election. Consider first the case of single authority. This follows the analysis of Ferejohn (1986) quite closely. The
voters determine a cutoff level of public goods consumption below which the incumbent is replaced.

Now suppose that there are two authorities who must agree before any policy outcome can be implemented. Persson et al [1997] observe that unless mutual agreement is insisted upon, then voters lose. In effect, there is a common pool problem in which both incumbents get to extract rents from the voters in an uncoordinated way. Under mutual agreement, they specify a bargaining game with one player having agenda setting powers. They show that if there are no productivity shocks, then voters are as well off as they were under a single policy maker – the rents are simply shared between the two politicians.

To get a strict benefit for the voters Persson et al [1997] show that a conflict of interest between the two politicians must be created. They do so by supposing that in the first stage an authority chooses the total budget while the authority that moves second chooses the division of resources between the authorities. The first stage player now has a stronger incentive to propose something that pleases the voters – since it loses out in the second stage rent division game.

In the presence of productivity shocks, the presence of two political agents who know the level of the shock makes it possible to glean better information about such shocks. This parallels well-known results in implementation theory (such as Maskin (1999)) which suggest that better information revelation is generally possible when more than one agent have the same information about some unobservable variable. Each authority can be asked to announce the value of the shock with punishments for equilibrium disagreements. Voters now gain from having more information entering the political process as they use a more exacting re-election rule.

Persson et al [2000] is concerned with the difference between parliamentary and presidential regimes. In the latter, the executive is directly elected, whereas a parliamentary regime has the executive being subordinated to the legislature. The key issue that they examine is the allocation of resources between political rents and public goods. They develop a model of legislature in which there is a legislature comprising three districts each of which can receive a transfer. In common with political agency models the voters use retrospective voting strategies defining a cut-off level to be re-elected.

The characteristic difference between the modeling of parliamentary and presidential systems is the nature of the agenda setting power. A parliamentary system unifies agenda setting powers while a presidential system diversifies it. The latter tends to reduce rent extraction and tends towards smaller government.48

Developing multi-agent extensions of the agency model is a useful step in the direction of realism. It is also an important step in understanding political organization which frequently involves multi-agent settings. However, our understand of these issues is far

48 Testa (2002) applies similar logic to the comparison of a bicameral and a unicameral legislature and shows when the former can reduce rent extraction by politicians.
CHAPTER 3. POLITICAL AGENCY AND ACCOUNTABILITY

from complete and developing it further remains a major challenge for the future.\textsuperscript{49}

3.5 Discussion

Our study of accountability and agency is useful if it assists us in understanding better how the world works and specifically ways of improving the functioning of democratic systems. In this section, I discuss the implications of the model for two areas. First, we discuss how the model casts light on the link between civic virtue and the quality of government, following on from the influential work of Putnam (1993). Second, we discuss the implications of the models for debates about the optimal degree of decentralization in the policy sphere.

3.5.1 Civic Virtue and the Quality of Government

In his celebrated book on politics in Italy, Putnam (1993) argues that civic virtue is a key value that affects the quality of democracy. The centerpiece of the analysis is the observed correlations between measures of civic virtue and the quality of government.\textsuperscript{50} His measures of civic virtue are participation in associations, the extent to which voters vote on the basis of preference for one candidate rather than on the basis of clientelistic relationships, and newspaper readership. There is plenty of interesting discussion in the book. But it is not entirely clear how to locate Putnam’s ideas in the context of political economy models.

The agency approach gives a persuasive way of thinking about the theory that underpins these ideas. Civic virtue in general could be thought of as an index of the probability that any politician elected to office is congruent and hence will pursue what voters want. Societies with a greater commitment to civic virtue will likely have a higher quality political class and hence higher quality government. Since a higher π

\textsuperscript{49}Myerson (2000) suggests the following interesting possibility where the separation of powers is dysfunctional. Suppose that an important observable outcome of government (like the inflation rate) depends on the efforts and the skill of two different elected officials, but the voters cannot directly monitor these officials’ separate activities that influence this outcome. Suppose also that the politicians A and B who currently hold these two offices are political rivals, expected to run against each other in the near future, but A’s office is more powerful and so has more effect on the observable outcome. Then voters should assume that the observed outcome is more an indication of the skills and efforts of politician A. In this situation, politician B may have an incentive to use his power counterproductively, in the hope of generating a bad outcome that would make A look inept to the voters. If the voters anticipate such secret sabotage by B then B’s incentive to sabotage becomes even stronger, because the voters could rationally infer from a bad outcome that B is a particularly capable manipulator of government who would be a more effective leader in a higher office.

\textsuperscript{50}Putnam uses a twelve dimensional measure of government quality which he analyzes by means of principal components.
may also create an increase in private voluntarism, we might expect Putnam’s observed correlation.

This observation is a point of departure between the first generation political agency models of Barro (1973) and Ferejohn (1986) with their exclusive focus on moral hazard which cannot give expression to Putnam’s ideas. The idea that there are civic minded types is however a key theme of Putnam’s discussion which fits with the broader intellectual focus of this book.

The agency model also fits Putnam’s observation that less clientelistic voting improves the quality of government. This can be seen using the analysis in section 4.1 where we observed that accountability is likely to be enhanced when there are fewer voters with entrenched attachments to the incumbent. In this case, even dissonant politicians will likely behave in a more disciplined fashion which increases voter welfare.

His third correlation concerns newspaper readership and the quality of government. Here, we can appeal to the results in section 4.2 on information which create a presumption that improving voter information would improve political discipline and raise the quality of government.

Making democracy work in the context of the agency models means finding effective accountability that improves discipline and selection. All three factors that Putnam correlates with quality of government are factors that affect the performance of politics as an accountability system holding fixed the formal political institutions and hence can explain why the same formal institutions can yield a very different political outcome. Putnam’s way of thinking about the quality of government sits naturally with the kind of agency model that we have studied in this chapter.

3.5.2 Decentralization versus Centralization

One of the key questions in choosing the tier of government at which certain goods and services should be provided is the extent to which government can be held accountable for its actions. The recent push towards the greater use of decentralized provision is built to a significant degree on the notion that having government closer to the people will lead to better governance. Political agency models are a good framework to think about these issues.

One of the key arguments for decentralization is often that it allows citizens to exercise “voice” more effectively. The latter is often interpreted as providing better information about policies and politicians to the voters. This has been the thrust of work by Bardhan and Mookherjee (2000). Our baseline model can be used to cast light on the issues. Using the model of section 4.2, we would expect a difference between centralized and decentralized government if $\chi$ (the probability that $\Delta$ becomes known) was higher under decentralization. There are clearly policies for which this is true. Suppose that one task of government is the removal of graffiti. Then voters can see whether this has been successful in their own location, but may have a poor sense of
whether this has been done in a larger jurisdiction.

But our model also reminds us that there are other preconditions for successful decentralization. The analysis of section 4.1 suggests that decentralized government will work less well if there is more polarization and less political competition in some jurisdictions. This is a real issue for the working of decentralized governments in regions with strong ethnic conflict or areas that are dominated by one particular group. It may be that at the national level these things aggregate to give much less bias for the incumbent than they would for any local incumbent. This line of argument suggests a more equivocal answer to whether decentralization necessarily improves accountability.\footnote{Bardhan and Mookherjee (2000) argue that another key issue is the propensity of government to be captured by special interests which may vary at the national and local levels.}

Decentralization may also be important because of the way that it induces competition between governments. One particular aspect of this concerns whether the electoral mechanism works better when voters can make yardstick comparisons between jurisdictions. We explore this in greater detail in the next chapter. The basic idea behind this comes from the literature on moral hazard in teams developed in Holmstrom (1982). A group of agents are asked to perform tasks where the outcome depends on some unobserved variable which is common to all. Effort put into making the outcome successful is not observable. It then becomes optimal to condition the incentives given to one agent on the outcomes achieved by others. This has applications in a number of areas, crucial to the regulation of firms with unknown costs.\footnote{Besley and Case (1995a) and Salmon (1995) observed that this kind of mechanism could have value in the context of decentralized political competition where a number of local or regional governments are being asked to perform a similar task against a backdrop of correlated private information. For example, the cost of building roads may be similar in jurisdictions that have similar topography. The theory of yardstick political competition is studied in Besley and Case (1995a), Besley and Smart (2003), Bordignon et al (2001), Hindriks and Belleflamme (2001) and Bodenstein and Ursprung (2001). We discuss these ideas in greater detail in the next chapter.} Yardstick competition can be thought of broadly as improving the information that voters have. Hence, our analysis from section 4.2 is once again relevant in thinking about it.

The empirical relevance of these ideas was formulated in Besley and Case (1995a) which considered interdependent tax setting for U.S. States using geographically neighboring states as a benchmark. They find that there is positive co-movement in the tax rates across state boundaries. Consistent with the agency model, this positive co-movement is not found among Governors who are not eligible for re-election. Similar co-movement results for European countries are found in Bordignon \textit{et al.} (2001) (Italy), Revelli (2001) (UK), and Schaltegger and Küttel (2002) (Switzerland).

Decentralization can also be important because of the constraints placed on politicians by resource mobility. Brennan and Buchanan (1980) ventured the idea that competition for resources among governments would place a desirable constraint on the power to tax given the tendencies towards revenue maximizing Leviathan governments.
From a normative perspective this claim is puzzling since the externalities due to tax competition tend to reduce the efficiency of revenue collection and hence enhance the possibilities for welfare improving collusion among governments.

Besley and Smart (2003) consider this argument in a political agency model and their analysis is discussed in greater detail in the next chapter. They point out that an apparently second-best policy of inducing tax competition can be welfare improving for voters because it induces a favorable change in the political equilibrium. This is because the initial outcome that voters can achieve is second-best – some dissonant politicians are ill-disciplined and others are re-elected. Hence, as in the general theory of the second best, a further distortion can be welfare improving.

Hindriks and Lockwood (2005) look at centralization versus decentralization in a political agency framework similar to the one that we use in the next chapter. They model centralization as a situation in which a politician is answerable to multiple constituencies subject to a common budget constraint. They show that in a model where politicians differ in competence, then decentralization unambiguously dominates whereas when politicians differ in their honesty, the outcome is ambiguous.

Myerson (2005) also studies decentralization issues in a agency model. He argues that the value of decentralization is that local leaders have incentives to cultivate reputations before succeeding to national office. Decentralization is then worthwhile because it improves political selection at the national level.

### 3.5.3 Autocracy versus Democracy

The structure studied can provide a useful framework for thinking through some basic differences between political accountability in autocratic and democratic regimes. While autocracies do not have elections to determine who will hold office, incumbents do typically rely on the support of particular groups in society to remain in office. In principle, some aspects of the agency model could be applied to thinking about an autocracy where a decisive group (such as the military) rather than the voters have to decide whether to keep an incumbent in power at the end of period one.\(^5\) It also seems consistent with standard discussions that there is a mixture of good and bad autocrats in the world. Notwithstanding, it is perhaps reasonable to suppose that \(\pi\) is lower under autocracy because there is more limited entry of congruent candidates.

In situations where the decisive group has the broad interest of citizens at heart, the model suggests that there could be little difference between accountability under democratic and non-democratic systems of government. The distortions may come in models such as that studied in section 4.5.1 where the autocratic incumbent targets a specific group (the decisive group for keeping him in power) at the expense of voters at

\(^5\) An example of this approach to studying autocracy is Gellego and Pitchik (2004). However, their focus is on turnover rather than accountability issues.
large.

Autocracies and democracies may also differ because the decisive group is informed about the incumbent’s type and may therefore have an interest in keeping $\pi = 0$ if that is a guarantee of future rents. There are also issues of what kind of contracts between decisive voters and autocrats are feasible and whether this is broader than what we allow in stylized representations of representative democracy.

Clearly, a complete account of these issues goes well beyond the analysis here. But the basic structure of agency models is a fruitful framework for thinking through the differences. Moreover, it will likely provide an understanding of what contributes towards “making autocracy work” to paraphrase Putnam.

### 3.5.4 Accountability to Whom?

Throughout the analysis, we have assumed that accountability is to the voters. Our discussion of the last section suggested that there is no particular reason to think that the ideas apply solely in this context. One of the classic critiques of representative democracy by thinkers of the left is that some types of voters are systematically ignored. Our model of multiple issues illustrated this idea in a very limited way. One important ground for ignoring some voters is that they are ignorant about policy outcomes. If the rich and better educated are better informed then the agency model can rationalize why these groups will tend to get more politicians to make policy choices that reflect their needs.\(^{54}\)

The model can also be extended to think about the possibility that politicians gain rewards after they have left office – we have so far concentrated on rewards that accrue while in office. In the spirit of the original career concerns model, the outside option may be indexed to behavior in office. If this outside option is strong enough and rewards politicians for producing $\Delta$ for voters, then this could reinforce good behavior and may even induce politicians to produce $\Delta$ in the second period. In such cases, outside options are a complement with achieving accountability. However, one could also imagine a (less agreeable) world in which politicians are rewarded in future for their rent seeking behavior while in office, reducing discipline among dissonant politicians. This depends on whether the characteristic of being dissonant or congruent is more valuable in the market place.

Recent important work by Diermeier, Keane and Merlo (2003) has looked at the way in which political careers are valued in the market place, although they are not able to say whether the reward is indexed to behavior while in office. However, they do find an overall positive premium. The agency model leads naturally to thinking about these issues and to the occupational choices that lie behind the profession of politics. We will

\(^{54}\)See also Stromberg (2004) who gets a result like this by supposing that politicians make ex ante policy commitments to groups of voters.
3.5. DISCUSSION

return to these ideas in Chapter 5.

3.5.5 Wage Policies for Politicians

So far, we have considered only the implications of re-election rules. However, there are other dimensions of the contract that is given to politicians that are potentially important. One key example is the wage. This can be thought of in the model as affecting $E$ – the benefit of holding office. Similarly, perquisites such as free accommodation, transportation and state banquets may be part of this. This raises the question of whether there is an optimal degree of generosity in the package that is offered to politicians.

The framework that we have considered provides an ideal vehicle for exploring this.\(^{55}\) On the electoral side of the model, as $E$ changes there will be discipline and selection effects. The first order effect is to increase the discipline of incumbents. This can be seen immediately by observing that $\lambda = G(\beta [\mu + E])$ which is increasing in $E$. This is like a rather standard efficiency wage argument and was observed in Ferejohn (1986). However, in a model of heterogeneous incumbents, this will tend to lead to relatively fewer congruent incumbents being sorted in to serve during the second term that an incumbent is in office. As observed by Gersbach and Liessem (2001a), there is a case for setting wages that depend on the seniority of the politician with higher wages in the second period to induce better behavior by incumbents. The conclusion that higher wages always improves political discipline is however fragile. This can be seen from the analysis of section 4.3 where we observed that timid behavior by congruent politicians becomes a possibility for high enough $E$.\(^{56}\)

The analysis in section 4.6.1 suggests that wages may also affect the quality of the pool of politicians. Recall that the equilibrium fraction of congruent politicians as a function of $E$ that we derived there was:

$$\pi (E) = \frac{\gamma}{\gamma + (1 - \gamma) \frac{w(E; \mu)}{E} \cdot \frac{W_c}{W_d}}.$$  

It is straightforward to see that this is increasing in $E$ so long as $W_c \geq W_d$. This is because standing for political office is more attractive on the margin to the dissonant politicians. It is interesting to note though that if we had added some sufficient public service motivation from holding public office to the congruent politicians this conclusion would be reversed.

Another issue that we have not considered is the possibility of designing incentive contracts that index the pay of politicians explicitly to their performance while in office. This is considered by Gersbach (2001) and Gersbach (2003). While this is an interesting

\(^{55}\)The discussion here is based on Besley (2004). Gersbach (2001), Gersbach and Liessem (2001a) and Gersbach (2003)) all discuss aspects of incentive contracts in political life.

\(^{56}\)This point is made in Smart and Sturm (2003b).
idea, we do not seem to see such schemes in practice. One reason is surely the fact that, by focusing attention on specific measurable goals, agents will tend to distort their efforts away from socially valuable tasks whose outputs are difficult to measure (Gersbach and Liessem (2001a)).

3.5.6 Behavioral versus Rational Choice Models

The agency model that we have studied here has rational politicians and rational voters – the latter who use Bayes rule to process information. There is some debate in the political science literature about how far the rationality postulate should be pushed in modeling political behavior. Our baseline model is consistent with voters using a simple behavioral rule where politicians who do something good for them get re-elected. This happens in this case to be consistent with Bayes rule. Thus, there is no difference in the baseline between a rational model of voters and sensible behavioral model.

More generally, it is interesting to postulate the implications of a behavioral model of accountability in which voters favor politicians who generate some utility for them during their term in office. This will not in general yield any form of “optimal” accountability of the kind studied here. However, some fraction of the time it will stand the voters in good stead. Going forward it would be interesting to understand better what the differences are between behavioral models of politics and the postulates of the strict rationality supposed here. It would be useful to understand when simple and sensible behavioral rules lead to large policy distortions.

3.6 Concluding Comments

Political agency models extend insights from contract theory and information economics to political situations. Their main focus is on how voters hold politicians to account for their actions. In this chapter, we have explored the insights that they offer and discussed a number of applications from the literature. This should give the reader some notion of their potential in explaining real situations. The models also put some structure on the data, in particular with regard to the term-limit effect.

As well as illustrating theoretical issues, we have also developed an empirical application to U.S. governors. The striking thing about that context is how well the basic framework applies. Governors are directly elected chief executives who are directly accountable. Given the murkiness of situations where chief executives are accountable to legislatures or parties, it is much harder to formulate the model’s implications. However, doing so constitutes a major challenge for future empirical applications in this area.

But how does all this relate back to the grand themes of chapter one? The agency problem is able to formulate the idea of policy in the interests of voters (the case where the voter gets $\Delta$) from that run by individuals with their own agendas. This is a core
idea at the heart of the political economy literature. However, the approach for the most part is simplistic. In a multi-issue world congruence is difficult to achieve and even to define. Thus democracy giving voters what they want is complex. As we discussed in the last chapter, then the achievements of democracy need to be judged according to some agreed upon social criterion.

Once this is recognized, the normative significance of congruence becomes problematic. It could well be that the role of politicians is to take unpopular actions for the social good even if they do not result in re-election. This is especially true when voters want inconsistent things. Politicians who simply follow voters interests may actually be worse for society. Of course, the dangers of populism have always been recognized and this is really only an instance of such concerns in an agency framework.

In the next chapter, I will develop an approach to political agency which is more specific in its ambitions. First, it is applied to public finance questions which means developing a more specific model. Second, we structure the model in a way that makes the normative issues more transparent.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governor won election</td>
<td>0.57</td>
<td>0.50</td>
</tr>
<tr>
<td>Vote share (percent)</td>
<td>57.5</td>
<td>8.73</td>
</tr>
<tr>
<td>Two year growth rate of real taxes per capita (%)</td>
<td>7.15</td>
<td>10.28</td>
</tr>
<tr>
<td>Two year growth rate of state income per capita (%)</td>
<td>14.31</td>
<td>6.36</td>
</tr>
<tr>
<td>Two year growth rate of state spending per capita (%)</td>
<td>6.77</td>
<td>9.56</td>
</tr>
<tr>
<td>Log of population</td>
<td>14.65</td>
<td>1.07</td>
</tr>
<tr>
<td>Governor’s age (years)</td>
<td>52.06</td>
<td>7.95</td>
</tr>
<tr>
<td>Governor is lawyer</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Years of work experience before Governorship</td>
<td>28.61</td>
<td>9.58</td>
</tr>
<tr>
<td>Fraction of previous experience in politics</td>
<td>0.35</td>
<td>0.24</td>
</tr>
<tr>
<td>Governor’s years of education</td>
<td>18.77</td>
<td>3.52</td>
</tr>
</tbody>
</table>

Notes: Data sources are as in Besley and Case (2003). Data is only for the 48 continental United States. The time period is 1950-96. The Governor’s characteristics are from the Book of the States.
Table 3.2: Accountability

<table>
<thead>
<tr>
<th></th>
<th>(1) Governor re-elected</th>
<th>(2) Governor re-elected</th>
<th>(3) Governor re-elected</th>
<th>(4) Governor re-elected</th>
</tr>
</thead>
<tbody>
<tr>
<td>growth in real taxes per capita</td>
<td>-0.704</td>
<td>-0.734</td>
<td>-0.932</td>
<td>-0.873</td>
</tr>
<tr>
<td></td>
<td>(2.49)*</td>
<td>(2.29)*</td>
<td>(3.22)**</td>
<td>(2.76)**</td>
</tr>
<tr>
<td>growth in real income per capita</td>
<td>1.808</td>
<td>2.501</td>
<td>1.475</td>
<td>2.350</td>
</tr>
<tr>
<td></td>
<td>(3.05)**</td>
<td>(4.73)**</td>
<td>(2.54)*</td>
<td>(4.82)**</td>
</tr>
<tr>
<td>growth in real expenditure per capita</td>
<td>0.132</td>
<td>-0.013</td>
<td>-0.035</td>
<td>-0.258</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.03)</td>
<td>(0.10)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>log of state population</td>
<td>-0.001</td>
<td>0.230</td>
<td>0.025</td>
<td>0.241</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(1.43)</td>
<td>(0.15)</td>
<td>(1.53)</td>
</tr>
<tr>
<td>Vote share in last election</td>
<td>0.004</td>
<td>0.010</td>
<td>-0.001</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(2.87)**</td>
<td>(0.17)</td>
<td>(2.09)*</td>
</tr>
<tr>
<td>Governor's age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.017</td>
<td>-0.013</td>
<td>(5.08)**</td>
<td>(2.77)**</td>
</tr>
<tr>
<td>Governor is trained as a lawyer</td>
<td>0.021</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of work experience before Governorship</td>
<td>0.018</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction of previous experience in politics</td>
<td>0.636</td>
<td>0.775</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5.58)**</td>
<td>(3.95)**</td>
</tr>
<tr>
<td>Years of education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.003</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.983</td>
<td>-3.131</td>
<td>-1.856</td>
<td>-4.186</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(1.31)</td>
<td>(0.87)</td>
<td>(1.76)</td>
</tr>
<tr>
<td>Observations</td>
<td>485</td>
<td>381</td>
<td>475</td>
<td>372</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.17</td>
<td>0.26</td>
<td>0.31</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Notes: Robust t statistics in parentheses (* denotes significant at 5%; ** denotes significant at 1%) Data sources are given in notes to Table 3.1.
<table>
<thead>
<tr>
<th></th>
<th>(1) % vote captured by the winner</th>
<th>(2) % vote captured by the winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>growth in real taxes per capita</td>
<td>-13.288 (2.50)*</td>
<td>-11.901 (2.18)*</td>
</tr>
<tr>
<td>growth in real income per capita</td>
<td>9.452 (1.10)</td>
<td>7.275 (0.82)</td>
</tr>
<tr>
<td>growth in real expenditure per capita</td>
<td>4.945 (0.85)</td>
<td>5.068 (0.83)</td>
</tr>
<tr>
<td>log of state population</td>
<td>-0.126 (0.28)</td>
<td>-0.175 (0.36)</td>
</tr>
<tr>
<td>Vote share in last election</td>
<td>0.432 (4.94)**</td>
<td>0.424 (4.84)**</td>
</tr>
<tr>
<td>Governor's age</td>
<td>-0.110 (0.66)</td>
<td></td>
</tr>
<tr>
<td>Governor is trained as a lawyer</td>
<td>1.592 (1.18)</td>
<td></td>
</tr>
<tr>
<td>Years of experience before governorship</td>
<td>-0.010 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Fraction of experience in politics</td>
<td>2.479 (0.97)</td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>0.147 (0.44)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>36.291 (3.98)**</td>
<td>38.904 (3.12)**</td>
</tr>
<tr>
<td>Observations</td>
<td>268</td>
<td>261</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.18</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Notes: Robust t statistics in parentheses (* denotes significant at 5%; ** denotes significant at 1%) Data sources are given in notes to Table 3.1.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real government spending per capita (thousands $1982)</td>
<td>1.04</td>
<td>0.50</td>
</tr>
<tr>
<td>Real government taxes per capita (thousands $1982)</td>
<td>0.49</td>
<td>0.25</td>
</tr>
<tr>
<td>Real government sales taxes per capita (thousands $1982)</td>
<td>0.31</td>
<td>0.14</td>
</tr>
<tr>
<td>Real government income taxes per capita (thousands $1982)</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Real government corporate taxes per capita (thousands $1982)</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Congruence (ADA) [-100 to 0]</td>
<td>-22.09</td>
<td>11.42</td>
</tr>
<tr>
<td>Congruence (COPE) [-100 to 0]</td>
<td>-25.22</td>
<td>11.96</td>
</tr>
<tr>
<td>Governor cannot run due to term limit</td>
<td>0.29</td>
<td>0.45</td>
</tr>
<tr>
<td>Log state income per-capita</td>
<td>16.63</td>
<td>1.56</td>
</tr>
<tr>
<td>Log state population</td>
<td>14.83</td>
<td>1.01</td>
</tr>
<tr>
<td>Proportion of population aged over 65 (percent)</td>
<td>10.71</td>
<td>1.01</td>
</tr>
<tr>
<td>Proportion of population aged 5-17 (percent)</td>
<td>22.41</td>
<td>3.46</td>
</tr>
<tr>
<td>Governor is a Democrat</td>
<td>0.56</td>
<td>0.50</td>
</tr>
<tr>
<td>Democrats are a majority in the state Senate</td>
<td>0.65</td>
<td>0.48</td>
</tr>
<tr>
<td>Democrats are a majority in the state House</td>
<td>0.67</td>
<td>0.47</td>
</tr>
<tr>
<td>Divided Government</td>
<td>0.38</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Notes: Sources as in Besley and Case (2003). Data is for the 48 continental states for the period 1950-2000. Nebraska is excluded from political variables as it does not have a two party system.
<table>
<thead>
<tr>
<th>Term-Limit Effects</th>
<th>(1) real government spending per capita ($1982)</th>
<th>(2) total taxes per capita ($1982)</th>
<th>(3) sales taxes per capita</th>
<th>(4) income taxes per capita</th>
<th>(5) Corporate taxes per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governor Cannot Run</td>
<td>0.034 (4.45)**</td>
<td>0.090 (1.81)</td>
<td>0.030 (0.83)</td>
<td>0.116 (3.35)**</td>
<td>0.028 (2.76)**</td>
</tr>
<tr>
<td>log of real income per capita ($1982)</td>
<td>-0.244 (4.53)**</td>
<td>1.015 (2.59)**</td>
<td>1.522 (5.52)**</td>
<td>-0.579 (1.80)</td>
<td>-0.142 (1.91)</td>
</tr>
<tr>
<td>log of state population</td>
<td>-0.047 (0.84)</td>
<td>-1.570 (3.80)**</td>
<td>-0.675 (2.05)*</td>
<td>0.184 (0.56)</td>
<td>-0.021 (0.26)</td>
</tr>
<tr>
<td>% population aged 65 and above</td>
<td>-0.851 (1.97)*</td>
<td>6.167 (2.39)*</td>
<td>9.202 (4.63)**</td>
<td>0.155 (0.06)</td>
<td>0.492 (0.93)</td>
</tr>
<tr>
<td>% population aged 17 and below</td>
<td>-0.571 (1.68)</td>
<td>6.063 (2.65)**</td>
<td>3.328 (2.20)*</td>
<td>7.241 (3.86)**</td>
<td>-0.051 (0.13)</td>
</tr>
<tr>
<td>Governor is a Democrat</td>
<td>0.020 (3.36)**</td>
<td>0.037 (1.03)</td>
<td>0.033 (1.33)</td>
<td>0.060 (2.06)*</td>
<td>-0.000 (0.06)</td>
</tr>
<tr>
<td>Democrats control Senate</td>
<td>0.032 (3.78)**</td>
<td>0.299 (5.26)**</td>
<td>0.099 (2.15)*</td>
<td>0.159 (3.30)**</td>
<td>0.021 (1.46)</td>
</tr>
<tr>
<td>Democrats control House</td>
<td>0.004 (0.39)</td>
<td>0.202 (3.39)**</td>
<td>0.049 (1.08)</td>
<td>0.103 (2.19)*</td>
<td>0.032 (2.23)*</td>
</tr>
<tr>
<td>Divided Government</td>
<td>-0.000 (0.03)</td>
<td>-0.103 (2.68)**</td>
<td>-0.039 (1.47)</td>
<td>0.030 (1.00)</td>
<td>-0.032 (3.72)**</td>
</tr>
<tr>
<td>Observations</td>
<td>2162</td>
<td>2203</td>
<td>2210</td>
<td>1749</td>
<td>1810</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.95</td>
<td>0.91</td>
<td>0.88</td>
<td>0.87</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Notes: Robust t statistics in parentheses (* denotes significant at 5%; ** denotes significant at 1%) Data sources are given in notes to Table 3.4. Data is for 47 states for period 1950-1997.
Table 3.6: Congruence and Term-Limits

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruence – ADA</td>
<td>1.173</td>
<td>2.383</td>
</tr>
<tr>
<td></td>
<td>(2.63)**</td>
<td>(4.40)**</td>
</tr>
<tr>
<td>log of real income per</td>
<td>-29.049</td>
<td>-22.964</td>
</tr>
<tr>
<td>capita ($1982)</td>
<td>(7.60)**</td>
<td>(4.90)**</td>
</tr>
<tr>
<td>log of state population</td>
<td>12.958</td>
<td>4.569</td>
</tr>
<tr>
<td></td>
<td>(2.88)**</td>
<td>(0.84)</td>
</tr>
<tr>
<td>% population aged 65 and above</td>
<td>-92.096</td>
<td>-139.090</td>
</tr>
<tr>
<td></td>
<td>(3.62)**</td>
<td>(4.14)**</td>
</tr>
<tr>
<td>% population aged 17 and below</td>
<td>-32.204</td>
<td>-7.249</td>
</tr>
<tr>
<td></td>
<td>(1.20)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Governor is a Democrat</td>
<td>1.651</td>
<td>2.104</td>
</tr>
<tr>
<td></td>
<td>(4.68)**</td>
<td>(4.78)**</td>
</tr>
<tr>
<td>Democrats control Senate</td>
<td>1.034</td>
<td>-0.818</td>
</tr>
<tr>
<td></td>
<td>(1.93)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>Democrats control House</td>
<td>-0.113</td>
<td>0.969</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Divided Government</td>
<td>-3.001</td>
<td>-3.499</td>
</tr>
<tr>
<td></td>
<td>(8.19)**</td>
<td>(7.84)**</td>
</tr>
<tr>
<td>Constant</td>
<td>343.609</td>
<td>360.278</td>
</tr>
<tr>
<td></td>
<td>(10.23)**</td>
<td>(8.41)**</td>
</tr>
<tr>
<td>Observations</td>
<td>1632</td>
<td>1632</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.72</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Notes: Robust t statistics in parentheses (* denotes significant at 5%; ** denotes significant at 1%) Data sources are given in notes to Table 3.4. Data is for 47 states for period 1950-1997.
Table 3.7: Correlations Between Print Freedom, Corruption and Political Longevity

<table>
<thead>
<tr>
<th></th>
<th>Print Freedom</th>
<th>Corruption (ICRG)</th>
<th>Corruption (Transparency International)</th>
<th>Corruption (Kaufmann et al)</th>
<th>Years in Office (Party)</th>
<th>Years in Office (Chief Executive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print Freedom</td>
<td>1.00</td>
<td>-0.7034</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>187</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corruption (ICRG)</td>
<td>-0.6974</td>
<td>0.8261</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>74</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corruption (Transparency International)</td>
<td>-0.6750</td>
<td>0.8253</td>
<td>0.9688</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>153</td>
<td>91</td>
<td>156</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years in Office (Party)</td>
<td>-0.2656</td>
<td>0.2990</td>
<td>0.1095</td>
<td>0.0822</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.008)</td>
<td>(0.3246)</td>
<td>(0.357)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>148</td>
<td>78</td>
<td>83</td>
<td>148</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Years in Office (Chief Executive)</td>
<td>-0.3466</td>
<td>0.1963</td>
<td>0.1630</td>
<td>0.1407</td>
<td>0.5018</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.062)</td>
<td>(0.127)</td>
<td>(0.086)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>172</td>
<td>91</td>
<td>89</td>
<td>150</td>
<td>148</td>
<td>172</td>
</tr>
</tbody>
</table>

Notes: Source is Besley and Prat (2004). p-values in parentheses. Third row is number of observations. Bold face denotes significant at a 10% level.
Chapter 4

Political Agency and Public Finance

(with Michael Smart)

“To make the argument for constitutional constraints, it is necessary, first of all, to derive the elementary logic of limits or constraints on the exercise of free or voluntary choice, whether at the level of an individual or a collectivity.” (Buchanan (1999 page 111)).

4.1 Introduction

The literature on the political economy of public finance falls into two main camps. On the one hand is the literature which uses the median voter model as a means of studying how the conflict of interest between different citizens plays out in the political process. This has been used to provide insights into the equilibrium level of the income tax and/or public goods. The main task is to view the political process as an aggregation mechanism – determining a set of social decisions from underlying preferences.

The other main tradition in political economy and public finance recognizes the conflict of interest between governors and the governed, recognizing the possibility that there is a limited ability to restrain the power of non-benevolent politicians. The political agency model discussed so far and the Leviathan model of Brennan and Buchanan (1980) are squarely in this tradition. This chapter tries to pull together some thinking on this theme using a simple model of public finance with an agency model of politics.¹

The objectives of the chapter are twofold. First, we seek to provide a textbook level treatment of public finance issues that illuminates many of the ideas of the last chapter in a more concrete policy setting. Second, we will use the model to address some substantive policy concerns. The first is whether there is a case for imposing restraints

¹See Hettich and Winer (2004) for an overview of approaches to the positive economics of taxation.
on government over and above the electoral mechanism. The second is whether there is an advantage in bringing non-governmental organizations to substitute for public provision of goods and services.

The political agency model is a useful starting point for the former exercise as it provides a rigorous underpinning for the view that elections (while important) can serve only a limited role in restraining politicians’ behavior. Thus, it is important to think about whether additional constraints above the need to be re-elected are a good idea.

Debates on restraining government bring into sharp relief the core theme of this book. If government were benevolent then restraints in a fiscal setting would be welfare reducing for voters as they would distort the policy process. Thus, the essence of the reason for considering fiscal restraints has to be that government is imperfect. But to provide a precise account of how restraints will work will inevitably require a model of how well government performs.

One response to imperfections in government resource allocation is to rely more on private organizations – non-governmental organizations (NGOs) – to substitute for government provision. These issues are widely debated. However, there is no generally agreed framework for thinking about this. One context where this has been brought into sharp relief is aid flows to developing countries. The political agency approach is ideally suited to thinking about these issues as it models the capacity of government to deliver services. It is also important to tackle head on the problems of accountability that are likely to arise in NGOs. We extend the model to allow for this.

The analysis developed here builds on a very simple model put forward in Besley and Smart (2003). A government must choose how much to tax and how to spend. Spending can be either on public goods which are valuable to voters or personal perks which we label generically as “rent”. We assume that rents are unobservable to voters. In principle, voters should be able to determine them once they know the level of spending and taxes. However, in practice this is difficult for a host of reasons. We capture this idea here in a very simple way by supposing that the “cost” of providing the kind of public spending that voters like is not observable. Thus taxes and spending levels are an imperfect means of detecting rent extraction.

In the spirit of chapter three, types of politicians may also vary between rent-seeking politicians and good politicians who serve the public interest. The role of the electoral process is both to sort in the right types and to discipline the rent seeking politicians if they are elected to office. The agency model developed here studies the role of electoral incentives in affecting public finance decisions – the overall level of taxation and the balance of spending between rent and public spending. We also allow for the government to raise revenues by issuing public debt.

The model has two time periods for policy making with an election occurring be-

\[\text{2Our good politicians are similar to the congruent politicians in the last chapter and our bad politicians are like the dissonant politicians.}\]
tween the two. Voters make inferences based on observed policy outcomes and use this information to update their assessments using Bayes rule. Incumbents who displease voters by extracting excessive rents are removed from office.

The setting sometimes yields some surprising insights. For example, some restraints on government are likely to be good for voters only when government is relatively benevolent. This is because these restraints work by improving selection which is most valuable when there is preponderance of good politicians in the population.

Many of the key findings in this chapter rest on the political theory of the second best. This was first articulated in Fischer and Summers (1989). This is an exact extension of the economic theory of the second best due to Lipsey and Lancaster (1956). When agency problems generate distortions in political behavior, then there is no guarantee that standard reasoning about how good government should be organized is valid. Measures that would not be sensible in a first best world of benevolent government earn a rationale. When it comes to instruments that restrain government, examples of the political theory of the second best abound.

The chapter is organized as follows. In the next section, we lay out the basic agency problem and the main assumptions. In section three, we go through the three basic kinds of agency model. The first has purely hidden type, the second only hidden action and the third both features. In section four, we draw out some implications of the model for the data – for example, patterns of re-election rates and the impact of term limits. These facts are consistent with the U.S. data.

In section five, we use the model to discuss a variety of means for restraining government. This allows us to re-examine the theoretical foundations of some ideas that have been central to the public choice approach and to see to what extent they hold up in an explicit agency model. Specifically, it considers a variety of proposals to restrain government – explicit restraints on taxation, tax competition, increased transparency, and yardstick competition. Modeling the micro-foundations of the political agency problem suggests that the issues can be quite subtle. At the heart of the agency problem are two main effects: discipline and selection. Greater discipline here means reduced levels of rent extraction. Changing the environment in which government operates can affect this. Selection refers to the information revealed to the voters to allow them to throw bad politicians out of office. There is a fundamental trade-off between these two – greater discipline makes it harder to tell who the bad guys are. The trade-off that this creates is important to appraising the effect of greater restraints on government.

The model is then extended to include public debt. If public debt is observable, then the model and its insights remain unchanged. However, difficulties in observing public debt levels do create a rather different spin on the model. This is discussed in section six. In section seven, we apply the ideas to whether an NGO is a more efficient provider of public spending. In section eight, we look at what happens if its is fiscal competence that is not observed. Section nine concludes.
Chapter 4. Political Agency and Public Finance

4.2 The Model

Government and the economy. There are two periods. In each period $t$, the government must decide on three things – the level of "valuable" public spending $G_t$, the level of public spending devoted to private ends $s_t$ and the level of taxation $x_t$. We assume that there is a maximal feasible level of government tax collections in any period; thus $x_t \in [0, X]$.

In each period, the unit cost $\theta$ of providing a public good is determined randomly by nature. The cost of the public good is independently and identically distributed in each period, with $\theta \in \{L, H\}$, $H > L$, and $\Pr(\theta = H) = q$. We denote the realized value in period $t$ by $\theta_t$.

For now we assume that tax collections in each period are constrained to equal total government spending. (We consider government debt in Section 6.) The government budget constraint in each period $t$ is

$$x_t = \theta_t G_t + s_t.$$ 

The representative voter derives utility from public goods, net of the cost of government spending. When the politician provides public goods in the amount $G$ and total spending is $x$, the welfare of the voter is $W(G, x) = G - \mu C(x)$, where $C$ is a strictly convex, increasing function and $\mu$ is an exogenous parameter that indexes the marginal cost of public funds in a simple way.

Optimal Policy. We take as a benchmark the pattern of taxing and spending that would be picked by a benevolent government which cared only about the voter welfare. In this case there will be no diversion of public resources to private ends, i.e. $s_t = 0$ for $t = 1, 2$. First define

$$G^*_\theta(\mu) = \arg \max \{G - \mu C(\theta G)\}$$ (4.1)

and let $x^*_\theta = \theta G^*_\theta$ with $W^g(\theta, \mu)$ being the level of voter welfare at these optimal levels. Finally, let

$$EW^g(\mu) = qW^g(H, \mu) + (1 - q) W^g(L, \mu).$$

We assume that $X > x^*_L$. The fact that $C' (\cdot)$ is increasing implies that $x^*_L > x^*_H$, i.e. government is optimally larger when the cost of public goods is lower. It is clear that $G^*_\theta(\mu)$ and $W^g(\mu)$ are decreasing in $\mu$ – citizens are better off when it is less costly to provide public goods.

Politics. We study an agency model of elections with two time periods; in each, the politician in office makes decisions about taxation, debt and government spending. Between periods, there is an election in which a voter chooses between the incumbent and
a challenger. The cost shock, $\theta$, and the amount of rent diverted for private purposes, $s$, is privately observed by the politician.

Politicians may be one of two types, “good” or “bad”. Thus we label the politician’s type by $i \in \{b, g\}$. A good politician simply chooses $G$ and $x$ in each period to maximize voter welfare, and places no value on rents diverted from public spending. In other words, he will pick the optimal policy bundle outlined in the previous sub-section.

In contrast, a bad politician behaves strategically, choosing policies to maximize the expected, discounted sum of rents $s_1 + \beta\sigma s_2$ extracted from government, where $\beta < 1$ is a discount factor, and $\sigma$ is the probability of re-election to second term. The determination of the re-election rule is discussed below.

The types $i \in \{g, b\}$ of first-period incumbent and challenger are independent draws from an identical distribution with $\Pr(i = g) = \pi$. The incumbent then observes the first-period cost shock $\theta_1$ and chooses public goods provision $G_1$ and rent diversion $s_1$. The voter then observes $G_1$ and government spending $x_1$ prior to the election at the end of the first period. However, the types of incumbent and challenger, the unit cost $\theta_1$, and the level of rent diversion $s_1$ are unobserved. In the second period, the politician then in office again chooses $G_2$, $x_2$ and $s_2$ given $\theta_2$. There are no further elections; thus even newly elected challengers are “lame ducks” in the second period. In casting their votes in the election, therefore, the voter must make an inference about the incumbent’s type based on observed performance and compare it to prior beliefs about the type of the challenger.

The interesting issues in the model pertain to incumbent behavior in period one and the determinants of electoral success. In the two period setting, there are no incentive issues to consider in period two.

The Leviathan model of government behavior of Brennan and Buchanan (1980) can be thought of as a special case of this model where $\pi = 0$, i.e. all politicians are known to be bad. In their model, as in this one, there are two kinds of public spending – rents to politicians and spending that voters value. Brennan and Buchanan (1980) assumes that these are produced in fixed proportions from any given tax revenue. However, in this model incumbents can choose how they allocate tax revenues between these ends.

### 4.3 Three Scenarios

The model can be used to develop three scenarios for the study of political agency problems. In the first, we suppose that incumbents cannot control the amount of the rent that they extract – bad ones take rent and good ones do not. The only problem for voters is then to pick them out. In the second case, we study moral hazard problems. All politicians are bad and re-election is a reward for limiting rent extraction. In the third case, we study the combined effect of adverse selection and moral hazard. We study the possibility that the bad incumbents mimic the good ones to gain re-election.
This has short run advantages in curbing rent seeking but long run disadvantages in reducing selection.

The model with adverse selection and moral hazard provides some bridge between the full blown public choice model and the benevolent government model. The only thing that prevents the benevolent outcome is, therefore, the inability of voters to identify who will behave benevolently (the selection problem).

4.3.1 Pure Adverse selection

We begin with pure adverse selection. This is defined by a case in which politicians have no ability to influence the amount of rents that they take. Good politicians deliver the outcomes preferred by voters while bad ones extract some pre-determined amount. We assume that $G$ is fixed at some commonly known level $\overline{G}$. While this model is not at all realistic, it will serve to initiate a discussion of selection problems that will be useful in the ensuing discussion and analysis.

If the levels of public spending chosen by the different types of politicians perfectly revealed their type, then pure adverse selection would be uninteresting. Hence, we assume that a bad politician extracts rents equal to $s = (H - L) \overline{G}$. Obviously this is a very specific assumption. However, it is the simplest way, in a two state model, for there to be one state of the world in which voters are unsure whether their incumbent is good or bad after they have seen the policy outcome. If they observe that public spending is $HG$, they are uncertain whether it is a good incumbent who has received a cost shock of $H$ or a bad one who has received a cost shock of $L$.

We can now compute the voting rule used by voters according to Bayes rule. It is clear that since $G$ is fixed at $\overline{G}$, voting depends exclusively on the level of taxation. Given our assumptions, $x \in \{LG, HG, (2H - L)G\}$. Two of these tax levels perfectly reveal the incumbent’s type. An incumbent that chooses $LG$ must be good and one that chooses $(2H - L)G$ must be bad. The key question is what inference the voters make when they see $HG$. Using Bayes rule, the probability that the politician is good conditional on seeing $HG$ is

$$\frac{q\pi}{q\pi + (1 - q)(1 - \pi)}.$$

To decide whether or not to re-elect the incumbent, they compare this with $\pi$, the probability that a randomly selected challenger is good.

It is easy to check that this is greater than $\pi$ if $q > \frac{1}{2}$. Then re-election is a good idea – it is more likely that the state is $H$ than otherwise and the incumbent is given the benefit of the doubt. If $q < \frac{1}{2}$, then it is always a good idea to remove an incumbent.

---

$^3$As we will see below, this level of rent has significance even in a model where rents and public spending can be chosen endogenously.
who chooses $H G$ even though this will lead to some good incumbents being removed from office. We summarize this as:

**Proposition 8** Under pure adverse selection, there are two possibilities:

(i) If $q \geq \frac{1}{2}$, then politicians are re-elected if they set taxes of $L G$ or $H G$ – all good politicians and some bad politicians are re-elected,

(ii) If $q < \frac{1}{2}$, then politicians are re-elected only if they set taxes of $L G$ – all bad politicians are thrown out of office along with some good ones.

A couple of features of this outcome are worth noting. First, voter behavior does not depend on the distribution of politicians in the population but purely on the likelihood of the state in which good and bad politicians are confused arises – specifically the likelihood ratio $\frac{q}{(1-q)}$. The model also provides a simple justification for voters being more likely to remove incumbents from office when they charge higher tax rates – which is consistent with both folk wisdom and empirical evidence. We discuss this in greater detail below.

It is clear that under adverse selection voters benefit from having better quality incumbent pools (higher $\pi$) and more information about the true state of the world on which they condition their voting decisions. The pure adverse selection model yields some stark insights. However, the idea that incumbents cannot make real choices which affect voters’ perception of them is too strong. We now consider models which relax this.

### 4.3.2 Pure Moral Hazard

As we discussed in the last chapter, the case of pure moral hazard was the subject of the “first generation” political agency models of Barro (1973) and Ferejohn (1986). We now show how this case can be studied in the framework introduced here.

Suppose then that the type of every politician is known to be bad so that they care primarily about rent extraction with voter welfare as a secondary concern. We allow both $G$ and $s$ to be chosen and assume that only $G$ and $x$ are observable. Voters need make no inference on the incumbent’s type. The role of voting is purely to achieve some level of discipline – rewarding politicians for delivering something closer to what voters would like (in this context lower rent seeking). We therefore consider a game in which the voter may announce – and commit to – an incentive-based re-election rule prior to the incumbent’s first period policy being chosen. Let the re-election rule used by voters be denoted by a function $\sigma(G, x)$. Then for fixed $G$, the incumbent’s rent extraction problem is to choose rent diversion in periods one and two denoted by $(s_1, s_2)$ to maximize $s_1 + \beta \sigma(G, \theta G + s_1) s_2$.

In period two, there are no re-election incentives to restrain rent seeking and every politician will choose $s_2 = X$. In period one, the incumbent can be rewarded for reducing
his rent seeking. Voters reason as follows. Since incumbents can always extract $X$ today and forego re-election, then they must be offered an incentive to prevent them from doing so. Since all they care about are rents, they are indifferent between all packages which yield them $X$. Hence, consider the threshold level of rent, $\tilde{s}(\sigma)$ which satisfies

$$\tilde{s}(\sigma) + \sigma \beta X = X.$$ 

As long as the level of "permissible" rent satisfies this, the incumbent will choose $\tilde{s}(\sigma)$ and get re-elected with probability $\sigma$. However, since the level of period one rent seeking that satisfies this incentive constraint is decreasing in $\sigma$, it is best for voters to set $\sigma = 1$. In the language of the principal–agent literature, we may therefore think of the condition $s_1 \geq (1 - \beta) X$ as the "participation constraint" that must be satisfied to insure that the incumbent produces the level of public spending desired by voters for each cost level $\theta = L, H$.

The pure moral hazard model is consistent with a term limit effect in politicians’ choices of the kind found by Besley and Case (1995b). However, given the choice in reality between a challenger who can serve an extra term in office and an incumbent who will be a lame duck, the model cannot explain why the voters would ever prefer a lame duck.

The model predicts that rent taking is lower by $\beta X$ when politicians face re-election incentives. The more patient are politicians, the larger the rent improvement for voters. The level of public spending is higher in the first period than the second and the level of taxation is $X$ in period two and less than $X$ in period one.

Even though information about $\theta$ is imperfect, it is revealed to voters in equilibrium via the choice of $G$. This is because politicians do not care intrinsically about $\theta$ – it does not affect the amount of rent that they can extract.

The model also makes voters indifferent between electing any politician in period two. There are no types in the model that make voters intrinsically prefer one kind of politician over another. This is because the model makes the most pessimistic assumptions about the political class – all are out to serve themselves.

Finally, even the pure moral hazard model is more optimistic about the prospects for electoral competition compared with the Leviathan model of Brennan and Buchanan (1980). The promise of future office can restrain rent seeking.

We now wish to determine the levels of spending and taxes $(G_\theta, x_\theta)$ that the voter will require from the incumbent in exchange for re-election. This is a "mechanism design" problem, which involves some technicalities not relevant to most of the material in this chapter. We therefore provide an informal discussion of the problem here, and a more complete treatment in an Appendix to the chapter.

Consider first the case of the high cost level. The incumbent is willing to supply any

\footnote{In this section, we ignore mixed strategy mechanisms by voters and incumbents.}
4.3. THREE SCENARIOS

fiscal policy \((G_H, x_H)\) that satisfies the participation constraint

\[ s_H = x_H - H G_H = (1 - \beta) X \]

If the cost level were known to the voter, the problem would then be merely to maximize voter welfare \(G_H - \mu C(x_H)\) subject to this constraint. This “first-best” choice is depicted in Figure 1, as the point of tangency \(A_1\) between the participation constraint for type \(H\), which has slope \(1/H\), and an indifference curve for the voter, with a tax level \(x_H^*\).\(^5\) (Observe that voter welfare is increasing in the “northwesterly” direction in the diagram, while the politician’s rents are increasing in the “southeasterly” direction.)

Figure 1 about here

Figure 1 also shows the difficulty with this outcome: If true cost is unobserved by the voter, an incumbent who faces cost \(\theta = L\) must be deterred from implementing the high-cost fiscal package \((G_H, x_H)\) and obtaining additional rents equal to the true cost difference \((H - L) G_H\). In other words, fiscal policies must also respect an “incentive constraint” for the low-cost type

\[ s_L \geq (1 - \beta) X + (H - L) G_H \]

In Figure 1, the indifference curves for type \(L\) are the straight lines with steeper slope \(1/L\). If the first-best spending point \(A_1\) for \(H\) were chosen, then the best policy for \(L\) is the point of tangency labelled \(B_1\), with associated level of spending \(x_L\). The \(L\) incumbent is indifferent between \(A_1\) and \(B_1\), but the latter gives the voter strictly higher welfare.

But the voter can do strictly better than this outcome. Consider a small reduction in spending along the participation constraint for \(H\), to the point labelled \(A_2\). Since \(G_H\) is now smaller, so is the level of rents that \(L\) may obtain. The incentive constraint for \(L\) is relaxed, and the voter can insist on higher spending from type \(L\), yielding higher welfare (the point labelled \(B_2\)).

This model therefore suggests why agency problems may lead to smaller government. Even if all politicians are known to have bad incentives, restrictions on government spending impose limits on the rents politicians can extract through “cost padding”.

The formal solution to this problem, given in the Appendix to this chapter, yields a number of other insights about the optimal levels of spending. In particular, we find:

**Proposition 9** The voter’s preferred incentive compatible spending level \(x_H\) is smaller:

---

\(^5\) As voter welfare is quasilinear with respect to \(G\), the equilibrium tax level is the same regardless of the equilibrium level of rent extraction.
1. when \( L \) is smaller; and

2. when the voter’s beliefs about costs are more optimistic (\( q \) is lower).

Thus the proposition states, somewhat counter-intuitively, that a reduction in true costs induces a reduction in government spending. Thus, reductions in the cost of government services may actually lead to further restrictions on government spending, as they worsen the potential agency problems created by bad politicians.

### 4.3.3 Combining Moral Hazard and Adverse Selection

As in chapter three, we now have two roles for elections – (i) selecting good politicians as in the adverse selection case and (ii) providing incentives to reduce rent-seeking as in the moral hazard case. The model defines a game of incomplete information between the incumbent politician and the representative voter. We seek to characterize perfect Bayesian equilibria of this game. As usual, the game is most easily solved by applying a type of backward induction. In the second period, the politician in office faces no further possibility of electoral discipline. Thus \( s_2 = X \) for \( i = b \) (bad politicians take maximal rents) while \( s_2 = 0 \) for \( i = g \).

Given that second-period strategies are identical for challenger and incumbent, the sequentially rational voting rule for the voter is to re-elect the incumbent if the posterior probability the incumbent is the good type exceeds the prior probability \( \pi \) that the challenger is good. The voter’s posterior beliefs depend in turn on the equilibrium strategy of the first-period incumbent. Since the good type cares only about voter welfare in the current period, he chooses \((G_H^*, x_H^*)\) with probability \( q \) and \((G_L^*, x_L^*)\) with probability \( 1 - q \). It follows that, in any perfect Bayesian equilibrium, the voter’s posterior beliefs assign probability zero to the good type at any other information set \((G;x)\). To economize on notation, we therefore write posterior beliefs as a function \( \Pr(g|x) \) of first-period spending alone.\(^6\)

We confine attention to three possible strategies for the bad incumbent, each associated with one of the three spending levels \((x_L^*, x_H^*, X)\) that are observed with positive probability along the equilibrium path. First, \( b \) might choose \( s_1 = 0 \) or \( s_1 = X \). Since future rents are discounted \((\beta < 1)\), however, the latter strategy dominates the former.\(^7\) Thus:

\[
\Pr(g|x_L^*) = 1
\]

\(^6\)The voters’ beliefs are not restricted by Bayes’ rule at nodes not reached in equilibrium. Since the good type’s actions are pinned down by our preference assumption, we impose the minimal restriction on out-of-equilibrium beliefs that \( \Pr(g|x) = 0 \) if \((G, x) \neq (G_{\theta}^*, x_{\theta}^*)\) for \( \theta \in \{L,H\} \). At any such information set, the voter would elect the challenger in the second period.

\(^7\)Regardless of true costs, \( s_1 = 0 \) yields a payoff to the bad type of \( \beta X \) if re-elected and zero otherwise, while \( s_1 = X \) pays \((1 + \beta)X \) if re-elected and \( X \) otherwise.
in any equilibrium, and the voter always re-elects when first-period spending is \( x_L^* \). Beliefs conditional on observing \( x_H^* \) are more complicated. A bad politician who faces low true costs may, instead of taking maximal rents, choose to produce \( G_H^* \) units of the public good and spend \( x_H^* \), diverting \( \hat{s}(\mu) \equiv (H - L)G_H^*(\mu) \) to private rent consumption. This strategy allows type \((b, L)\) to “pool” with type \((g, H)\), and doing so may be desirable, if it brings a positive probability of re-election. Accordingly, let

\[
\lambda = \Pr(x = x_H^*|\theta = L, i = b)
\]

denote the probability type \((b, L)\) exercises restraint in this sense, and let \( \sigma \) denote the probability of re-election when the voter observes \( x_H^* \). The posterior probability that spending \( x_H^* \) was generated by a good politician is

\[
\Pr(g|x_H) = \frac{\pi q}{\pi q + (1 - \pi)(1 - q)\lambda}
\]

A best response for the voter is to re-elect with positive probability \((\sigma > 0)\) only if \( \Pr(g|x_H) \geq \pi \) or, equivalently, \( \lambda \leq q/(1 - q) \). If the inequality is strict, then \( \sigma = 1 \). Further, type \((b, L)\) prefers to exercise restraint instead of diverting maximal rents \((\lambda > 0)\), only if \( \hat{s}(\mu) + \beta \sigma X \geq X \). When this inequality is strict, then \( \lambda = 1 \).

Collecting these observations, there are three possible equilibrium configurations.

First, equilibrium may be pooling, as type \((b, L)\) chooses \( s_1 = \hat{s}(\mu) \) and so is indistinguishable from type \((g, H)\). Second, equilibrium may be separating, as type \((b, L)\) chooses \( s_1 = X \) and is revealed \textit{ex post}. Third, equilibrium may be a hybrid one, in which type \((b, L)\) adopts a strictly mixed strategy on actions \( s_1 = \hat{s}(\mu) \) and \( s_1 = X \), so that type is revealed with positive probability strictly less than one. The following result fully characterizes the possible configurations.

**Proposition 10** (Besley and Smart, 2003) An equilibrium exists for all values of parameters and is generically unique.

1. A pooling equilibrium, with \( \lambda = \sigma = 1 \), exists if and only if

\[
q \geq \frac{1}{2} \quad \& \quad \hat{s}(\mu) \geq (1 - \beta)X
\]

2. A hybrid equilibrium, with \( \lambda = q/(1 - q) \) and \( \sigma = (X - \hat{s}(\mu))/(\beta X) \), exists if and only if

\[
q < \frac{1}{2} \quad \& \quad \hat{s}(\mu) \geq (1 - \beta)X
\]

3. A separating equilibrium, with \( \lambda = 0 \) and \( \sigma = 1 \), exists if and only if

\[
\hat{s}(\mu) \leq (1 - \beta)X
\]
For a formal proof of this, the reader is referred to Besley and Smart (2003). Thus, there are three possible kinds of equilibria. In the separating equilibrium, the bad politician takes maximal rents and is detected with certainty and replaced by the challenger. This equilibrium outcome is therefore equivalent to that which would be obtained if voters could observe the cost shock directly. In the pooling and hybrid equilibria, the incumbent is taking less than maximal rents and his type is revealed with lower probability.\(^8\) The latter is most likely when \(\tilde{s}(\mu)\) is high and the incumbent discounts the future heavily so that he prizes rents earned in period one.

The result has features of the adverse selection and moral hazard models. Like the adverse selection model, there is some possibility that a bad incumbent is re-elected. In the pooling and separating equilibria, and unlike pure adverse selection, there is no chance that a good incumbent is removed from office – it is always optimal to re-elect any incumbent who chooses \(x_H^I\). In line with the moral hazard model, the incumbent also restrains his rent seeking endogenously to get re-elected. However, this is incomplete – when the cost shock is \(H\) there is no restraint. This arises because (unlike the pure moral hazard moral) the voters are not indifferent between a bad incumbent and a challenger because of the existence of good politicians in the pool of potential politicians. The model also predicts that, even if they exercise restraint, bad incumbents choose a higher level of rents than under pure moral hazard when they face the cost shock \(H\).\(^9\)

### 4.4 Implications

In this section, we draw out some positive and normative implications of the equilibrium described in Proposition 3.

We begin with some normative implications. We first look at how selection and incentive effects combine to determine voter welfare. We also look at how the model predicts the case for government intervention in the economy is circumscribed by the agency costs implied the political equilibrium. We also look at the implications of having only a few good politicians available to voters and whether this can be worse than having only bad politicians.

For positive implications, we look first at how the model links taxes, spending and political turnover. We then look at what the agency model predicts about the existence of fiscal cycles.

---

\(^8\) In the pooling equilibrium, the outcome is similar to that assumed in the Leviathan model where rents are assumed to be positively related to the amount of “legitimate” public spending.

\(^9\) For bad incumbents with cost shock \(L\), the level of rents may not be higher - it depends on how much the high-cost fiscal policy is distorted under the pure moral hazard model.
4.4. IMPLICATIONS

4.4.1 Equilibrium Voter Welfare

To understand how the electoral process affects political decision-making in the model, it is useful to calculate expected voter welfare in equilibrium. As a benchmark for the analysis, suppose that politicians are removed from office each period with certainty. Expected voter welfare when type $g$ is in office would then be

$$EW^g(\mu) = qW^g(H, \mu) + (1-q)W^g(L, \mu),$$

whereas welfare with $b$ in office would simply be $W^b(\mu) = -\mu C(X)$, since “lame-duck” bad politicians divert maximal rents. To simplify notation, let $W^0(\mu) = \pi EW^g(\mu) + (1-\pi)W^b(\mu)$ be the unconditional expected per-period welfare in this case. It follows that the present value of expected welfare is just $(1 + \beta)W^0(\mu)$ when there is no chance of re-election.

Equilibrium welfare in the equilibrium described in Proposition 3 can be written as

$$EW(\lambda, \sigma, \mu) = (1 + \beta)W^0(\mu) + (1-\pi)(1-q)\lambda \Delta(\mu) + \beta(\pi_2 - \pi)\Sigma(\mu)$$

where:

$$\Delta(\mu) = W^g(H, \mu) - W^b(\mu)$$

$$\Sigma(\mu) = EW^g(\mu) - W^b(\mu)$$

and

$$\pi_2 - \pi = \pi(1 - \pi)[q\sigma + (1-q)(1-\sigma\lambda)]$$

Equation (4.2) has a simple interpretation, as the last two terms represent the deviation from welfare in the absence of re-election possibilities. The term $\Delta(\mu)$ is the discipline effect induced by re-election possibilities – the fact that bad incumbents restrain rent seeking – while $\Sigma(\mu)$ represents selection – the fact that elections improve the quality of period two policy by weeding out bad incumbents. Multiplying each effect is the probability that the relevant events occur and these benefits of elections are reaped.

In summary, the possibility of re-electing incumbents can increase voter welfare both by improving average quality of office-holders (the “selection effect”) and by offering prospective incentives for low-quality incumbents (the “discipline effect”). To understand the impact of changes in the fiscal regime, the effects on selection and incentives in elections must be understood. This in turn requires understanding how equilibrium political behavior (as represented by $\lambda$ and $\sigma$) is affected.

Differentiating (4.2) with respect to $\lambda$ yields:

$$\frac{\partial EW(\lambda, \sigma, \mu)}{\partial \lambda} = (1-q)(1-\pi)[\Delta(\mu) - \beta\pi\sigma\Sigma(\mu)].$$

---

10This is similar to the expression $\Pi - \pi$ in chapter 3.

11Some care is needed in interpreting this heuristic argument. As we have seen, $\lambda$ is an equilibrium outcome. To have a larger $\lambda$, we need to have a lower $X$, a higher $\beta$, a lower $\mu$, a higher $H$, and a
CHAPTER 4. POLITICAL AGENCY AND PUBLIC FINANCE

The sign of this depends on the magnitude of the discipline effect ($\Delta (\mu)$) and the selection effect ($\Sigma (\mu)$). A higher $\lambda$ (more discipline) is good for voters if $\Delta (\mu) > \beta \pi \Sigma (\mu)$.

The formula in (4.2) also gives a precise characterization of agency costs in government. This gives a criterion for government to intervene in the economy which is more conservative than that which is studied by the standard public economics model. As a point of comparison, we take an extreme case where the government does not intervene at all and sets $G = x = 0$. Below, in section five, we consider more subtle possibilities that restrain government rather than closing it down completely.

A purely benevolent view of government would see the case for intervention in terms of optimal citizen welfare under government intervention. In this case, government is worthwhile if and only if $(1 + \beta) EW^g(\mu) > 0$. A welfare economist operating in a world of benevolent government would advocate this case for government.

The agency model suggests that this is too optimistic. However, the identification of specific agency costs makes this more conservative case for government more precise. For the sake of illustration suppose that the equilibrium being played is separating so that $\sigma = 1$ and $\lambda = 0$. The reader will find it straightforward to work through other cases. Then, welfare in the political equilibrium is:

$$ (1 + \beta) EW^g(\mu) - [EW^g(\mu) - W^b(\mu)] (1 - \pi) (1 + \beta (1 - \pi)) \tag{4.7} $$

The second term in this expression (which is proportional to $(1 - \pi)$) represents the agency cost of running the government. The cost is higher, the greater is the difference between having a good and bad politician in office. This gives an exact sense of why a political economy approach is less sanguine about the case for government. The criterion $(1 + \beta) EW^g(\mu) > 0$ is too optimistic.

The above expression gives precise content to the Public Choice critique of welfare economics in this context.\footnote{See Besley and Coate (2003) for a model of the public choice critique.} This way of thinking is reminiscent of the analysis of Acemoglu and Verdier (2000) which looks at how imperfections in the policy process can militate against government intervention. It gives sense to the idea that failures of government (here due to the existence of rent-seeking politicians and imperfect information) should be weighed up against the positive case for government. This could be thought of in terms of the market failure that underpins the case for having government spending on $G$ in the first place.

The expression in (4.7) also shows how the model that we have laid out here can be thought of as lying between the pure public choice view and the standard benevolent government view. If $\pi$ tends to one, there are only benevolent politicians and the agency cost goes to zero leaving us with the standard welfare economic criterion for lower $L$. But a change in these parameters also affects the voter welfare. For example, if $\lambda$ goes up by lowering $X$, then the condition for more discipline being better will be weaker than $\Delta (\mu) > \beta \pi \Sigma (\mu)$ because a reduction in $X$ also mitigates the negative effect of re-electing a pooling bad incumbent.
government. As \( \pi \) gets close to zero, then the expression gets close to \( (1 + \beta) W^b(\mu) = -(1 + \beta) \mu C(X) < 0 \) and it is not worthwhile to have the government. Thus, whether having the government function depends in a discernible way on the size of agency costs and on whether government operates closer to the benevolent or public choice model.

### 4.4.2 Are Good Politicians Necessarily Good for Voters?

Comparing the models in sections 3.2 and 3.3, it is interesting to contrast the outcomes in which there are a few good politicians with those that would obtain were all politicians bad. We will observe in this case that things could actually be better if all politicians were bad. To this end, we will contrast welfare in the pure moral hazard case and that where there is some fraction of good politicians.

In both cases, all bad politicians will extract maximal period two rents. However, if there is any chance that a good politician will be found, then having some good politicians is clearly better for voters. Now consider period one. The bad politician under pure moral hazard takes rents of \( (1 - \beta) X \) for sure in period one. Suppose that the equilibrium in the combined moral hazard/ adverse selection model is separating so that \( \lambda = 0 \) and \( \sigma = 1 \). Then the period one rent will be \( X \) if there is a bad incumbent and zero if there is a good one. It is now easy to see that the pure moral hazard case can be better.

Voter welfare under pure moral hazard is:

\[
q \left[ \frac{x_H^H}{H} - \mu C(x_H) \right] + (1 - q) W^g_L - \frac{\delta}{H} - (1 - q) \frac{H - L}{HL} x_H + \beta E W^b
\]

where \( \psi = q/H + (1 - q)/L \), while under moral hazard and adverse selection, it is (rewriting (4.7)):

\[
(1 + \beta) W^b(\mu) + \pi (1 + \beta (2 - \pi)) \left[ EW^g(\mu) - W^b(\mu) \right].
\]

in the case of a separating equilibrium. It is now easy to see that, for small enough \( \pi \), welfare is higher under pure moral hazard. It is also clear that as \( \pi \) tends to one, welfare is unambiguously higher when there is both moral hazard and adverse selection.\(^{13}\)

This may seem paradoxical. However, the reasoning behind it is clear after a little thought. Once there are good politicians in the population, they create a benchmark against which bad politicians are measured when they seek re-election. Proposition 3

\(^{13}\)In the case of a pooling equilibrium, the above expression becomes

\[
(1 + \beta) W^b(\mu) + \pi (1 + \beta + (1 - \pi) q) \left[ EW^g(\mu) - W^b(\mu) \right] + (1 - \pi)(1 - q)[W^g(H, \mu) - W^b(\mu)].
\]

So for small enough \( \pi \), welfare under pure moral hazard is very likely to be higher because \( W^g(H, \mu) \) is smaller than \( EW^g(\mu) \) and because \( \psi (1 - \beta) X \) is small due to the condition for a pooling equilibrium. So the same logic basically applies when we compare to a pooling equilibrium.
tells is that it is worthwhile to mimic the good types when there is sufficient rent for doing so available to the bad types. Otherwise, they would prefer to take $X$. Thus, the good types create an externality which affects the bad type’s behavior. However if bad types are not very numerous, they have a negligible direct effect on the voters’ payoffs. Thus, voters are worse off if there are just a few good types in the population.

Another way to look at this in terms of the voters’ ability to commit. With pure moral hazard, voters are indifferent (ex post) between voting for the incumbent and a randomly selected challenger. Hence, they can pick a standard for incumbents to meet that creates the best possible incentives for incumbents to reduce their rent-extraction. Voters cannot commit to this voting rule, however, when the incumbent could be a good type even though they would prefer to commit to the voting rule that is used under moral hazard.

This finding suggests that the pure moral hazard model is rather fragile to a small variation in the model to include some good types of politicians. This is because the strict indifference rule that underpins incentives in that case allows the voters to commit. Once this indifference is broken, then there is actually a constraint on optimal voting strategies which can make things worse.

### 4.4.3 Turnover of Politicians

As we discussed in Chapter 3, one of the distinctive features of the political agency model is how it links turnover rates among politicians and government policy. Given the particular economic concerns being studied here, this refers to links between taxes, public spending and voting behavior. The model can underpin the empirical prediction that higher taxes are bad for re-election chances and can also look at the link between public spending and political turnover.

Using Proposition 3, the model predicts that the retention rate of a randomly selected first period politician is:

$$f(\lambda, \sigma) = 1 - [\pi q (1 - \sigma) + (1 - \pi) (q + (1 - q) (1 - \lambda \sigma))]$$

It is straightforward to verify that the retention rate is increasing in $\lambda$ and $\sigma$.

Empirically, the interesting issue is not how turnover relates to $\lambda$ and $\sigma$ since these equilibrium strategies are not observable. Instead, the data come in the form of observations about $x$ and $G$ and how these relate to whether politicians are re-elected. To tease out these empirical implications turns out to be a little complicated since the probability of re-election depends on two variables: total taxation and the amount of valuable public spending undertaken. To summarize the result implied by Proposition 3:

$$\sigma(x_L^*, G_L) = 1, \sigma(x_H^*, G_H) \in [0, 1], \text{ and } \sigma(X, 0) = 0. \quad (4.8)$$
4.4. IMPLICATIONS

There is one clear-cut prediction – a politician who provides more valuable public spending is always good for re-election chances. This is rather obvious as higher public spending is always brought forth by good incumbents. In practice, however, testing this would require being able to separate expenditure items into their $G$ and $s$ components which cannot easily be done.

Thus, the more interesting empirical prediction links total taxation with re-election prospects. However, as is plain from (4.8), there is a potential non-monotonicity in the relationship between taxes and re-election if $\sigma(x_H^*, G_H^*) < 1$ since $x_H^* < x_L^*$. This is because the state in which good politicians differentiate themselves from bad ones is when the cost of providing public goods is low ($\theta = L$). The low spending state can be consistent with bad politicians extracting some rent and masquerading as good politicians.

In either a pooling or separating equilibrium, retention probability is constant for low levels of spending – incumbents are re-elected for sure when $x = x_L^*$ or $x_H^*$. However, for all levels of taxation above $x_L^*$, the incumbent is removed from office.

In a hybrid equilibrium, $\sigma(x_H^*, G_H^*) = (S - s(\mu))/\beta S < 1$. In this case, there is a non-monotonicity and while high tax incumbents are still thrown out for sure, the middle range (those who pick $x_L^*$) are retained more often than those who pick low total taxes $x_H^*$.

This model gives a more direct underpinning for the empirical results on U.S. Governors that we discussed in Chapter 3. There we looked at patterns of taxation and spending and the way in which they drove re-election chances and were subject to term limit effects. The results on re-election as a function of taxes are consistent with this model.

4.4.4 The Spending Cycle

Political agency models predict a spending cycle depending on whether or not politicians are able to stand for re-election. In the current model this comes out only in the comparison between first and second period taxation decisions. It is straightforward to study this in the model. We will show that the sign of the electoral cycle – comparing two incumbents who can and cannot run again – is ambiguous. It depends on the balance of the discipline and selection effects.

Consider, first of all, the difference between the tax policy of second period (term-limited) and first period (non term-limited) incumbent. In a separating equilibrium, expected period one expenditures are:

$$Ex_1 = \pi (qx_H^* + (1 - q)x_L^*) + (1 - \pi) X$$

while expected period two taxes are

$$Ex_2 = (\pi + (1 - \pi) \pi) (qx_H^* + (1 - q)x_L^*) + (1 - \pi)^2 X.$$
Comparing the two yields:

\[ Ex_1 - Ex_2 = (1 - \pi) \pi (X - (q x_h^* + (1 - q) x_L^*)) > 0. \]

The effect of selection is to reduce the number of bad incumbents and hence to reduce the average level of taxation between the first and second periods of office.

Now we look at what happens in the case of pooling. Using the result in Proposition 3, we have that period one expected taxes are:

\[ Ex_1 = \pi (q x_h^* + (1 - q) x_L^*) + (1 - \pi) [(1 - q) x_H^* + q X]. \]

In period two, expected taxes are:

\[ Ex_2 = \pi [1 + (1 - \pi) q] (q x_h^* + (1 - q) x_L^*) + (1 - \pi) (1 - q \pi) X. \]

For low enough \( x_h^* \) (assuming that the conditions for the equilibrium to be pooling continue to hold), then it is possible to have period one expected taxes being below period two expected taxes.\(^{14}\) In effect, this is because the gain from having a bad politician exercise discipline in period one is large. This is consistent with the results of Besley and Case (1995b) and the empirical results that we presented in Chapter 3 – Governors who face a binding term limit seem to spend and tax more.\(^{15}\)

Even though the sign of the electoral taxation cycle is ambiguous and hence an empirical question, its existence is not. It is central to the agency model that political agents will behave differently depending on whether they face the prospect of being re-elected.

### 4.5 Restraining Government

Advocating methods to restrain government as a means to raising voter welfare is a staple prescription of the public choice approach to public finance. It has been hugely influential in the policy sphere as well as in the classroom. In a nutshell, the idea is that voters are unable to control the taxing and spending proclivities of incumbent politicians and hence would benefit by developing institutions and constitutional rules that will restrain government further in addition to elections.

In this section, we explore these arguments in the framework developed above. We will exploit the fact that our political economy model has micro foundations to see how various parameters that affect the workings of the political process impact on policy outcomes.

\(^{14}\) It is also straightforward to show that as \( x_h^* \) gets close to \( x_L^* \), period one expected taxes are higher than those in period two as the selection effect now dominates.

\(^{15}\) Taken literally, this empirical result would suggest that, in US states, the equilibrium is not the separating one implying that \( \hat{s}(\mu) > (1 - \beta)X \).
4.5. RESTRAINING GOVERNMENT

One general lesson that emerges is that incentives should be thought about in a second-best framework in which the core incentive mechanism is the (highly imperfect) electoral mechanism. In cases where this has no power in achieving improved discipline or better selection of politicians, then the extra-electoral restraints advocated by public choice economists are the only sources of restraint and it is fairly easy to see their implications. At the other extreme, if there are sufficient institutions to completely determine the outcome that voters want, then further incremental constraints have no value. This analysis operates in the middle ground where electoral incentives have some incomplete ability to discipline and select incumbents. Any extra restraints must then be viewed in this kind of second-best context and can have more subtle effects.

We divide our discussion of restraints on government into two parts. We first consider direct restraints – imposing limits on taxing and spending. We then consider a set of more indirect restraints which might serve to make politics work better. We specifically look at measures to increase the cost of taxation. In practice, these come in a variety of forms. One of them is by introducing restrictions on tax bases and tax rates. In the U.S., these have often been implemented though direct democracy, i.e. citizens’ initiatives. Increasing the intensity of tax competition is also a means of raising the marginal cost of public funds. A second class of indirect measures work through increasing the amount of information available to voters. There are two main possible sources of this: improvements in budgetary transparency and greater use of yardstick comparisons of politicians.

Measures to restrain the size of government are important in practice. For example, many states in the U.S. have passed restrictions on taxing and spending. These fall into three broad categories: (i) indexed limits on the growth of revenues or expenditures, for example, to the population growth rate; (ii) requirements that voters approve all new taxes; and (iii) supermajority requirements that require anywhere between three-fifths and three-quarters of the legislature to approve tax increases. There are twenty-four states with indexed limits, thirteen of which allow an override with a supermajority vote. Five of these states require a simple majority if the governor has declared a state of emergency. Half the states with indexed restrict the growth in state expenditures to the growth rate in personal income averaged over some previous period. Five others restrict the size of appropriations to a specified percentage of state income, while four others restrict growth to an index of population growth and inflation. Three other states restrict the absolute expenditure growth rate. Spending on capital projects is excluded, as are federally funded projects. Half of the limits in place are constitutional, with the

---

16 There is a huge literature on the economics of tax competition surveyed in Wilson (1999). The main focus has been on implications of increased mobility of goods and services for tax setting by governments. The literature shows that allowing competitive determination of taxes will lead to externalities that skews taxes towards immobile factors such as labor and away from mobile factors such as capital. These negative externalities between competing governments lead to an increase in the marginal cost of public funds.
remainder being statutory.\textsuperscript{17} Most tax and expenditure limitations were introduced in the 1970s, which many believe reflected a general disillusionment with government and a view that spending was out of tune with what a majority of voters preferred. While the institutional detail may be complex, the thrust of such rules is to diminish the discretion that politicians enjoy.

Many tax and expenditure limitations were the result of citizen power exercised through citizens’ initiatives. The latter are constitutional rules that allow citizens to place legislation directly on the ballot. At the present time, twenty-three U.S. states permit an initiative process whereby citizens can place ballot propositions, which are voted on subsequently. Tax and expenditure initiatives are a large fraction of those that are brought.\textsuperscript{18} Perhaps the best known example is Proposition 13, passed in California in 1978, which restricted property tax increases for current residents.

The theoretical approach to initiatives suggested by Denzau, Mackay and Weaver (1981) is explicitly premised on the idea that they are a discipline device on politicians who do not act as faithful servants of the people. Thus, they stem from a view that the electoral process is insufficient to provide a check on such behavior. There is a good deal of empirical analysis in this spirit. Zax (1989) investigates how access to initiatives affects state expenditures per capita, in a cross-section of fifty states for 1980. Contrary to the idea that initiatives promote smaller government, he finds that state spending is significantly higher in states that permit direct statutory initiatives.\textsuperscript{19} Farnham (1990) estimates the cross-sectional effect of citizens’ initiatives and referenda using data on 735 communities, taking the log of community expenditures as the dependent variable, finding little or no evidence that access to the initiative is important. Matsusaka (1995) regresses government expenditures and revenues in 49 states (Alaska is excluded) on a number of control variables for a panel of states sampled over a 30 year period at five year intervals from 1960 to 1990. His main finding is a strong negative effect on expenditures of access to the initiative. He also finds some evidence that the effect is strongest where the signature requirement to hold an initiative is low.\textsuperscript{20} This is in tune with the idea that initiatives may have a role to play in reducing agency problems that result in the state being too large.

These findings are consistent with empirical evidence of Knight (2000) and Rueben (2000) which show that tax and expenditure limitations tend to reduce government spending with the presence of citizens’ initiatives being used as an instrument to control for the potential endogeneity of such limitations.

\textsuperscript{17}See Rueben (2000) for a useful overview of the history and content of limitations in category (i) above.

\textsuperscript{18}For existing theoretical treatments of initiatives, see Gerber (1996) and Besley and Coate (2001).

\textsuperscript{19}This may, of course, be suggestive of reverse causality - states with higher spending also tend to have citizens’ initiatives.

\textsuperscript{20}To table an initiatives requires a minimum number of voters to support it. A low threshold is a situation where this minimum number is low making it easier to get an initiative on the ballot.
Apart from work on the United States, there is extensive work on the consequences of direct democracy in Switzerland. Feld and Kirchgassner (2000) provides an excellent overview of studies that use the cross-Canton variation in direct democracy to identify its effects. In similar spirit to the U.S. findings, they argue that the weight of studies find that public expenditure and public debt are lower in Cantons that make greater use of direct democracy. However, unlike the evidence above, initiatives are not used to place direct restrictions on politicians. They argue that the mechanism at work is through increasing the supply of information that voters have about policy issues. Thus, their interpretation of the role of initiatives fits better with the discussion of information provision to voters which we discuss below.

However, finding that limitations reduce spending allows us to draw no welfare conclusions. The effect of such limitations may drive out public spending that voters like rather than reducing rent seeking. Indeed this is bound to be the case when all politicians are good. Hence, the balanced likelihood of good politicians in office will likely be a key factor in determining whether such limitations are valuable.

### 4.5.1 A Direct Restraint on the Size of Government

Suppose that the government creates a constitutional restriction on the size of government which lowers $X$, the maximum tax level that the government can levy. This has the direct advantage of reducing the rent that a bad politician can extract and hence improves voter welfare (assuming that the limitation does not distort the behavior of good politicians). Thus, its direct consequences for voters is favorable. Thus a tax limitation that leaves the behavior of the good politicians unaltered and does not change the political equilibrium is welfare improving for the voters.\(^{21}\)

However, to understand its impact fully, we also need to understand how it affects the political equilibrium. This depends on the balance of discipline and selection effects. Proposition 3 makes clear why lowering $X$ increases the incentive of incumbents to pool, and so to make the political equilibrium less informative. This suggests that a tax limitation will be attractive when selection is less important than discipline. Referring to equation (4.6), this happens when $\beta \pi \Sigma (\mu) < \Delta (\mu).^{22}$ Thus, a tax limitation is more likely to be desirable when there is a predominance of self-interested politicians – when $\pi$ is small. This is in line with the traditional Public Choice view about the value of tax limitations.

Thus, we have:

**Proposition 11** *(Besley and Smart (2003))* Suppose that a limit is imposed on the size

\(^{21}\)Here we consider only the case where the maximum spending level $X$ remains high enough to leave the behavior of good politicians unchanged.

\(^{22}\)As pointed out in our discussion of equation (6), this condition is not exact. A change in $X$ also affects $\Sigma (\mu)$ and $\Delta (\mu)$ through $W^b (\mu) = -\mu C (X)$. 
of government (as measured by $X$). Then there exists a $\hat{\pi}$ such that voter welfare increases if $\pi < \hat{\pi}$.

For high values of $\pi$, the result may go the other way if it shifts the political equilibrium towards a pooling outcome, since the importance of selection may dominate.

### 4.5.2 Indirect Restraints

We now discuss other ways of restraining government that work less directly. We begin by studying the argument that allowing taxes to be collected in a more inefficient manner can actually be good for voters. This argument has been around in various guises. For example, Becker and Mulligan (2003) consider how changing the efficiency of the tax system alters the resources devoted to political influence. If greater deadweight cost in the tax system leads to reductions in influence activities that more than offset the direct costs to taxpayers then it can be good from a societal point of view. Krusell, Quadrini and Rios-Rull (1996) argue that income taxation can be attractive relative to consumption taxation even when the latter has less deadweight loss since it leads to a lower level of transfer activity in equilibrium.

Raising the marginal cost of public finance could work by constitutional restrictions on the tax bases that can be used. For example, many of the U.S. states did not traditionally allow corporate and personal income taxes. They could also come from specific restrictions on particular taxes implemented via citizens initiatives. Tax competition is also seen as a way of making it more difficult for governments to raise taxes, and hence as a break on government. This was suggested, for example, by Brennan and Buchanan (1980). Finally, they could come from technological progress in tax collection technologies.\(^{23}\)

#### Increasing the Cost of Taxation

In terms of the model, we consider the effect of an increase in the marginal cost of public goods $\mu$. We are interested in when this leads to an increase in voter welfare. It is clear that rent diversion in a pooling or hybrid equilibrium is decreasing in $\mu$.\(^{24}\)

By reducing the level of spending by benevolent governments, increasing inefficiency in the tax system restricts the amount of wasteful spending that can be undertaken by self-interested officials without fear of detection. Thus, increased inefficiency in the tax system (as represented by an increase in $\mu$) that leaves the political equilibrium unchanged reduces voter welfare, even if it reduces rent diversion by bad politicians. This result holds in spite of the fact that increasing inefficiency of the tax system does (sometimes) lower

\(^{23}\)Peltzman (1980) argues that technological differences in the ability to collect taxes have been important historically and in comparing developed and developing countries.

\(^{24}\)Observe that $\hat{s}(\mu) = (H - L)G^*_H(\mu)$ is decreasing in $\mu$. 
rent extraction by bad incumbents. This fails to deliver a benefit to voters in any of the equilibria described in Proposition 3.

When the equilibrium is separating, rent seeking is maximal anyway and voters would prefer to be “robbed” efficiently. Inefficiency in the tax system only increases the costs of venality—even when, as in the pooling or hybrid equilibrium, greater tax inefficiency leads to reduced rent seeking. To see this, observe that voter welfare in period one in the pooling/hybrid equilibrium is:

\[
\pi EW^g(\mu) + (1 - \pi)[(q + (1 - q)(1 - \lambda))W^b(\mu) + (1 - q)\lambda W^g(H, \mu)].
\]

As the political equilibrium is unchanged, this expression falls with \(\mu\). The key to understanding the logic is to consider the case where bad incumbents imitate good ones. Although the rent extraction goes down, voter welfare itself does not depend on the amount of rent extracted because the level of welfare is the same as when the good incumbents face a high cost shock, i.e. \(W^g(H, \mu)\), which is decreasing with \(\mu\).

The key assumption in the above argument is that the political equilibrium remains unchanged. Returning to Proposition 3, it is clear that increasing inefficiencies in the tax system can lead to a move from a pooling or hybrid equilibrium to a separating equilibrium. Specifically, defining \(\mu\) from \(\hat{s}(\mu) = (1 - \beta)X\), then for all \(\mu > \mu_c\) there will be separation between the good and bad incumbents. This occurs because equilibrium rents are proportional to the size of government in the pooling or hybrid cases. As the size of government is reduced by greater inefficiency in the tax system, rent extraction possibilities are limited, making it more likely that a bad incumbent will “go for broke” and extract maximal rents. In this instance, all equilibrium information (about \(\theta\) and the type of the politician) is revealed in equilibrium. Hence, to ascertain the welfare effects of an increase in the cost of tax collection which increases \(\mu\) above \(\mu_c\), we need to compare full information welfare with that in the equilibrium with \(\mu < \mu_c\).

Intuitively, the impact on voter welfare of a move to separation involves a trade-off between the short-run costs of reduced discipline, and the long-run benefits that result when bad politicians reveal type and are removed from office. Computing welfare for the two cases from (4.2) shows that the selection effect outweighs the discipline effect, so that welfare is higher in the separating equilibrium only if \(\beta \pi \Sigma(\mu) \geq \Delta(\mu)\).

Besley and Smart (2003) derive the following result on the welfare effect of tax inefficiencies:

**Proposition 12** (Besley and Smart (2003)) There exists a \(\pi^* > 0\) such that increasing the inefficiency of the tax system (as represented by an increase in \(\mu\)) unambiguously reduces voter welfare for all \(\pi < \pi^*\). For \(\pi \geq \pi^*\), an increase in the inefficiency of the tax system which moves the equilibrium from hybrid or pooling to separating may increase voter welfare if it induces a shift from a hybrid or pooling equilibrium towards a separating equilibrium.
This result says that an increase in the inefficiency of the tax system can enhance voter welfare only if it leads to an increase in the ability of the voter to detect bad incumbents. This leads to some rather paradoxical implications. First, a sufficiently large increase in the cost of public funds may indeed increase equilibrium welfare, but only if the change increases the amount of wasteful spending in the first period. Voter welfare will be higher from improved selection when the fraction of good types in the population $\pi$ is sufficiently high. In other words, raising the marginal cost of funds can increase voter welfare only when the political process is closer to the benevolent government paradigm.

Information Provision

We now discuss means of restraining government by generating better information for voters.

**Fiscal Transparency** Transparency is fast becoming the motherhood and apple pie of good governance. However, the emerging literature on transparency in government is often quite vague about what exactly is the essence of transparent government. There is actually a long history of discussions related to this. For example, John Adams, the first Vice President and second President of the United States, argued that “(l)iberty cannot be preserved without a general knowledge among the people who have a right ... and a desire to know... the characters and conduct of their rulers.”

In practice, there is a huge amount of variation in the world in the way that governments allow their citizens to observe the workings of government and the basis of social decisions. Polities also differ in the extent to which there are developed private institutions such as the media and independent watchdogs that are effective sources of commentary and information about the quality of policies and politicians. There seems little doubt that in most developed polities, the rise of mass media has increased the scrutiny to which most governments’ fiscal decisions are exposed. However, there remain many countries around the world whose media systems remain closed and repressed.

Given the concerns of this chapter, our main interest is how transparency plays out in the fiscal process. One widely cited definition of transparency in a public finance setting is that of Kopits and Craig (1998) who suggest that “[F]iscal transparency is defined ... as openness toward the public at large about government structure and functions, fiscal policy intentions, public sector accounts, and projections” (page 1). This definition is quite broad – applying to a whole array of fiscal institutions.

However, when transparency is defined in a specific framework or model, it is often required to be more precise about which variables are being varied. Within public

---

finance, one of the primary concerns has been in the accounting procedures used by government. This comes to the fore in many discussions of public deficits. In particular it is possible for the citizens to observe the true fiscal stance being taken by a government from the reporting procedures which it has. This is the focus of contributions by Milesi-Ferretti (2004), Alt, Dreyer Lassen and Skilling (2002), and Alt and Dreyer Lassen (2002). The latter focus on the issue of how greater transparency affects budget deficits from both a theoretical and empirical viewpoints. In both papers, their underlying model is political agency. However, unlike our approach, they use a model where government must make an intertemporal financing decision – tax or debt. We will develop a model that includes public debt in section six. We will then revisit some of the transparency issues.

In a related contribution, Shi and Svensson (2000) model transparency as reflecting the fraction of the informed voters in the population. They show that a greater degree of transparency will tend to sharpen re-election incentives, making it more likely that an incompetent incumbent will be removed from office. Ferejohn (1999) studies the link between transparency and the size of government. He argues that greater transparency will see voters being willing to put government in charge of a larger amount of tax revenue.

In Chapter 3, we showed that whether more information is good or bad in agency models depends crucially on how that information is modeled. The model of this chapter suggests at least three dimensions of transparency which could be important:

- Information about the past records of candidates which reveals information about their underlying type, permitting voters to “pre-screen” candidates and leading to an increase in $\pi$.

- Information about the fiscal outcomes: better observations of $s$ or, in a more general setting, better information about taxing and spending.

- Information about the cost of public spending ($\theta$)

The only unambiguous case is transparency which raises $\pi$ – which can be thought of as improvements in public scrutiny which make it more likely that a politician devoted to serving voters’ interests is elected. The reason for this being welfare improving is clear – the equilibrium strategies are unaltered and voters always benefit from having a “good” politician in office.

Our analysis of discipline and selection effects spelled out so far confirms that, even if voters are provided with better information about $s$ or $\theta$, there is no guarantee that voter welfare will be higher. This will depend upon whether changes in incumbent discipline and are more or less important than changes in selection. To model this, suppose that, after the incumbent has chosen $s_1$ and before the period one election is held, the voter may learn about the true cost of public services $\theta$. Specifically, the true value of $\theta$ is
revealed with probability $\xi$; otherwise, no signal is received by the voter. The result in Proposition 3 is now modified. The payoff when a bad politician pools with a good one is $s(\mu) + (1 - \xi) \beta X$ while, if he chooses to reveal his type, it is $X$. Pooling is now worthwhile if and only if $s(\mu) > (1 - (1 - \xi) \beta) X$, a more stringent condition than in the absence of an informative signal (see Proposition 3). Moreover, pooling is less likely to be optimal the closer is $\xi$ to one. Indeed, if $\xi = 1$, then the only possible period one equilibrium is separating.\(^{26}\)

Better information therefore tends to reduce discipline and increase first-term rent seeking. At the same time, however, it improves selection, as bad incumbents are less likely to survive re-election. By evaluating the trade-off between agency costs of first-term and second-term incumbents, we can assess whether such a change in the political equilibrium is worthwhile. Assume that $q > 1/2$, the only comparison is between pooling ($\lambda = 1$) and separation ($\lambda = 0$). Evaluating welfare from ((4.2)) gives the difference between separating and pooling equilibrium welfare as

$$(1 - \pi)(1 - q)(\beta \pi \Sigma(\mu) - \Delta(\mu)).$$

The following result is due to Besley and Smart (2003):

**Proposition 13** (Besley and Smart (2003)) Suppose $q > 1/2$ and that the voter receives an informative signal about the cost of providing public goods. The signal improves voter welfare only if the selection effect of elections dominates the discipline effect, i.e. $\beta \pi \Sigma(\mu) \geq \Delta(\mu)$.\(^{26}\)

The above expression reveals the determinants of this trade-off. The selection effect is larger, the better is the pool of incumbents ($\pi$ close to one) and the lower the level of discounting.\(^{27}\)

This result builds on the findings of section 4.2 in Chapter 3. In political agency models, more information need not be desirable for voters. This is related to the broader theoretical literature on the value of information in agency relationships. In the standard complete contracts model of Holmstrom (1979), more information is better. Dewatripont, Jewitt and Tirole (1999) show that coarser information may sometimes be better in motivating agents in a career concerns model where incentives are implicit. This reflects a kind of second-best reasoning – with incomplete incentives, an otherwise welfare improving change (more information) may have deleterious effects on equilibrium behavior that more than offset the direct welfare impact. In our model, incentives are incomplete since the threat of not being re-elected is the only mechanism with which the voter may discipline incumbents.

\(^{26}\)A similar argument can be made if we instead supposed that with some probability information about whether $s_1$ is positive comes to light before the election.

\(^{27}\)This is also the case the higher is $W^\pi(L, \mu)$, the lower is $q$, and the lower is $W^\pi(H, \mu)$. 
Yardstick Competition  The last sub-section considered the direct effect of information provision. However, there are also indirect means of generating information about incumbents. One important example of this applies to decentralized governments where inter-governmental yardstick competition can be important.

This idea was first discussed in the context of governments by Salmon (1987) and its empirical relevance established in Besley and Case (1995a).\textsuperscript{28} The theory of yardstick competition is also studied in Bodenstein and Ursprung (2001), Bordignon et al (2001), Hindriks and Belleflamme (2001) and Sand-Zantman (2004). Increasingly, public spending functions in a number of countries are being decentralized to local governments and agencies. Decentralization is motivated in part by a desire to generate performance comparisons among decision-makers and so to enhance incentives for efficient provision of public services. If there is a positive correlation in the cost shocks, then this creates an informational externality. However, unlike the pure transparency model, studying this requires a model of equilibrium behavior generating information, i.e., equilibrium in both jurisdictions. Whether such yardstick competition is welfare improving also depends on how it affects the balance of discipline and selection effects in government.

To extend the model to include yardstick comparisons, suppose now that there are two identical jurisdictions, labeled “domestic” and “foreign”; variables that apply to the foreign jurisdiction will be denoted by the prime symbol. To focus on symmetric equilibria of the game among incumbents and voters in the two jurisdictions, assume that the joint probability distribution function of cost shocks $\Pr(\theta, \theta')$ is symmetric, with

\begin{equation}
\begin{align*}
\Pr(H, H) &= \Pr(L, L) = \frac{\rho}{2} \\
\Pr(H, L) &= \Pr(L, H) = \frac{1 - \rho}{2}
\end{align*}
\end{equation}

Moreover, we work with the case where $\rho > 1/2$, so that cost shocks in the two jurisdictions are positively correlated. To further simplify the analysis, we assume that $\delta(\mu) > (1 - \beta)X$, so that a separating equilibrium cannot exist. Since the marginal p.d.f. has $q = \Pr(\theta = H) = 1/2$, it follows from Proposition 3 that the unique equilibrium of the game without yardstick competition is one with pooling. We now show that, depending on the value of $\pi$, both hybrid and pooling equilibria are possible with yardstick competition.

When performance of foreign as well as domestic officials is observable, voters may base their decision to re-elect the incumbent or not on relative performance in the two jurisdictions. Voters will now condition their voting behavior on tax setting in both the domestic and foreign jurisdictions. Accordingly, let the probability of re-election in the domestic jurisdiction be $\sigma(x, x')$ when observed spending levels in the domestic and foreign jurisdictions are $x$ and $x'$ respectively. We say the voter’s strategy involves

yardstick competition when re-election occurs with positive probability if spending is high in both jurisdictions, but the probability of re-election is zero if domestic spending is high and foreign spending is low. That is, a re-election rule with yardstick competition has \( \sigma(x_H, x_H) = \sigma \) for some \( \sigma > 0 \) and \( \sigma(x_H, x_L) = 0 \).

As before, let \( \lambda \) denote the probability that type \((b, L)\) chooses \( s_1 = \hat{s}(\mu) \). Since we look for an equilibrium in which the strategies adopted by domestic and foreign incumbents are symmetric \((\lambda = \lambda')\), the p.d.f. \( \Pr(x, x' | i) \) of domestic and foreign spending conditional on the type of the domestic politician can be calculated as

\[
\begin{align*}
\Pr(x_H, x_H | g) &= \pi \frac{\rho}{2} + (1 - \pi) \frac{1 - \rho}{2}, \\
\Pr(x_H, x_H | b) &= \pi \lambda \frac{1 - \rho}{2} + (1 - \pi) \lambda^2 \frac{\rho}{2}.
\end{align*}
\]

(There are two terms in each probability because \( x_0 = x_H \) might have been generated by a good foreign politician facing high costs or a bad foreign politician facing low costs.) Voters’ posterior beliefs about the incumbent can therefore be calculated from Bayes’ rule:

\[
\Pr(g | x_H, x_H) = \frac{\pi}{\pi + (1 - \pi) \ell(\lambda, \rho, \pi)}
\]

where \( \ell(\lambda, \rho, \pi) = \Pr(x_H, x_H | b) / \Pr(x_H, x_H | g) \) is the likelihood ratio that \((x_H, x_H)\) was generated by a bad rather than good incumbent. Key to understanding the logic of the ensuing results is the fact that \( \ell(\cdot) \) is a decreasing function of \( \pi \) – the worse the initial reputation of the incumbent, the more likely it is that \((x_H, x_H)\) was generated by a bad incumbent. This is because, at low \( \pi \), \((x_H, x_H)\) is more likely generated by a foreign bad incumbent with cost of \( L \) than a foreign good incumbent with a cost of \( H \). But with positive correlation in costs, it is also more likely that the cost at home is \( L \) and hence that the domestic incumbent is bad.

A necessary and sufficient condition for an equilibrium with yardstick competition to exist is that \( \Pr(g | x_H, x_L) < \pi \), so that the voter prefers to remove the incumbent from office when domestic spending is high and foreign spending is low. Moreover, the equilibrium is pooling if \( \Pr(g | x_H, x_H) > \pi \) for \( \lambda = 1 \), and is hybrid otherwise. After some tedious manipulation, these conditions reduce to a simple one, given in the following result.

**Proposition 14** (Besley and Smart (2003)) Suppose that \( \hat{s}(\mu) > (1 - \beta) X \). Then voters use yardstick competition in equilibrium. A pooling equilibrium exists if and only if \( \pi \geq 1/2 \), and a hybrid equilibrium exists if and only if \( \pi < 1/2 \).

To interpret this, recall that, in the absence of yardstick comparisons, the equilibrium would have bad incumbents choosing \( x_H \) when the state is \( \theta = L \), yielding a pooling
equilibrium. Compared to this benchmark, the case of yardstick competition deviates in three ways. First, a bad domestic incumbent may not be re-elected when he chooses $x_H$, if the foreign incumbent is good and gets a low cost draw. Second, a good domestic incumbent is retained in office when costs are high, and the foreign politician chooses maximal rents. These changes to the voter’s strategy reflect the clear-cut information advantage from yardstick competition. Third, pooling may no longer be optimal for incumbents when the foreign incumbent has a poor initial reputation. To see this, observe that the likelihood ratio $\ell(\lambda, \rho, \pi)$ is decreasing in $\pi$, as it depends on the voter’s assessment of the quality of the incumbent in the other jurisdiction. Thus facing a foreign incumbent with a low reputation makes it relatively less likely that the $(x_H, x_H)$ outcome is generated by a good domestic incumbent, and hence that voters will re-elect an incumbent who picks $x_H$. The equilibrium now has the bad incumbent reducing the probability that he chooses $\hat{s}$ in order to raise the signaling value of the outcome $x_H$. A foreign incumbent with a poor reputation inflicts a reputational externality on a domestic bad incumbent and reduces his incentive to pool with a good incumbent. Moreover, this aspect of yardstick competition increases rent seeking.

Since improved information available through yardstick comparisons has countervailing effects on incentives and selection of politicians, its net impact on voter welfare is unclear. The following result shows that the reputations of politicians are key to understanding this.

**Proposition 15** (Besley and Smart (2003)) Suppose that $\hat{s}(\mu) > (1 - \beta)X$. Then there exist parameters $0 < \tilde{\pi}_a < \tilde{\pi}_b < 1/2$ such that voter welfare is lower when yardstick comparisons are available than when they are not if $\pi < \tilde{\pi}_a$, and the converse is true if $\pi > \tilde{\pi}_b$. This result emphasizes that voters who are better informed about the fiscal environment may be worse off in equilibrium, as bad politicians put less effort into building a reputation when they first take office. This insight explains the above result. In some circumstances ($\pi$ low), voters would be better off if they could commit to ignoring the fiscal performance in the other jurisdiction in the course of a domestic election.

Yardstick competition is welfare decreasing when politicians’ reputations are poor because rents are increased with little advantage from the improved information generated as most politicians who are kicked out are replaced by an incumbent of the same type.

### 4.5.3 Summary

To summarize, the following lessons learned here are central to understanding what follows:
Improving information about the state of the world need not be welfare improving – selection and discipline effects can work in opposite directions.

Changing features of the environment in which political agency plays out has direct effects and indirect effects – the latter working through changes in the political equilibrium. These welfare consequences of the direct and indirect effects can work in different directions.

Together these propositions make clear how a micro-founded model of politics is important to understand the economic logic of fiscal restraints.

4.6 Debt and Deficits

So far, we have focused on the case where government can only finance public spending through taxation. In this section, we extend the model to allow for public debt. One reason for doing so is simply that debt is an important tool of government and hence it should be considered in a dynamic model of public finance. However, debt plays an especially prominent role in the political economy of public finance. A central theme of the political economy literature is that public debt is unlikely to be chosen optimally from a social point of view.

There are two main arguments around to this effect:

- Strategic use of debt: This has been argued most prominently by Tabellini and Alesina (1990) and Persson and Svensson (1989). They suggest that an incumbent may choose to run a budget deficit in order to constrain the actions of future incumbents. The most popular application of these ideas is to the large Federal budget deficits of the Reagan era which provided a constraint on the spending patterns of future Democratic incumbents. This argument gives a reason for tendency towards excessive budget deficits since it is only putting future pressure on the government budget constraint will it have any effect.

- Fiscal Illusion: The other main concern with debt arises from the possibility that it is a less visible means of public finance than raising taxes. One aspect of this argument concerns the complex accounting conventions that surround public debt. Another possibility is that politicians deliberately find ways to cover up their deficits in order to make hay until eventually discipline is required. The most prominent advocate of the fiscal illusion view of excessive debt finance is Buchanan and Wagner (1977) who argue that “complex and indirect payment structures create a fiscal illusion that will systematically produce higher levels of public outlay than those that would be observed under simple-payments structures” (page 129) and that “events of fiscal history strongly support the hypothesis that unconstrained
access to public borrowing will tend to generate excessive public spending” (page 142).

While the first is clearly an important and interesting point, it has no direct bearing on political agency relationships being studied here. This is because period one incumbents in the model studied here derive no utility from period two fiscal policies unless they are re-elected. Fiscal illusion stories are frequently treated with a good deal of suspicion in the modern political economy literature (see, for example, Alesina and Perotti (1995)). The concern is that it attributes a systematic bias which appears to rest on voter irrationality. The fiscal illusion argument is, however, more interesting as one interpretation of it is precisely about the likely information that voters have on which to base their fiscal decisions.

In this section we explore two things. First, we consider how adding public debt into the model changes the insights in a world where political equilibria balance discipline and selection effects. Second, we investigate the need for a limit on the size of the deficit using this approach.

Debt has a dramatic impact on incentives of bad politicians. If they face no (non-economic) external constraint on debt issuance, they can steal tax revenue in period one and period two, and divert it to their own ends. This dramatically reduces their incentive to behave like good politicians in period one. This bears out the standard public choice suspicion about debt as an instrument.

**Public finance and government debt.** We begin by outlining a simple extension of our two period model to incorporate government debt. Suppose that, after observing $\theta_1$, the incumbent may issue debt $D$ in period 1 which is observable to the voter, incurring a gross liability payable in period 2 for the government of $R(D)$, where $R(D) \geq D$ for $D \geq 0$ and $R(D) < D$ for $D < 0$. We also assume that $R(\cdot)$ is increasing and strictly convex.\(^{30}\) Furthermore, we will assume that $\beta R(0) = 1$, so that (as we shall see) the role of debt for a good government is to smooth taxation between periods in view of the shocks to the cost of providing public goods.

In the presence of debt, the government’s budget constraints in the two periods become

$$x_1 = \theta_1 G_1 + s_1 - D$$

and

$$x_2 = \theta_2 G_2 + s_2 + R(D).$$

The optimal fiscal policy may again be solved by backward induction. In period two:

$$G^*_{\theta_2} = \arg \max \{G - \mu C (\theta_2 G + R(D))\},$$

\(^{30}\)The convexity ensures an interior solution for the debt level. For $D > 0$, it crudely captures the possibility of default as perceived by investors. For $D < 0$, it captures the absence of an infinitely elastic supply of quality investment opportunities.
which implies that:

\[ 1 = \theta_2 \mu C' (\theta_2 G_{\theta_2} + R (D)) \, . \]

Thus period two taxes will, as in section 2, be equal to \( x^*_\theta \) with spending equal to \( x^*_\theta / \theta - R (D) / \theta \). Thus debt reduces period two resources devoted to valuable public spending but does not affect period two taxation levels. Expected period two welfare is therefore equal to \( \text{EW}^*(\mu) - \psi R (D) \) where \( \psi = q / H + (1 - q) / L \) as before.

It is now straightforward to determine the period one optimal policy. The level of public spending will solve:

\[ G^*_1 = \arg \max \{ G - \mu C (\theta_1 G - D) \} . \]

As above, we confirm the principle that the level of period one taxation is set equal to \( x^*_\theta \) with spending being \( G^*_1 = x^*_\theta / \theta + D / \theta \). So the incidence of debt is purely to increase public spending.

The optimal debt level is used to smooth the cost of public spending across the two periods with taxes remaining constant. At an interior solution, the optimal level of debt, \( D (\theta_1) \), satisfies:

\[ \frac{1}{\theta_1} = \beta \psi R' (D (\theta_1)) . \]

Since \( \beta R' (0) = 1 \), the model predicts that the government will run a budget surplus when costs are high \( (D^*_H < 0) \) and a budget deficit when costs are low \( (D^*_L > 0) \). This is a direct application of the Barro (1979) tax smoothing idea. Compared to the case where \( D = 0 \), the government runs a deficit when the size of the public sector is large and a surplus when it is small.

**The politics of debt.** We turn to characterizing the effect of debt on our political equilibrium. Let \( \bar{D} \) be the highest level of debt that can be incurred in period one. In the absence of any constitutional restrictions on debt, we will have \( R(\bar{D}) = X \); however, we also consider tighter limits below. A bad politician who decides to go-for-broke in period one will now take rent equal to \( X + D \). At the same time, the amount of rent that he can extract by mimicking a good incumbent with cost shock \( H \) is lower in both periods one and two than in the case of no debt. This is because he will have to repay the debt out of any rents that he can capture tomorrow and the tax-smoothing argument implies that public spending will be lower today than it would be in the absence of debt (that is, \( D^*_H < 0 \)). Thus, the condition for being willing to exercise restraint is:

\[ (H - L) \left[ \frac{\hat{\sigma}_H + D^*_H}{H} \right] + \beta [X - R (D^*_H)] \geq X + D . \]

Comparing this inequality with \( (\mu) \geq (1 - \beta)X \) in Proposition 3, it is clear that the condition for incumbent discipline is more difficult to satisfy. Thus we have:
Proposition 16 Suppose that the government can raise resources though both debt and taxation. Then, public debt reduces incumbent discipline. More specifically, a pooling equilibrium with \( \lambda = \sigma = 1 \) exists if (4.11) holds. The equilibrium is separating otherwise.

The fact that debt leads to ill discipline motivates the observation that the citizens would wish to impose a restriction on the deficit that can be raised. If the political equilibrium is pooling, then a relatively weak restriction on debt \( \tilde{D} > D^*_L \), which does not change the political equilibrium, must improve voter welfare. It leaves equilibrium behavior for good incumbents and bad incumbents in state \( L \) unchanged. However, it reduces the long-run harm that bad incumbents can do in state \( H \). If the political equilibrium is separating, then again good incumbent behavior is also unchanged while bad incumbent behavior is now restrained in every state. In all cases, political discipline is improved and the ability of the political process to select good incumbents is unchanged.

Thus, we have

Proposition 17 A cap on the size of the deficit of \( \tilde{D} \) which does not change the political equilibrium raises voter welfare.

If the debt limit is sufficiently tight that the political equilibrium is changed, it is in the direction of increasing incentives for pooling. A switch from separating to pooling reduces the information generated in the political equilibrium and hence is most valuable when selection is not very important (low \( \pi \)). Thus, this view of reducing deficits fits much better with the conventional public choice view that restraining government is important in two senses. First, it is a restraint on bad governments. Second, the resulting changes in political behavior have positive effects so long as the likelihood of a bad government is high.

This focuses on a rather different distortion due to public debt than has been put forward in either the strategic debt model or the fiscal illusion view. The key idea is that debt allows bad politicians not only to extract rent from current taxpayers but also from future ones. Thus, the quality of policy is more susceptible to political opportunism. Without placing a clear restraint on the amount of debt that politicians can run up, there is a risk that ill-discipline will now be much more of a problem.

The result also points to the limitations of elections as a disciplinary device, and so the potential role for fiscal restraints. Unlike the previous results on fiscal restraints, this one does not rest on a balance of selection and incentive effects. Over the range that we considered \( (\tilde{D} \geq D^*_L) \), the political equilibrium is the same as it would be without the debt restriction – there is a pure discipline effect. Whether further debt restrictions below this level are desirable depends on whether distorting the behavior of good politicians and/or changes in the political equilibrium are desirable.
4.7 Governments versus NGOs

The model so far has assumed that government is the only means of delivering provision of public goods. However, the kind of agency problems studied here are a key motivating force behind efforts to delegate responsibility for public goods provision to non-governmental organizations. Indeed pessimism about the state in developing countries has seen increasing amounts of aid assistance being dispersed in this way.

Interest in NGOs versus government provision is not confined to a developing country context. One structural distinction between countries concerns the extent to which private rather than state delivery of services is used. Moreover, there is much debate about where the border should lie. For example, George W. Bush announced a far-reaching proposal to use "faith-based" organizations in service delivery.

A burning issue in the debate concerns the way in which NGOs are held accountable for the quality of service that they deliver. Unlike (democratic) governments, there is no direct accountability to the beneficiary groups that they help. To understand their role in providing more effective aid provision, it is necessary to understand whether this is an issue. At one extreme, NGOs could be thought of as being free from the kinds of incentive problems that are faced by government. Indeed, the rose-tinted spectacles that have been used to describe NGOs in some quarters perpetuate this characterization. But we need only remind ourselves of the problems that can arise from ignoring incentive problems in government to realize the potential folly of this.

Here, we adopt a level playing field approach. NGOs will be agents just as governments are. We will explore aspects of their behavior as equilibrium phenomena. However, one one hand, we will accept the possibility that there are well meaning NGOs in the world. However, on the other hand, there is a selection problem in identifying them.

We motivate these ideas with the following example. Suppose that an external agency is deciding to fund an increment of public spending and it can do so either by offering resources to a government or via a non-governmental organization. It cares that these resources are turned into public goods that are valued by the citizens.

There are two ways in which NGOs could improve on the outcomes implemented by governments in the kind of setting that we are considering. First, it may be possible to write contracts with NGOs that are more complete than those offered to government. This is because the government is primarily accountable to its citizens via elections. The NGO could, in principle, be subject to all manner of other accountability mechanisms. This should manifest itself in better discipline, i.e., lower rent seeking.

Second, the outcome could improve if the composition of individuals who work for NGOs is drawn from a better pool. What makes NGOs unique among possible services providers is the idea that they are not merely staffed by self-interested individuals. Working for an NGO is not like working in the private sector. This raises the potential for NGO delivery of services having an advantage over government in that the agents are
more likely to be of the “good type” i.e., maximizing the interests of the beneficiaries. On the downside, it is also pointed out that, unlike governments, NGOs are not accountable to those whom they serve. There are no elections for NGO leaders. This raises the possibility of a trade-off between improved selection from having an NGO undertake spending tasks and reduced accountability because there is no electoral mechanism.

We begin by studying the case where the NGO faces the same extent of accountability as does the government. Specifically, the NGO can be fired if it does not do its job properly, but there are no other sanctions available. The NGO accountability mechanism now looks very similar to elections as means of generating incentives.

4.7.1 Framework

Suppose that an external organization that cares about $G$ is giving “aid” to a country. The level of aid is denoted by $T$. We remain within a two period horizon and the information structure is exactly as above. The decision by the aid provider is whether to give the aid to the NGO or to the government.

We maintain the parallel structure to our model of government by supposing that there are two kinds of NGOs. Good NGOs spend all aid resources on public goods whereas bad NGOs consume the resources privately. We also assume complete symmetry in the technology that governments and NGOs have for producing public goods. Thus each faces the same structure of costs shocks to the provision of public goods. For simplicity here, we focus on the case where $q > 1/2$. This rules out the hybrid equilibrium in Proposition 3 and allows us to focus on pooling or separation.

The difference between an NGO and a democratic government is that the former is not directly accountable to the people whom it serves. The main incentive that we consider is retaining the right to provide public goods on behalf of the aid provider. Thus, the aid agency cannot commit to an incentive scheme ex ante. It must make a decision whether or not to retain the NGO ex post.

We assume that there is a potential pool of NGOs from which an aid agency can select. The probability that a randomly selected NGO is good is $\nu$. Suppose that the aid provider decides to go with an NGO, then the timing of the NGO/aid donor game is as follows.

1. Nature chooses an NGO at random and a cost of public goods provision $\theta$.
2. The NGO decides how much of the aid, $T$, to spend on public goods and how much on its private interest.
3. The aid agency then decides whether or not to retain the NGO against another randomly selected NGO.
4. The period two value of $\theta$ is realized.
5. The incumbent NGO receives \( T \) and decides whether or not to spend its resources on public goods or private spending.

This gives rise to a game that is very similar to the one played between a government and its citizens. We now consider the equilibrium policy choice of the NGO.

A good NGO spends \( T/\theta \) on public goods. Just as in the case of a bad government, a bad NGO can seek to mimic a good one and retain “its franchise” to provide the public good. In the state where \( \theta = L \), it can mimic the good NGO by choosing:

\[
G = \frac{T}{H}
\]

and collecting a rent of \( \left( \frac{H-L}{H} \right) T \). Hence the condition for restraint being optimal (assuming that the NGO is retained) is that

\[
\left( \frac{H-L}{H} \right) T + \beta T > T \text{ or } \beta > \frac{L}{H}.
\]

Let \( \lambda_n \) be the probability that a bad NGO picks \( G = T/H \) when the state is \( \theta = L \) and let \( \sigma_n \) be the probability that the NGO is retained when the observed level of public goods is \( T/H \). We now have the following proposition describing the equilibrium behavior of the NGO.\(^{31}\)

**Proposition 18** Suppose that \( q > 1/2 \). Then an equilibrium exists for all values and is generically unique. There are two cases:

1. A pooling equilibrium with \( \lambda_n = \sigma_n = 1 \) exists if and only if

\[
\beta \geq \frac{L}{H}
\]

2. A separating equilibrium with \( \lambda_n = 0 \) and \( \sigma_n = 1 \) exists if and only if

\[
\beta < \frac{L}{H}.
\]

The reasoning is exactly the same as for the case of government accountability. The Proposition shows that there can be an equilibrium where a bad NGO mimics a good one provided that it is sufficiently patient. The expected discounted level of public goods being provided viewed from the perspective of period one is now easy to calculate for each kind of equilibrium. In a pooling equilibrium this is

\[
T \left[ \nu \psi \left( 1 + \beta \right) + (1 - \nu) \left( \left( \frac{1-q}{H} \right) + q \nu \psi \beta \right) \right]
\]

\(^{31}\)As in our model of government, we assume that an NGO that does anything which could not be generated by a good NGO is assumed to be bad.
where $\psi = q/H + (1 - q)/L$ as before. In a separating equilibrium, it is:

$$T \left[ \nu \psi (1 + \beta) + (1 - \nu) \nu \psi \beta \right].$$

Whether, from an ex ante point of view, there are more public goods produced in one case than another is not clear. This is true under pooling if

$$\frac{1}{H} > \nu \psi \beta.$$

Since $\frac{1}{H} < \psi$, this can hold only if $\beta$ and $\nu$ are low enough. Intuitively, this is clear since pooling is preferable if there is a low level of patience (other things being equal) as the gains from better discipline occur in the short run. It is also clear that pooling is preferable if the NGO is more likely to be bad.

This analysis shows that the balance of selection and discipline effects are also important in understanding the role of NGOs. Hence the model achieves the level playing field that we desired to investigate the role of NGOs. We now consider what would happen if the aid provider were instead to give the aid to government.

### 4.7.2 Aid to the Government

The alternative to financing an NGO is giving aid directly a government. Thus, suppose instead that the aid donor chooses to give $T$ extra resources to a government. Since it plays no part in the analysis, we set $\mu = 1$. We assume that the level of aid is observable to the citizens/voters and they can condition their voting behavior on it. Let

$$G_p^*(T) = \arg \max G - C(\theta G - T)$$

be the level of public goods chosen by a good politician when aid is $T$. It is clear that if the government acts in the citizens’ interests, then this aid will increase in public spending of $T/\theta$ – exactly the same amount as a faithful NGO does. Therefore, $\partial G^*/\partial T = 1/\theta$. Modifying the analysis above, the critical rent level at which a bad politician can mimic a good one is now:

$$\hat{s}(T) \equiv (H - L)G_{H}^*(T)$$

Access to aid resources therefore increases the level of rent extraction that is feasible and consistent with re-election.

The result in Proposition 3 is now modified as follows:\(^{32}\)

**Proposition 19** An equilibrium exists for all values and is generically unique. There are two cases:

\(^{32}\)Recall that we assuming that $q > 1/2$ so that a hybrid equilibrium is ruled out.
1. A pooling equilibrium, with $\lambda = \sigma = 1$, exists if and only if
   \[ \hat{s}(T) \geq (1 - \beta)(X + T) \]

2. A separating equilibrium, with $\lambda = 0$ and $\sigma = 1$, if and only if
   \[ \hat{s}(T) < (1 - \beta)(X + T) . \]

A bad politician will mimic a good one now if $\hat{s}(T) + \beta(X + T) > X + T$. This result demonstrates an often made claim that aid payments change government incentives and specifically may create worse discipline among incumbents. Aid destroys government incentives to pool in cases where $\beta < \frac{L}{H}$. Thus, a government that is suddenly given a large injection of aid may indeed decide to make hay even though there may be an electoral penalty from doing so. Such considerations are clearly important when choosing whether or not to run aid programs through governments.

We can now compute the incremental level of public goods spending by government under each kind of equilibrium. Under pooling, it is:

\[ T[\pi \psi(1 + \beta) + (1 - \pi)\left(\left(\frac{1 - q}{H}\right) + q\pi \psi \beta\right)] \]

while in a separating equilibrium it is:

\[ T[\pi \psi(1 + \beta) + (1 - \pi)\pi \psi \beta] . \]

Note that both payoffs are identical to what we obtained for the case of an NGO, except that $\nu$ is replaced by $\pi$.

### 4.7.3 Comparisons

Our first observation comes from comparing the expressions for expected additional public good payments by governments and NGOs. Here, we have:

**Proposition 20** Suppose that an external agency wishes to give a transfer to a government to promote public spending. Suppose also that the NGO and the government either both pool or both separate, then voter welfare is higher when the transfer is given to an NGO if and only if $\nu > \pi$.

This result says that provided that NGOs and government have the same incentives, then whether it is desirable to direct aid via an NGO comes down purely to whether

---

This is exactly the condition for a bad NGO to choose separation.
there is a selection advantage. In other words, a preference for NGO provision comes from believing that they have access to a pool of workers who are more likely to run the program in the interest of citizens at large.

A key assumption is that the NGO and the government play the same equilibrium strategy. However, comparing Propositions 11 and 12, it is clear that there is no reason why this should be the case. If it is not, then the result in Proposition 13 needs to be modified. Given that they are dealing both with taxes and aid payments, it is clear that bad governments may have a larger incentive to pool. The government’s “pooling” condition can be written as:

\[ \hat{s}(0) + T \left( 1 - \frac{L}{H} \right) > (1 - \beta)(X + T) \]

Hence, in the case where the NGO will pool, this reinforces the incentive of the government to pool. Suppose that the government pools and the NGO separates, then NGO provision is better if pooling is valuable, i.e.,

\[ (\pi - \nu)\psi(1 + \beta + \beta q(1 - \pi - \nu)) + (1 - q)(1 - \pi) > (1 - q)(1 - \nu)\nu \psi \beta. \]

If the quality of government and NGOs is the same then this condition boils down to:

\[ \frac{1}{H} > \pi \psi \beta, \]

in which case then government provision is better if the quality of governments and NGOs is very low. If NGOs are of better quality than government, then government provision can still be better if the difference in quality is not large and if the quality of government is low.

To summarize, our model makes three key predictions about when NGOs are more likely to be better than government in delivering public goods:

- If NGOs and governments have the same equilibrium strategies then NGOs needs to be staffed by individuals who are more motivated towards the public interest (better selection).
- If NGOs separate and governments pool, then we need a large difference in quality between NGOs and government or to have a government of good quality.
- If NGOs pool and governments separate, then we need a small difference in the quality of NGOs and government and the quality of NGOs to be low.

---

34 This has a parallel with Besley and Ghatak (2001) who show that NGO provision is preferred to government provision in an incomplete contracting framework with public goods only if the NGO values the welfare of the target group more than does the government.

35 This happens only if \( \hat{s}(0) > (1 - \beta)X \), i.e. without aid the government would pool.
Obviously the model is simple and stylized. However, it is clear that the model has the key ingredients that are needed to study the issue. First, the model models the imperfection in government behavior that generates the case for an NGO endogenously. Second, it models how aid to government changes the incentives of the government. Third, it models the contracting process between the aid agency and the NGO, given the weak accountability that the NGO has.

4.7.4 Further Issues
There are a number of further issues that could be usefully be studied in the framework developed here. In particular, there are a number of asymmetries in the accountability of NGOs and governments that we have set aside.

As noted above, the analysis assumes that the kinds of contracts that can be reached between NGOs and aid providers are exactly the same as those that can be reached between governments, citizens/voters and NGOs. This facilitated a simple analysis, but it arguably misses an important point. If more sophisticated contracts can be used, then this would enhance the case for NGOs. However, a similar comment applies to government-aid provider relations. We did not allow any direct elements of conditionality into the contracts there and supposed that there was a sovereign political process which governs this. Clearly, meaningful notions of joint accountability to voters and aid providers create some interesting common agency problems which deserve further work and could better illuminate the meaning of conditionality in exchange for aid alongside domestic political processes. This discussion is a first step down this route.

Another feature of the model that could be developed is to suppose that democratic accountability may utilize information about government performance better. We have assumed that aid providers have as much information about how much of the aid finance goes into public goods as do voters. However, it appears more likely that information will be greater among voters. This is particularly true when aid delivery is by local governments. While it might be assumed that this would go in favour of government provision, our observation that better information about the local conditions need not improve voter welfare (see section 5.2.2) would be an important caveat to the fad for decentralization based on the assertion that this will enhance the use of local information. Developing a model in this direction would cast light on the exact benefit and cost of aid provision through NGOs versus local governments.

Another key difference between NGO and government accountability concerns the scope of their operations. The government is usually a multi-task agency while NGOs have a more narrow domain of responsibility. This surfaced in a limited way in the analysis so far in the fact that the incentive conditions for NGOs and governments diverged. In general, this observation suggests that it should be easier to design an incentive scheme for NGOs because the specific task that they are engaged in can be written into the contract whereas citizens hold government to account on a wide variety
of performance outcomes. The tasks being funded by aid may simply be too small to have a strong impact on government accountability (for better or worse).

### 4.8 Competence

The model has focused on agency problems due to politicians not serving voters interest. It is interesting to see how the model might instead be used to study politicians who differ from each other in terms of competence.

To study this, we set up a model where there is an inherent characteristic of politicians which affects the cost of providing public goods. This is similar in spirit to the way that we chose to model competence in section 4.3 chapter 2. If incompetence were the only issue, then incompetent politicians might be happy to fall on their sword and be voted out. To create an agency problem, it must be that incompetents earn a rent. Suppose then that there is an “ego” rent (i.e., an increment to utility) rather than a monetary payoff which we denote by \( \eta > 0 \).

To incorporate competence we suppose that the cost of producing public goods depends on \( \theta \) (a cost shock unobservable to the voter) and also on a competence parameter \( \kappa \). Specifically, the cost is

\[
\begin{align*}
    \theta G & \quad \text{if the incumbent is competent} \\
    \kappa \theta G & \quad \kappa > 1: \text{if the incumbent is incompetent}.
\end{align*}
\]

We assume that an incompetent incumbent can mimic a competent one at a personal cost to himself. This cost is proportional to the amount of cost-reducing effort that he puts in. To lower costs per unit of public goods by an amount \( \delta \) requires committing cost reducing effort of \( c\delta G \).

The timing of the model is basically the same as in the baseline model. Nature first determines the type of the incumbent and the cost of providing public goods. The incumbent then chooses a level of taxation and public goods and the amount of cost reducing effort to put in. The rest of the game structure is as before.

The features of the story developed in section 3.3 now go through. A competent incumbent will choose \( G_\kappa^* (\mu) \). An incompetent incumbent observes the cost shock and decides whether to commit cost reducing effort and mimic a competent incumbent or make no such effort. In the latter case voters will observe that they pay taxes of \( \kappa \theta G_{r,\theta}^* (\mu) \) and will vote the incumbent out of office.

Now consider the incentive of the incumbent to pool. Suppose first that the cost shock is \( L \) and let \( \lambda \) be the probability that he pools (i.e. picks sufficient effort to mimic the competent incumbent). He can pick effort of \( (\kappa - 1) L G_\lambda^* (\mu) \) or \( (\kappa L - H) G_H^* (\mu) \).

It is straightforward to check that he will always choose to put in effort to mimic the cost shock \( H \). Then he can choose the level of public goods \( G_H^* (\mu) \) and then “subsidize” the
level of public goods to mask his incompetence to the tune of \((\kappa L - H) G_H^* (\mu)\). This is costly today, but if it results in the incumbent being re-elected, this is worthwhile if:

\[
\beta \eta - c (\kappa L - H) G_H^* (\mu) > 0.
\]

Now consider what happens in the state \(H\). He has to commit cost reducing effort of \((\kappa - 1) H G_H^*\) to mimic a competent incumbent with a cost shock of \(H\). It is now straightforward to see that the condition required for a pooling equilibrium to exist is:

\[
\beta \eta - c (\kappa L - H) G_H^* (\mu) > 0 > \beta \eta - c (\kappa - 1) H G_H^* (\mu).
\]

Otherwise, the equilibrium will be separating.\(^{36}\) If \(q > 1/2\), then the above condition is both necessary and sufficient for pooling.

The model with competence differences between politicians will exhibit the same selection and incentive issues that we found in the model where the difference between politicians is due to motivational differences.

### 4.9 Conclusions

The main aim of this chapter has been to set the political agency model in a less abstract setting than the previous chapter. Our economic setting has been a simple public finance model where a government must determine a level of public spending and finance this out of taxes and debt. We have presented a simple and stylized political model which determines the equilibrium policy path consistent with voters optimally using re-election incentives.

The core idea, which carries over from chapter three, is that the political equilibrium can be understood as the balance of selection and incentive effects in politics. The way in which they interact determines the welfare properties of political equilibria.

Despite being very simple, the model offers some suggestions about the links between re-election and policy outcomes. It also gives some insight into tax and spending cycles when politicians cannot face re-election at the end of the second term.

The main analysis was devoted to studying the implications of the agency model for ways of adding additional constraints on government which affect its ability to tax and spend. We couched this as a kind of political second best theory. While elections may help to combat agency problems due to information problems, they fall short – some amount of rent seeking remains along the equilibrium policy path. The problem is then to view additional restraints on government in this second-best context. Restraints can now have two effects – direct effects when the restraint changes the economic constraint

\(^{36}\)Observe that if the incompetent incumbents always provide \(G_H^*\) and \(x = H G_H^*\) regardless of the cost shock (ie. if \(\beta \eta - c (\kappa - 1) H G_H^* (\mu) > 0\)): then the voter ousts any incumbent providing \(G_H^*\) and \(x = H G_H^*\). In this case, incompetent incumbents lose an incentive to mimic.
set and indirect effects when they lead to politicians changing their political behavior. The latter are mediated through the selection and incentive effects in politics.

We also considered the possibility of channelling resources through non-governmental organizations. Here, the framework is useful since it allows us to model incentive problems in government and NGOs. The agency model provided some useful insights in this regard.

The model is only illustrative. It is certainly too simple and abstract to provide a guide to the details of policy debates. However, as a way of thinking it provides a balanced perspective, in a public finance setting, on the conflict of interest that arises between voters and government. This is not the only relevant dimension when we think about the political economy of public finance. However, it is an important one and clearly the correct one for addressing policy debates about the need to impose additional constraints on governments. It also has the attraction of being rooted in empirical observations on taxes, spending and political turnover.

Appendix: Optimal fiscal policy with pure moral hazard

Section 4.3.2 described the voter’s optimal fiscal policy in the case of pure moral hazard, in which politicians are known to be the bad type, and the voter can commit to a re-election rule in advance. Here we provide a formal solution of the problem.

The voter’s problem is to choose levels of spending and taxation \((G_L, x_L)\) and \((G_H, x_H)\) to maximize expected welfare

\[
q \left[ G_H - \mu C(x_H) \right] + (1-q) \left[ G_L - \mu C(x_L) \right]
\]

subject to the participation constraint guaranteeing that type \(H\) prefers to produce \((G_H, x_H)\) to extracting \(X\) in rents immediately:

\[
s_H = x_H - HG_H \geq (1 - \beta)X \quad \text{(PC}_H)\]

and to the incentive constraint guaranteeing that type \(L\) prefers to produce \((G_L, x_L)\) than \((G_H, x_H)\):

\[
s_L = x_L - LG_L \geq x_H - LG_H = s_H + (H - L)G_H \quad \text{(IC}_L)\]

This is the “relaxed” problem, which ignores the analogous participation constraint for type \(L\) and incentive constraint for type \(H\). It is immediately clear however that any solution satisfying the above constraints with \(G_H \geq 0\) will also satisfy \(L\)’s participation constraint (since \(L < H\)). That \(H\)’s incentive constraint is also satisfied is less evident, however. Our approach will be to solve the relaxed problem, and then check that \(H\)’s incentives are also correctly aligned.
Since higher taxation $x_\theta$ decreases welfare of the voter but increases rents to the politician, the solution to (4.12) must have both (PC$_H$) and (IC$_L$) satisfied as equalities. The first-order conditions for the voter’s problem can then be simplified to

$$\mu C''(x_L) = \frac{1}{L}$$  \hspace{1cm} (13)

$$\mu C''(x_H) = \frac{1}{H} \left( 1 - \frac{1 - q H - L}{q} \right)$$  \hspace{1cm} (14)

$$x_H = HG_H + (1 - \beta)X$$

$$x_L = LG_L + (1 - \beta)X + (H - L)G_H$$

Equation (13) shows that low-cost fiscal policies are undistorted: the marginal benefit of spending is set equal to its marginal tax cost, as represented by the point of tangency $B_2$ in Figure ???. Equation (14) shows that spending in the high-cost state is distorted downwards: $\mu C''(x_H) < 1/H$.

Note that (14) implies that $x_H$ (and $G_H$) is an increasing function of $L$ and $q$. This confirms Proposition 2 in the text that lower costs in the sense of first-order stochastic dominance lead to greater spending in the high-cost state. Thus, reductions in the cost of government services may actually lead to further restrictions on government spending, as they worsen the potential agency problems created by bad politicians.

To complete the derivation, we must check that a politician facing high costs does not prefer to adopt the fiscal policies intended for the low-cost type. The relevant incentive constraint is

$$s_H = x_H - LG_H \geq x_L - HG_L = s_L - (H - L)G_L \quad \text{(IC}_H\text{)}$$

Substituting into (IC$_L$), if the two incentive constraints are both two satisfied at any pair of fiscal policies, then we must have

$$(H - L)(G_L - G_H) \geq 0 \quad \text{(IC}_H\text{)}$$

or simply $G_L - G_H \geq 0$. To verify that this necessary condition for implementation holds, we may invert the first-order conditions (13)–(14) to obtain

$$L(G_L - G_H) = \frac{1}{\mu} C''^{-1} \left( \frac{1}{L} \right) - \frac{1}{\mu} C''^{-1} \left( \frac{1}{H} \left( 1 - \frac{1 - q H - L}{q} \right) \right) \geq 0$$
\[ G \times (1 - \beta)X \]

\[ A_1 x^H H \]

\[ B_1 \]

\[ B_2 \]

\[ x^L \]

\[ (1 - \beta)X \]

\[ x^H \]

\[ x^H \]

\[ x^L \]

Figure 1: Optimal spending with bad politicians
Chapter 5

Final Comments

The main aim of this book has been to look at the problem of good government through the lens of modern political economy. This is an evolving field and much has still to be learned. In this final brief chapter, I will discuss some topics for future research which follow on from the discussions of the previous chapters.

The topic of who chooses to become a politician and how institutional structure shapes this decision is relatively neglected in political economy.\(^1\) However, our analysis has shown that the quality of the political class is an important determinant of good policy. We showed in chapter two that this was a source of political failure. It was also a recurrent theme in agency models with adverse selection and moral hazard in chapters three and four. Thus, it would be useful to have an understanding of how elections work to select good politicians.

While elections are now accepted as the main vehicle for picking politicians, this was not always the case. For a period of time, ancient Athens filled seats on its legislative council by drawing lots from among its citizens (Manin, 1997). Each citizen served for one year and there was a restriction to two terms in a lifetime. The Greeks understood the downside of this method in terms of ensuring good politicians. They did impose safeguards in the form of a kind of confirmation hearing in which the character and competence of the selected candidate was scrutinized. However, the basic premise behind selection by lot is that civic virtue was widely distributed in the population, so that random selection made it relatively unlikely that anyone picked by the lottery would be a bad politician. Selection by lot was deemed preferable to elections for three main reasons. First, it guaranteed rotation in office, so that politicians were guaranteed to experience both political and everyday life. Second, selection by lot guaranteed the widest possible access to public office and hence was viewed as egalitarian. Third, lots

\(^1\)The ensuing discussion of the history of selection is based on Besley (2005).
seemed more likely to maintain a unity of purpose in the community, while elections increased the chance that citizens would group into factions.

The use of lottery makes a lot of sense in a relatively homogenous city state such as Athens. For similar reasons, lotteries were also used in the Italian city states of Venice and Florence. However, even political thinkers such as Montesquieu and Rousseau who took the idea of political selection by lot seriously in their writings ultimately favored elections, principally because they believed that elections helped in the selection of a natural aristocracy of the talented and virtuous. After all, selection by lot does not favor those with greater political competence over those with less.

This view heavily influenced the founding fathers of the United States, who similarly saw the task of political selection as selecting a ruling class that was different from the citizens at large – superior in their talents and mental capacities. Indeed the term “natural aristocracy” originates with Thomas Jefferson (1813), in a letter written to John Adams. Jefferson wrote: “I agree with you that there is a natural aristocracy among men. The grounds of this are virtue and talents. . . . May we not even say, that that form of government is the best, which provides the most effectually for a pure selection of these natural aristoi into the offices of government?” Jefferson continues to argue that he favors laws to break up large inheritances and support public education as methods of creating a situation in which the natural aristocracy can rise and be selected.

The idea of a natural aristocracy produced by election contrasts with the prevailing norm in early modern Europe, where many countries still relied heavily on hereditary aristocracy. The idea of selection by blood-line makes sense only if qualities required for making policy are passed between generations, either genetically or by some form of social conditioning, and offers no safeguards if the transmission of such traits fails. It also has an advantage when it permits the ruling class to take a long-run view. The hereditary aristocracy had some ability to co-opt the most talented and successful citizens in any generation – but it was much harder to demote those hereditary aristocrats who lacked political competence. Hereditary aristocracy also has a built in bias towards the interests of the rich. In such a setting, broad-based policy is on the basis of noblesse oblige and/or the threat of revolution.

Some countries, notably the United Kingdom, still have elements of hereditary selection of their political class. Even representative democracies have prominent political dynasties. The Nehru dynasty in India spawned three prime-ministers who “ruled” India between 1947 and 1989 with only a two year hiatus. The current Bush dynasty in the United States appears likely to have held the presidency for 12 out of 20 years. The advantage in name recognition is palpable. But whether politician quality is transmitted inter-generationally is far from clear. Debates about hereditary in politics are reminiscent in this regard to discussions about whether family-owned firms have advantages in the market place.

When hereditary aristocracy is considered through the lens of political selection, it is no great surprise that this institution has largely disappeared over time. There is
little to commend it as a means of picking the best people to make policy. The principle of heredity was dealt its severest blow by the increasing influence of Enlightenment thinkers, such as John Locke, who championed the idea that government could only gain its legitimacy from consent given freely by the citizens as expressed through the ballot box. This claim about perceived legitimacy has nothing directly to do with whether elections provide incentives to discipline rules; instead, it’s a claim about the importance of selection to a well-functioning political system.

In the modern world, most autocrats rule by force rather than inheritance. Autocracy’s selection rule is to pick those leaders who can must the greatest loyalty from coercive institutions such as the army and/or police. While autocracies display a fair amount of heterogeneity in their performance, it seems fair to say that the most brutal regimes in human history have typically been of this kind.

Selection by force does not appear to be an effective method either for selecting people who will offer wide-ranging political representation or those who will be the most competent to hold public office. Jones and Olken (2004) offer some evidence that the low quality government under autocracy is rooted in selection problems. They look at economic growth before and after the death in office of world leaders to provide evidence on whether leader quality matters. Their main finding is that the death of an autocrat leads to an improvement in growth. Democratically elected leaders show no corresponding pattern. This finding is consistent selection by force having a disadvantage in selecting good leaders compared to democratic selection.

Even though elections are now the most commonly used institution for selecting a political class, there is enormous diversity in institutional structure. There are differences in the wages paid to politicians, the kind of public duties that are required, and the kind of policy discretion that the politician enjoys. There are diverse rules about who can stand for office, such as age, wealth or literacy requirements. Rules about the conduct of elections also differ across democratic systems, including whether there are primaries for multiple candidates from the same party or runoffs between the top two candidates. Rules also differ with regard to the conduct of campaigns and the kind of financial support that candidates may receive.

For the most part, political economy model has studied how institutions shape incentives to offer policies either before or after elections have taken place. As a result, our understanding of electoral institutions shape political selection is in its infancy. Creating a better understanding of this (theoretically and empirically) creates a challenge for the future. This requires understanding who choose to become a politician in the first instance and how this fits into broader career concerns.²

The approach throughout most of this book is quite individualistic. But much political activity takes place in collectives such as cabinets, legislatures, parties, interest groups and NGOs. The dominant tradition in political economy has been treat them as

²See Matozzi and Merlo (2005) for an interesting start on such a project.
unitary actors with coherent preferences. But there is much to be done to understand the internal workings of such organizations and the implications for selection and incentives. For example, parties work as filters for picking politicians. Whether this results in the best quality people being picked to run for office is poorly understood. There are also incentive questions concerning how parties form collective reputations and whether they have tools to discipline agents for acting against the collective interest.

There is an extensive literature on cabinets, coalitions and legislatures. However, this has not yet been embedded in an agency framework. This has a bearing on the optimal structure of decision making rules. A key issue is whether delegating to multiple decision makers worsens or improves the quality of public decision making. The examples discussed in chapter 3 suggest that this is likely to depend on the details of the institutional structure. Understanding the mapping from institutional structures into policy outcomes in this area presents a challenge for future research.

The analysis above focused mainly on the way in which political processes shape policy. However, political choice is only one dimension of the issue. Policies have to be implemented, typically by bureaucrats. Bureaucrats are typically powerful because they have a lot of information about policy implementation. This creates a multi-layered agency problem between voters, politicians and bureaucrats which has not been studied extensively in the agency model. Bureaucrat accountability is rather different from politician accountability, in part because bureaucrats can be given more explicit incentives for project delivery and are typically focused on a more narrow set of task than politicians. Bureaucrats may also be more sensitive to market based incentives. Related to the issue of political selection is that of bureaucratic selection. When does a potential public servant choose to become a politician or a bureaucrat?

A more complete book than this would delve into issues of bureaucracy as well as politics in determining good policy. Just as the old-style public choice reasoning has now been replaced by more sophisticated models of agency in politics, so there is a need to get beyond the old-fashioned budget maximizing models of bureaucracy. A full treatment of these issues would require looking at models which are sensitive to the issues that arise in very different bureaucratic environments – from the armed forces, to schools, to hospitals and town planning. The new literature in political economy still has some way to go to provide a comprehensive approach that covers all these areas.

Another missing element in the discussion above is the exercise of judicial power. As we discussed above, Maskin and Tirole (2004) have used the agency approach to contrast different methods of judicial selection and their consequences for outcomes. It is clear that much important policy making – for example competition policy – makes extensive use of judicial processes in implementing policy. How far, courts (rather than legislatures) can and should be given a policy role raises fascinating questions and one

---

3See Carillo and Mariotti (2000) for an analysis in this direction.
4See Alesina and Tabellini (2004) for a move in this direction.
dimension of good government is surely getting the mix of responsibilities right. There is scope for much more work on these issues.

The field of political economy is booming. It is now widely appreciated that incentives in government need to be understood to improve the quality of policy making. The traditional mode of public economics has been to study the pros and cons of different policies. While this has been extremely powerful in shaping policy debates, it is clear that this is only part of the story.

First, we need to understand why good policies are enacted only some of the time. Trade protection based on tariffs is a case in point. It would be useful to understand the source of political failure at work here to see if there is anyway to reform the policy process to improve the quality of government policy.

Second, many proposed reforms involve changes in institutional structures. For example, governments wish to understand the pros and cons of decentralizing policy making to lower tiers of government or creating an independent agency to carry out policy or increasing transparency in policy making. In such cases, it is necessary to work with models of the policy process in which implications for selection and incentives from institutional change can be understood.

In response to both of these cases, the field of political economy is providing the knowledge and tools that are needed to improve the workings of government.

The material in this book shapes broad thinking on the competence of government. We have begun from a position that while markets have their limits in allocating resources, so do governments. It is evident that the economics profession is now providing tools to meet the challenge of deciding where the boundaries lie between public and private responsibility. There is a section of opinion that equates good government to small government. Moreover, this has been a dominant tradition in political economy in the past. However, there is nothing in modern political economics to support this claim even if attitudes towards government intervention are more cautious than in the past. A political economy approach can also fuel optimism – if we can understand the logic of good government, then perhaps this is the first step towards creating it.
Bibliography


[154] Le Borgne, Eric and Ben Lockwood, [2001b], “Candidate Entry, Screening, and the Political Budget Cycle,” typescript.


[165] Madison, James, [1961], The Federalist Papers: A collection of essays in support of the constitution of the United States (from the original text), News York: Doubleday.


