9-4-2012

Three Essays in Comparative Political Economy

Ugur Ozdemir
Washington University in St. Louis

WASHINGTON UNIVERSITY

Department of Political Science

Dissertation Examination Committee:

Randall Calvert, Co-chair

Nathan M. Jensen, Co-chair

Steven Fazzari

Elizabeth Maggie Penn

M. Remzi Sanver

Norman Schofield

Three Essays in Comparative Political Economy

by

Ugur Ozdemir

A dissertation presented to the

Graduate School of Arts and Sciences

of Washington University in

partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

August 2012

St. Louis, Missouri

ABSTRACT

Three Essays in Comparative Political Economy

by

Ozdemir, Ugur

Doctor of Philosophy in Political Science,

Washington University in St. Louis, August, 2012.

Professor Randall Calvert and Nathan M. Jensen, Co-chairs

This thesis consists of three chapters. The first two chapters offer game theoretical models for different linkages between political parties and the citizens: programmatic linkage and clientelistic linkage. Empirical implications of these models are discussed using real world data on Turkey and USA through statistical analysis and simulation. The last chapter steps back from the specific strategies of political agents to examine in a more general context the concepts of institutional maintenance and institutional change.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the guidance and the help of several individuals who in many different ways contributed and extended their valuable assistance.

I would like to express my deepest gratitude to my advisors. Dr. Randall Calvert provided excellent guidance, patience and a vision of serious research. His truly careful attention enriched my growth as a social scientist. Throughout my thesis-writing period, Dr. Nathan Jensen provided encouragement, sound advice, great company, and lots of good ideas. I would have been lost without him.

A very special thanks goes to Dr. Norman Schofield. Working with him as a student, as a research assistant and as a coauthor was one of the greatest privileges of my life. I learnt so much from him. His extraordinarily rich intellectual world had indelible impact on my own intellectual maturation.

Dr. Elizabeth Maggie Penn was always there when I needed her. The only regret I have is that I was not able to meet her earlier in my doctoral education.

I gratefully thank Dr. Steven Fazzari for his constructive comments on this thesis. I am thankful that in the midst of all his activity, he accepted to be a member of the committee.

Dr. Remzi Sanver is much more than a committee member. It is not possible to express his importance in my academic and personal life in a paragraph. Since we met almost ten years ago, he became the single most important person in my life. His existence as a moral and intellectual figure is one of the reasons why I keep my hopes about a better world alive.

This thesis benefitted a lot from the seminar courses I took. Among them, the courses I took from Dr. Andy Sobel and Dr. Guillermo Rosas laid the foundations for two of the chapters in this dissertation. I would like to offer my special thanks to them for the critical discussions we had.

My special thanks are extended to the administrative staff of the Department of Political Science. Heather and Rachel always knew what I did not and should have known and showed me the way.

My colleagues at Istanbul Bilgi University provided great encouragement and support in my final year while I was a visiting scholar. Among them, I would like to express my special thanks to Dr. Koray Akay and Dr. Hasan Kirmanoglu who provided valuable insights about my projects in our discussions.

I am grateful to all my fellow graduate students for providing a family-like environment. Their warmth was so crucial for a student who was 11.000 miles away from home.

My friends in different parts of the world made life bearable with their support during the most stressful times. Thank you, Gazihan Alankus, Mustafa Avci, Burak

Can, Ceyhun Coban, Aydin Danaci, Emre Dogan, Mehmet Morali, Ali Ihsan Ozkes, Hadi Sahin, Ismail Sergin, and Burak Uras.

Last but not least, I would like to express my very great gratitude to my family. My parents, who have always supported, encouraged and believed in me, in all my endeavors. My parents-in-law and sister-in-law who provided valuable support throughout this period. Betul, my wife, without whom this effort would have been worth nothing. Your love, support and constant patience have taught me so much about sacrifice and compromise. And my son, Ahmet Selim, who had to spend the first 15 months of his life with a father who is writing a dissertation. I am deeply sorry for the time you had to watch Baby TV instead of being at the seaside where all your friends were having fun. This thesis is dedicated to you and your mother.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

PREFACE

I did my undergraduate study in Industrial Engineering. Although I was in top 200 in a country-wide entrance exam taken by more than 1 million students, I was quite unsuccessful as an engineering student. The reason was clear: at the end of the first semester it had turned out that I did not like what I study. I wanted to be a social scientist. Who was to blame? After thinking and reading for quite a long time I had found the answer: institutions ! The education institutions which were unsuccessful in discovering what I wanted and directing me in that direction, and my family, as an informal institution, who always motivated me for the highest score in the exams, not the highest level of happiness. I was lucky though. Another previously unhappy engineer who had become a very successful social scientist, M. Remzi Sanver, held my hand and I wrote a M.Sc thesis about Arrow's Theorem. At the end of that two years, it was all clear to me: mathematical social science was the shortest possible summary of what I would like to do in the rest of my life and institutions will have to be in the picture. And I ended up at one of the best possible places for this purpose.

Starting the PhD program in Washington University was scary (I had not taken a single Political Science course before) but very exciting. With Douglas North living upstairs, Randall Calvert and Norman Schofield living next door, it was a heaven for me. I still remember how I got excited when Dr. Schofield was talking about why financial crises was a singularity.

Apart from providing me a distinguished education and research vision, my department provided a family-like environment and I spent 5 great years in St.Louis. Finally, I was starting to feel that what I want to do and what I do are the one and the same thing. This thesis is a focal point in this long effort, which took me almost ten years to accomplish.

1. Introduction

What links citizens to political parties or candidates? The democratic theorists argue that citizens vote for a platform hence between political candidates in order to exercise political authority make rulers accountable to the citizens [Kitschelt and Wilkinson, 2007b]. This responsible party model ignores a quite a different type of party-voter linkage which exists even in the advanced industrial democracies: the clientelistic linkage. The most common way of establishing this linkage is through direct material benefits transfered to voters who trade their votes for a price. The first two chapters of this thesis offer theoretical models for these two different linkages between political parties and the citizens and analyzes their empirical implications through statistical analysis and simulation. These two chapters focus on the possibility and the structure of equilibria -hence stability- in political systems. The last chapter steps back from the specific strategies of a given party to examine in a more general context the concepts of institutional maintenance and institutional change.

The spatial model of voting has been the dominant approach to electoral politics since it was first introduced [Downs, 1957]. In this model voters have preferences over some salient policy dimensions. Office seeking parties and politicians announce electoral platforms they commit to enact if elected, and voters opt for the program which they are closest to. The important theoretical question is than that given a set of voters, what should we expect to observe in terms policy choices of the political candidates? The theoretical prediction is that the political candidates should converge to the electoral mean [Schofield, 2004]. This result, known as the median voter theorem, however is not in accord with the empirical observations [Schofield and Sened, 2006].

One way of explaining this contradictory nature focuses on the concept of valence. It was Stokes, 1963 who introduced the concept of valence in a very famous critique. In a later work, Stokes, 1992 defines valence issues as issues, on which parties or leaders are differentiated not by what they advocate but by the degree to which they are linked in the publics mind with conditions or goals or symbols of which almost everyone approves or disapproves". The consequence of introducing the valence term in the analysis is that when all candidates choose the mean of the electoral distribution as their position, they are not going to be treated as identical by the voters. Hence the convergence to the mean will not be an optimum vote maximizing strategy for all candidates.

Since Stokes' original work, different interpretations of the valence concept have emerged. While some scholars, staying closer to Stokes' original definition, equate the term with policy related factors [Budge and Farlie, 1983, Bélanger and Meguid, 2008], others used it to refer to nonpolicy related factors [Fiorina, 1978, Enelow and Hinich, 1982, Alvarez and Nagler, 1995, Grose, 2005, Clark, 2009]. In the former approach, voters evaluate the candidates on the basis of competence in handling issues rather than their platforms on the issue space. In the latter approach, valence refers to voters judgements about positively or negatively evaluated aspects of candidates, or party leaders, which cannot be ascribed to the policy choice of the party or the candidate.

Chapter 1[1] uses a stochastic electoral model where party leaders or candidates are differentiated by differing valences— the electoral perception of the quality of the party leader. If valence is simply intrinsic, then it can be shown that there is a convergence coefficient, defined in terms of the empirical parameters, that must be bounded above by the dimension of the space, in order for the electoral mean to be a Nash equilibrium. This model is applied to elections in Turkey in 1999 and 2002.

Chapter 2 extends the model in Chapter 1 to handle endogenous valences in order to provide a theoretical framework for the clientelistic linkage based on distribution of materialistic benefits. This chapter introduces a spatial model which formalizes clientelistic linkage as valence-buying phenomenon. The model allows for the existence of more than one machine party. Moreover the policy choice of the political parties and their decision regarding to whom to bribe and how much are simultaneously determined in the equilibrium subject to their budget constraints. This approach opens the way to many interesting comparative statics which in fact do have implications outside the vote-buying literature, such as the controversial subject of bribing in international organizations. Using simulation techniques we illustrate how budget differentials among parties affect the political equilibrium.

[1]This chapter is co-authored with Norman Schofield. The author is solely responsible for data collection, statistical analysis, calculations, coding and simulations. Sections on Turkey politics are co-written.

The first two chapters studies the possibility and the structure of equilibria - hence stability- in political systems. Change however is a natural characteristic of any sociopolitical system. In particular, while a party organization persists, it pursues certain policy positions (and a pattern of changes in those positions), partly at the behest of its particular activists and partly in reaction to the moves of other parties. From time to time, however, those organizations or coalitions break up and recombine, a classic instance of institutional change. Chapter 3 offers a model of institutional change and persistence as a generalization of this phenomenon.

2. A Stochastic Model of Elections in Turkey[1]

2.1 Introduction

The early work in modeling elections focused on two-party competition, and assumed a one-dimensional policy space, X, and "deterministic" voter choice. The models showed the existence of a "core" point, unbeaten under majority rule vote, at the median of the electoral distribution. Such models implied that there would be strong centripetal political forces causing parties to converge to the electoral center [Hotelling, 1929, Downs, 1957]. In higher dimensions, such two party "pure strategy Nash equilibria" (PNE) generally do not exist, so the theory did not cover empirical situations where two or more policy dimensions were relevant.[2] It has been shown, however, that there would exist *mixed* strategy Nash equilibria whose support lies within a subset of the policy space known as the "uncovered set."[3] "Attractors" of the political process, such as the "core", the "uncovered set" or the "heart" [Schofield, 1999] are centrally located with respect to the distribution of voters' ideal points. The theoretical prediction that political candidates converge to the center is very much at

[1] This chapter is co-authored with Norman Schofield. The author is solely responsible for data collection, statistical analysis, calculations, coding and simulations. Sections on Turkey politics are co-written. Some findings in this chapter have appeared in [Schofield et al., 2009].

[2] See [Saari, 1997] and the survey in [Austen-Smith and Banks, 2000].

[3] [Banks and Duggan, 2006].

odds with empirical evidence from U.S. presidential elections that political candidates do not locate themselves close to the electoral center.[4]

The deterministic electoral model is also ill-suited to deal with the multiparty case. (Here multiparty refers to the situation where the number of candidates or parties, p, is at least three.) As a result, recent work has focused on "stochastic" models which are, in principle, compatible with empirical models of voter choice.[5] In such models, the behavior of each voter is modeled by a vector of choice probabilities. Various theoretical results for this class of models suggested that vote maximizing parties would converge to the mean of the electoral distribution of voter ideal points.[6]

Empirical estimates of party positions in European multiparty polities can be constructed on the basis of various techniques of content analysis of party manifestos.[7] More recent analyses have been based on factor analysis of electoral survey data to obtain a multidimensional description of the main political issues in various countries. All these empirical analyses have obtained policy spaces that are two dimensional. These techniques allow for the estimation of the positions of the parties in the empirically inferred policy space. These estimates have found no general tendency for parties to converge to the center.[8]

The various empirical electoral models can be combined with simulation techniques to determine how parties should respond to electoral incentives in order to

[4] [Poole and Rosenthal, 1984]; [Schofield et al., 2003]. See also the empirical work in [Schofield et al., 2011].

[5] [Schofield and Sened, 2006].

[6] [Hinich, 1977]; [Lin et al., 1999]; [Banks and Duggan, 2005]; [McKelvey and Patty, 2006].

[7] See [Laver and Hunt, 1992]. [Benoit and Laver, 2006] use expert estimates.

[8] [Adams and Merril III, 1999], for example.

maximize their vote shares. [Schofield and Sened, 2006], in their simulation of elections in Israel in the period 1988 to 1996, found that vote maximizing parties did not converge to the electoral origin. It may be objected that factor analysis of survey data gives only a crude estimate of the variation in voter preferences, while vote maximization disregards the complex incentives that parties face. Nonetheless, as a modeling exercise, the stochastic model for Israel seemed to provide a plausible account of the nature of individual choice[9] as well as the party positioning decision. Although the simulated equilibrium positions of the parties in Israel were not identical to the estimated positions, the positions were generally far from the origin, and for some of the parties very close to their estimated positions. The purpose of this paper is to attempt to extend the stochastic empirical model so as to close the apparent disparity between the simulated equilibrium positions of the parties, and the estimated positions.

The key to the contradiction between the non-convergence result of Schofield and Sened, and the convergence result in other work on the formal stochastic model was the incorporation of an asymmetry in the perception of the quality of the party leaders, expressed in terms of *valence* [Stokes, 1992, Stokes, 1963].

In the model presented here, the average weight, in the voter calculus, given to the perceived quality of the leader of the j^{th} party is called the party's *intrinsic or exogenous valence.* In empirical models this valence is assumed to be exogenous, so it is independent of the party's position. The valence coefficients for each party are

[9] Over 60% of the individual votes were correctly modeled.

generated by the estimation of the stochastic model, based on the "multinomial logit" (MNL) assumption that the stochastic errors have a "Type I extreme value or Gumbel distribution" [Dow and Endersby, 2004]. These valence terms add to the statistical significance of the model. In general, valence reflects the overall degree to which the party is perceived to have shown itself able to govern effectively in the past, or is likely to be able to govern well in the future. [10]

Appendix A considers a pure spatial stochastic vote model,with party specific exogenous valences, based on the same distribution assumption, and on the assumption that each party leader attempts to locally maximize the party's vote share. Results from [Schofield and Miller, 2007], [Schofield, 2006] give the necessary and sufficient conditions under which there is a "local pure strategy Nash equilibrium" (LNE) of this model at the *joint electoral mean* (that is, where each party adopts the same position, z_0, at the mean of the electoral distribution).[11] Theorem 2 in Appendix A shows that a "convergence coefficient", c, incorporating all the parameters of the model, can be defined. This coefficient, c, involves the differences in the valences of the party leaders, and the "spatial coefficient" β. When the policy space, X, is

[10]See [Penn, 2009]. Notice, however, that valence refers to the *perception* by voters of the quality of political leaders. Recent work by [Westen, 2007], for the United States, suggests that voters' perceptions of the characteristics of political candidates are very important. Moreover, [Schofield et al., 2009] shows that voter perception of character traits has a strong effect on candidate positions in the United States.

[11]A local Nash equilibrium under vote maximization is just a vector of positions such that no small unilateral move by a party can increase its vote. The usual notion of a pure strategy Nash equilibrium (PNE) cannot be used because in the games we study there may exist no PNE.

assumed to be of dimension w, then the necessary condition for existence of an LNE at the electoral center[12] is that the coefficient, c, is bounded above by w.

When the necessary condition fails, then parties, in equilibrium, will adopt divergent positions. Because a pure strategy Nash equilibrium must be an LNE, the failure of existence of LNE when all parties are at the electoral mean implies non existence of such a centrist PNE. In this case, a party whose leader has the lowest valence will have the greatest electoral incentive to move away from the electoral mean. As the party moves away from the electoral mean, it increases the probability that voters on the electoral periphery will vote for it.[13] Other low valence parties will follow suit, and the local equilibrium will be one where parties are distributed along a "principal electoral axis."[14] The general conclusion is that, with all other parameters fixed, then a convergent LNE can be guaranteed only when β is "sufficiently" small. Thus, divergence away from the electoral mean becomes more likely the greater are β, the valence differences and the variance of the electoral distribution.[15]

The innovation of this paper is that in additional to exogenous valence, we also incorporate "sociodemographic valence." These party specific valence terms are associated with different groups in the society, and are defined by dichotomous or contin-

[12] Again, the electoral center, or origin, is defined to be the mean of the distribution of voter ideal points.

[13] This follows for theoretical reasons as shown in [Schofield, 2006]. When $c > w$, at least one of the eigenvalues of the Hessian of the vote share function of a low valence party will be large and positive at the origin. As it moves from the origin, it will lose votes from centrist voters, but gain votes from more radical voters. Simulation of empirical models for Israel [Schofield and Sened, 2006] has shown this to be the case.

[14] The principal electoral axis is defined to be the one dimensional subspace along which the variance of the distribution of voter ideal points is maximum.

[15] These results are presented for the reader's convenience in the context of the more general model described in Appendix A.

Party Name	.	% Vote	Seats	% Seats
Democratic Left Party	DSP	22.19	136	25
Nationalist Action Party	MHP	17.98	129	23
Virtue Party	FP	15.41	111	20
Motherland Party	ANAP	13.22	86	16
True Path Party	DYP	12.01	85	15
Republican People's Party	CHP	8.71	—	—
People's Democracy Party	HADEP	4.75	—	—
Others	—	4.86	—	—
Independents	—	0.87	3	1
Total			550	

Table 2.1
Turkish Election Results 1999

uous characteristics of different subgroups in the population. This model is shown to be statistically superior to the spatial model with exogenous valence. This is the case because the exogenous valence model assumes that all voters have the same perception of the quality of the party leaders, whereas with the sociodemographic variables, these perceptions are allowed to vary across different subgroups.

We apply this valence model by considering in some detail a sequence of elections in Turkey from 1999 to 2007. The election results are given in Tables 2.1, 2.2 and 2.3, which also provide the acronyms for the various parties.

As in other related work, the empirical models were based on factor analyses of voter surveys.[16] Figures 2.1 and 2.2 show the electoral distributions (based on sample

[16]The estimations presented below are based on factor analyses of sample surveys conducted by Veri Arastima for TUSES.

Party Name		% Vote	Seats	% Seats
Justice and Development Party	AKP	34.28	363	66
Republican People's Party	CHP	19.39	178	32
True Path Party	DYP	9.54	—	—
Nationalist Action Party	MHP	8.36	—	—
Young Party	GP	7.25	—	—
People's Democracy Party	HADEP	6.22	—	—
Motherland Party	ANAP	5.13	—	—
Felicity Party	SP	2.49	—	—
Democratic Left Party	DSP	1.22	—	—
Others and Independents	—	6.12	9	2
	—			
Total			550	

Table 2.2
Turkish Election Results 2002

Party Name		% Vote	Seats	% Seats
Justice and Development Party	AKP	46.6	340	61.8
Republican People's Party	CHP	20.9	112	20.3
Nationalist Movement Party	MHP	14.3	71	12.9
Democrat Party	DP	5.4	—	—
Young Party	GP	3.0	—	—
Felicity Party	SP	2.3	—	—
Independents	—	5.2	27	4.9
Others	—	2.3	—	—
Total		100	550	100

Table 2.3
Turkish Election Results 2007

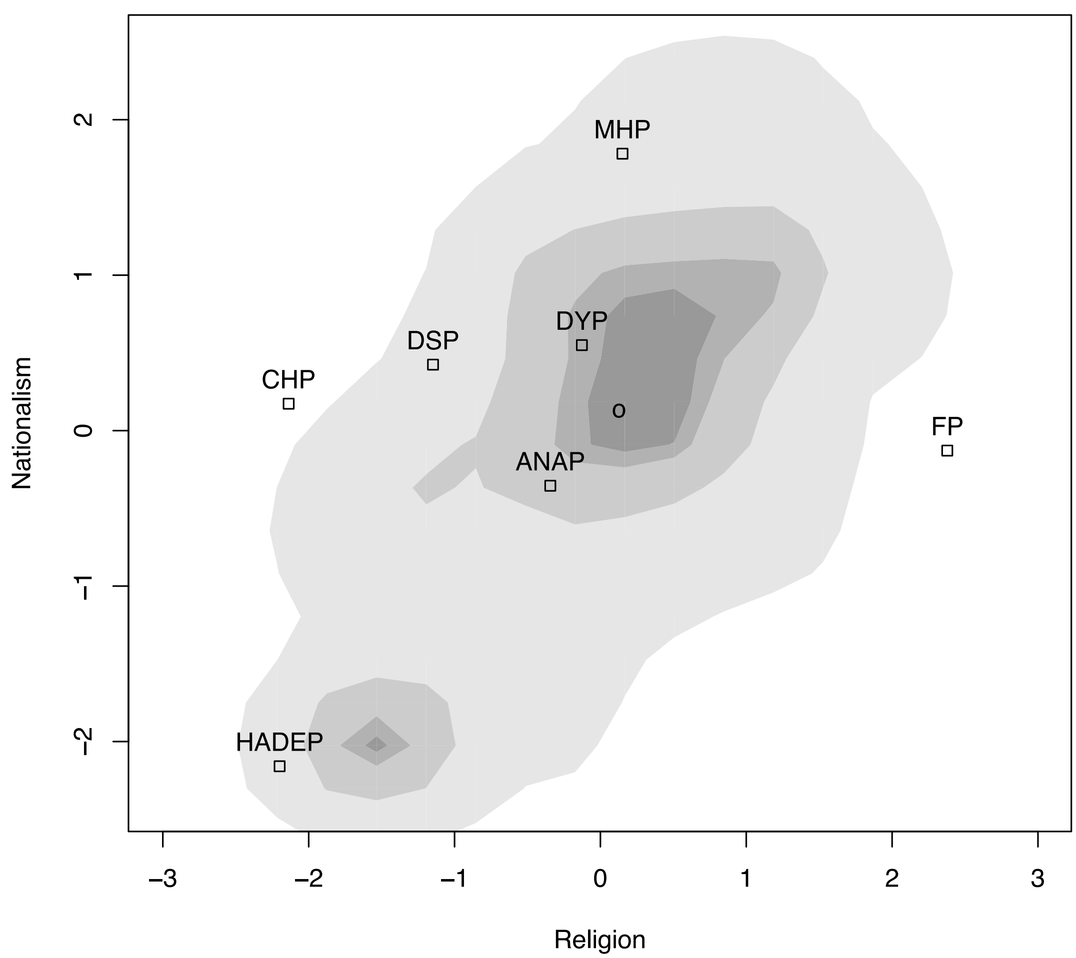

Figure 2.1. Party positions and voter distribution in Turkey in 1999

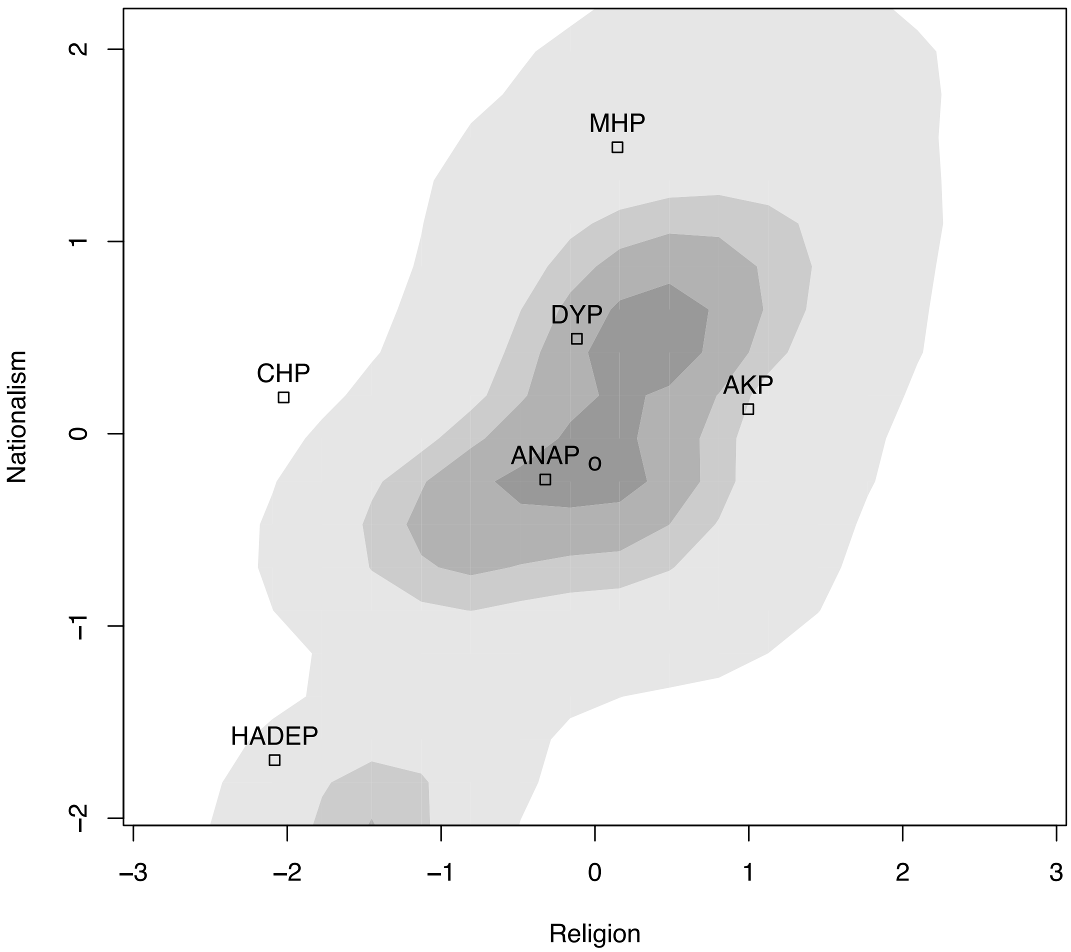

Figure 2.2. Party positions and voter distribution in Turkey in 2002

surveys of sizes 635 and 483 respectively) and estimates of party positions for 1999 and 2002.[17] The two dimensions in both years were a "left-right" religion axis and a "north-south" Nationalism axis, with secularism or "Kemalism" on the left and Turkish nationalism to the north. (See also [Carkoglu and Hinich, 2006] for a spatial model of the 1999 election).

Minor differences between these two figures include the disappearance of the Virtue Party (FP) which was banned by the Constitutional Court in 2001, and the change of the name of the pro-Kurdish party from HADEP to DEHAP.[18] The most important change is the appearance of the new Justice and Development Party (AKP) in 2002, essentially substituting for the outlawed Virtue Party.

In 1999, a DSP minority government formed, supported by ANAP and DYP. This only lasted about 4 months, and was replaced by a DSP-ANAP-MHP coalition, indicating the difficulty of negotiating a coalition compromise across the disparate policy positions of the coalition members. During the period 1999–2002, Turkey experienced two severe economic crises. As Tables 2.1 and 2.2 show, the vote shares of the parties in the governing coalition went from about 53 percent in 1999 to less than 15 percent in 2002. In 2002, a 10% cut-off rule was instituted. As Table 2.3 makes clear, seven parties obtained less than 10% of the vote in 2002, and won no seats. The AKP won 34% of the vote, but because of the cut-off rule, it obtained a majority of the seats (363 out of 550). In 2007, the AKP did even better, taking about

[17]The party positions were estimated using expert analyses, in the same way as the work by [Benoit and Laver, 2006].

[18]For simplicity, the pro-Kurdish party is denoted HADEP in the various Figures and Tables. Notice that the HADEP position in Figures 1 and 2 is interpreted as secular and non-nationalistic.

46% of the vote, against 21% for the CHP. The Kurdish Freedom and Solidarity Party avoided the 10% cut-off rule, by contesting the elections as independent non-party candidates, winning 24 seats with less that 5% of the vote.

The point of this example is that a comparison of Figures 2.1 and 2.2 suggest that there was very little change in policy positions of the parties between 1999 and 2002. The basis of support for the AKP may be regarded as a similar to that of the banned FP, which suggests that the leader of this party changed the party's policy position on the religion axis, adopting a much less radical position.

In sum, the standard spatial model is unable to explain the change in the electoral outcome, taken together with the relative unchanged positioning of the parties between 1999 and 2002.

Section 2 of the chapter considers the details of the multinomial logit (MNL) model for Turkey for 1999 and 2002. In particular, this section shows that the pure spatial model with exogenous valence predicts that the parties diverge away from the origin. To illustrate, Table 5 shows that the lowest valence party in 2002 was the Motherland Party (ANAP) while the Republican People's Party (CHP) had the highest valence. The convergence coefficient was computed to be 5.94, far greater than the upper bound of 2. Figure 2.3 presents an estimate of one of the LNE obtained from simulation of vote maximizing behavior of the parties, under the assumption of the pure spatial model with exogenous valence. As expected from the theoretical result, the LNE is non centrist. Note however, the the LNE positions for the pure spatial model given in Figure 3 are quite different from the estimated positions in Figure 2.2.

To improve the prediction of the model, we incorporated the sociodemographic variables. Estimating the LNE for this sociodemographic model gave a better prediction. To explain the difference between the estimated positions of the parties, and the LNE from the sociodemographic model, we then added the the influence of party activists to the model. Since sociodemographic variables can be interpreted as specific valences associated with different subgroups of the electorate, we can use these sociodemographic valences to estimate the influence of group-specific activists on party positions.

Theorem 1 in Appendix A[19] gives the first order *balance* condition for local equilibrium in the stochastic electoral model involving sociodemographic valences and activists. The condition requires the balancing of a *centrifugal marginal activist pull (or gradient)* against a *marginal electoral pull.* In general, if the exogenous valence of a party leader falls, then the marginal electoral pull also falls, so balance requires that the leader adopt a position closer to the preferred position of the party activists.

The pure spatial model, with exogenous valences, and a joint model, with sociodemographic valences, but without activists, are compared using simulation to determine the LNE in these models. This allows us to determine which model better explains the party positions. For example, Figure 2.4 shows the LNE based on a joint sociodemographic model for 2002. In this figure, the LNE position for the Kurdish

[19]The results in the Appendix A extend the version of the activist model originally proposed by [Aldrich, 1983] and developed in [Schofield, 2006].

party, HADEP, is a consequence of the high electoral pull by Kurdish voters located in the lower left of the figure. Similarly, the position of the CHP on the left of the figure is estimated to be due to the electoral pull by Alevi voters who are Shia, rather than Sunni and can be regarder as supporters of the secular state. Although Figure 2.4 gives a superior prediction of the party positions than Figure 2.3, there is still a discrepancy between the estimated positions of Figure 2.2 and the LNE in Figure 2.4. We argue that the difference between these two vectors of party positions, as presented in Figures 2.2 and 2.4, can be used to provide an estimation of the marginal activist pulls influencing the parties.

More generally, we suggest that the combined model, with sociodemographic variables and activists, can be used as a tool with which to study the political configuration of such a complex society. In the conclusion we suggest that the full model involving activists may be applicable to the study of what [Epstein et al., 2006] call "partial democracies", where a political leader must maintain popular support, not just by winning elections, but by maintaining the allegiance of powerful activist groups in the society.

2.2 Elections in Turkey 1999-2007

Appendix A defines an empirical electoral model, denoted $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta; \mathbf{V})$ which utilizes socio-demographic variables, denoted θ.

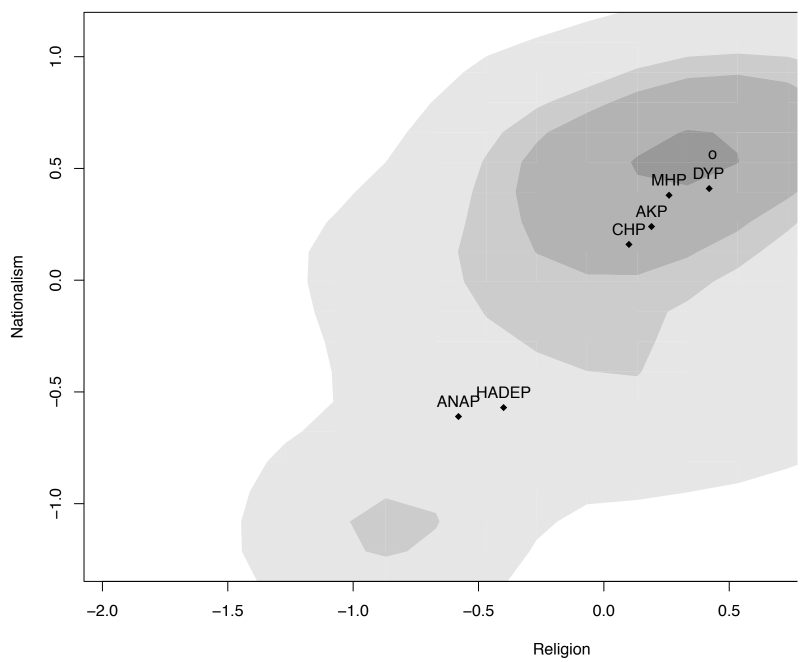

Figure 2.3. A Local Nash Equilibrium for the joint model in 2002

The symbol,$\mathbf{V}$, denotes a family of egalitarian vote functions, one for each party, and under which all voters are counted equally. The formal model of Appendix A considers a more general class of vote functions where the voters vary in their weights, thus allowing for complex electoral rules. In Appendix A, the egalitarian family is denoted $\mathbf{V}_e$. The symbol, Ψ, denotes the Gumbel stochastic distribution on the errors. To simplify notation in the applications that follow we delete reference to $\mathbf{V}$ and Ψ.

This empirical model asumes that the utility function of voter i is given by the expression

$$u_{ij}(x_i, z_j) = \Lambda_j + (\theta_j \centerdot \eta_i) - \beta\|x_i - z_j\|^2 + \varepsilon_j.$$

Here, the *spatial coefficient* is denoted β and $\mathbf{\Lambda}$ =$\{\Lambda_j : j \in P\}$ are the *exogenous valences* (relative to a baseline party,k^*).[20] The relative *exogenous* valence, Λ_j, gives the average belief of the voters in the electorate concerning the quality of the leader of party j in comparison to the leader of the baseline party, k^*. The symbol, θ, denotes a set of m—vectors $\{\theta_j\}$ representing the effect of the m different sociodemographic parameters (class, domicile, education, income, religious orientation, etc.) on the beliefs of the various subgroups in the polity on the competence of party j. The symbol η_i is an m-vector denoting the i^{th} individual's relevant "sociodemographic" characteristics. The composition $(\theta_j \centerdot \eta_i)$ is the scalar product and can be interpreted as the group specific valence ascribed to party j as a consequence of the various

[20]Note that in the empirical models discussed below, these are specified relative to the baseline party, the DYP.

sociodemographic characteristics of voter i. Again, these sociodemographic variables will be normalized with respect to the baseline party k^*, essentially by estimating $((\theta_j - \theta_{k^*}) \cdot \eta_i)$. This scalar term is called the *total sociodemographic valence* of voter i for party j. The t^{th} term in this scalar is called the *sociodemographic valence* of i as a result of membership by i of the t^{th} group, or, more briefly, the t^{th} group specific sociodemographic valence for the leader of party j.[21]

The vector $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ is the set of party positions, while $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$ is the set of ideal points of the voters in N. When β is assumed zero then the model is called pure *sociodemographic* (SD), and denoted $\mathbb{M}(\mathbf{\Lambda}, \theta)$. When $\{\theta_j\}$ are all assumed zero then the model is called *pure spatial*, and denoted $\mathbb{M}(\mathbf{\Lambda}, \beta)$. The pure spatial model implicitly assumes that the ranking over valence is identical among voters.The empirical model, $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta)$, including the sociodemographic terms is called *joint*. These socio-demographic variables allow us to incorporate characteristics common to specific groups of supporters of any party, and this permits the valence ranking to vary among voters in a way which depends on sociodemographics. Not accounting for these characteristics in the analysis will bias the estimates of the exogenous valences of the parties.

Tables 2.4 and 2.5 give the details of the pure spatial MNL models for the elections of 1999 and 2002 in Turkey, while Tables 2.6 and 2.7 give the details of the joint MNL

[21] For example, in Table 6 and 7 there are 6 sociodemographic variables, so $m = 6$. An individual who is Alevi has $\eta_{i,Alevi} = 1$. The coefficient for the CHP party for an Alevi is 3.089 in 1999, and this is the group-specific valence that a voter who is a member of the group of Alevi voters has for this party. Note again that this is specified relative to the baseline party, the DYP. These valences may be the result of the perception of the leader's ability, as displayed in the past, or of the particular partiality of these voters to choose the party, independently of the party's policy position.

Party Name		Λ_k	Std.error	t-value
Democratic Left Party	DSP	0.724	0.153	4.73*
Nationalist Action Party	MHP	0.666	0.147	4.53*
Virtue Party	FP	-0.159	0.175	0.9
Motherland Party	ANAP	0.336	0.153	2.19
True Path Party	DYP	—	—	—
Republican People's Party	CHP	0.734*	0.178	4.12*
People's Democracy Party	HADEP	-0.071	0.232	0.3
(Normalized with respect to DYP)				
Spatial Coefficient β	0.375*		0.088	4.26*
Convergence Coefficient c	1.49		0.22	6.77*
$n = 635$				
Log marginal likelihood (LML) $= -1183$				
*=Significant with probability < 0.001.				

Table 2.4
Pure Spatial Model of the Turkish Election 1999

models. The differences in log marginal likelihoods for the three different models then gives the log Bayes' factor for the pairwise comparisons.[22] The log Bayes' factors show that the joint and pure spatial MNL models were clearly superior to the SD models. In addition the joint models were superior to the pure spatial models.[23] We can infer that, though the sociodemographic variables are useful, by themselves they do not give an accurate model of voter choice.[24] It is necessary to combine the pure spatial model, including the valence terms, with the sociodemographic valences to obtain a superior estimation of voter choice.

[22]Since the Bayes' factor [Kass and Raftery, 1995] for a comparison of two models is simply the ratio of marginal likelihoods, the log Bayes' factor is the difference in log likelihoods.

[23]The log Bayes factors for the joint models over the sociodemographic models were highly significant at +31 in 1999 and +58 in 2002. The Bayes' factors for the joint over the spatial models were also significant, and estimated to be +6 and +5 in 1999 and 2002, respectively.

[24]Sociodemographic models are standard in the empirical voting literature.

Party Name		Λ_k	Std. error	t-value
Justice and Development Party	AKP	0.78	0.15	5.2*
Republican People's Party	CHP	1.33	0.18	7.4*
True Path Party	DYP	—	—	—
Nationalist Action Party	MHP	-0.12	0.18	0.66
Young Party	GP	—	—	—
People's Democracy Party	HADEP	0.43	0.21	2.0
Motherland Party	ANAP	-0.31	0.19	1.63
(Normalized with respect to DYP)		—	—	—
Spatial Coefficient β	1.52*		0.12	12.66*
Convergence Coefficient c	5.94*		0.27	22.0*
$n = 483$				
Log marginal likelihood (LML) $= -737$				
*=Significant with probability < 0.001.				

Table 2.5
Pure Spatial Model of the Turkish Election 2002

Comparing Tables 2.4 and 2.5, it is clear that the relative valences of the ANAP and MHP, under the pure spatial model, dropped between 1999 and 2002. In 1999, the estimated Λ_{ANAP} was +0.336, while the confidence interval on Λ_{ANAP} for 1999 in Table 4 shows that the hypothesis that $\Lambda_{ANAP} = 0$ should be rejected. In contrast the estimated value of Λ_{ANAP} for 2002 was −0.31, and the confidence interval on Λ_{ANAP} does not allow us to reject the hypothesis that $\Lambda_{ANAP} = 0$. [25] Similarly Λ_{MHP} fell from a significant value of +0.666 in 1999 to −0.12 in 2002. The estimated relative valence, Λ_{AKP}, of the new Justice and Development Party (AKP) in 2002 was +0.78, in comparison to the valence of the FP of −0.159 in 1999. Since the AKP can be regarded as a transformed FP, under the leadership of Recep Tayyip Erdogan, we can infer from the confidence intervals on these two relative valences that this was a significant change due to Erdogan's leadership.[26]

It should be noted that the β coefficients for the pure spatial models were 0.375 in 1999, and 1.52 in 2002. Both of these are estimated to be non-zero at the 0.001 level. Indeed, they are significantly different from each other,[27] suggesting that electoral preferences over policy had become more intense.

We first use the results of the formal pure spatial model given in the Appendix A to compute estimates of the convergence coefficients. These computations suggest that convergence to an electoral center is not to be expected in these elections. We then

[25]These tables show the standard errors of the coefficients, as well as the *t-values*, the ratios of the estimated coefficient to the standard error.

[26]Although Erdogan was the party leader, Abdullah Gul became Prime Minister after the November 2002 election because Erdogan was banned from holding office. Erdogan took over as Prime Minister after winning a by-election in March 2003.

[27]The 95% confidence interval for β_{1999} is [0.2,0.55] and for β_{2002} it is [1.28,1.76]

		Coefficient		95% Confidence Interval	
Variable	Party	Est	Std Err	Lower Bound	Upper Bound
Spatial Coeff. β		0.456	0.104	0.243	0.648
Relative Valence Λ_k	ANAP	-0.114	0.727	-1.513	1.227
	CHP	-0.673	0.770	-2.166	0.786
	DSP	0.463	0.720	-0.930	1.825
	FP	1.015	0.878	-0.709	2.755
	HADEP	-0.610	1.230	-3.004	1.803
	MHP	2.447	0.669	1.167	3.664
Age	ANAP	0.001	0.012	-0.021	0.023
	CHP	-0.009	0.013	-0.033	0.016
	DSP	-0.008	0.012	-0.031	0.014
	FP	-0.023	0.014	-0.050	0.003
	HADEP	-0.053	0.023	-0.103	-0.014
	MHP	-0.044	0.012	-0.067	-0.022
Education	ANAP	0.006	0.065	-0.115	0.130
	CHP	0.106	0.063	-0.012	0.232
	DSP	0.077	0.058	-0.024	0.197
	FP	-0.129	0.081	-0.285	0.018
	HADEP	0.144	0.097	-0.038	0.335
	MHP	-0.060	0.061	-0.175	0.070
Kurd	ANAP	1.132	0.924	-0.410	3.138
	CHP	1.715	0.911	0.194	3.637
	DSP	-0.102	1.083	-2.650	2.098
	FP	1.116	0.972	-0.733	3.024
	HADEP	5.898	0.926	4.290	7.904
	MHP	0.063	0.933	-1.751	2.148
Soc. Econ. Status	ANAP	0.080	0.165	-0.302	0.394
	CHP	0.163	0.176	-0.195	0.499
	DSP	-0.010	0.158	-0.322	0.333
	FP	0.120	0.179	-0.230	0.458
	HADEP	-0.119	0.264	-0.598	0.384
	MHP	0.168	0.159	-0.147	0.469
Alevi	ANAP	-0.697	0.972	-2.687	1.168
	CHP	3.089	0.693	1.965	4.715
	DSP	0.934	0.729	-0.383	2.423
	FP	0.346	0.939	-1.374	2.007
	HADEP	1.355	0.972	-0.332	3.605
	MHP	-0.873	0.925	-3.225	0.676
n=635			Log marginal likelihood = -1178		

Table 2.6
Joint Model of the 1999 Election in Turkey

	Coefficient			95% Confidence Interval	
Variable	Party	Est	Std Dev	Lower Bound	Upper Bound
Spatial Coeff β		1.445	0.143	1.180	1.723
Valence Λ_k	AKP	1.968	0.667	0.708	3.432
	CHP	1.103	0.797	-0.579	2.615
	HADEP	2.596	1.246	-0.254	5.049
	MHP	1.714	0.889	-0.021	3.426
	ANAP	-0.567	0.880	-2.487	1.133
Age	AKP	-0.031	0.011	-0.052	-0.010
	CHP	-0.019	0.013	-0.045	0.005
	HADEP	-0.060	0.024	-0.110	-0.014
	MHP	-0.067	0.017	-0.103	-0.034
	ANAP	-0.004	0.014	-0.031	0.022
Education	AKP	-0.070	0.062	-0.185	0.045
	CHP	-0.007	0.068	-0.136	0.115
	HADEP	-0.142	0.108	-0.365	0.079
	MHP	-0.048	0.079	-0.202	0.106
	ANAP	-0.078	0.076	-0.237	0.064
Kurd	AKP	2.086	1.105	0.203	4.596
	CHP	1.251	1.171	-0.891	3.839
	HADEP	5.996	1.208	3.960	8.945
	MHP	1.595	1.312	-0.960	4.258
	ANAP	1.603	1.199	-0.535	4.358
Soc. Econ. Status	AKP	0.142	0.160	-0.160	0.457
	CHP	0.198	0.191	-0.196	0.560
	DEHAP	-0.217	0.281	-0.755	0.301
	MHP	0.317	0.204	-0.083	0.703
	ANAP	0.214	0.209	-0.182	0.613
Alevi	AKP	-0.249	0.983	-2.125	1.743
	CHP	2.567	0.817	1.111	4.489
	DEHAP	0.377	1.045	-1.519	2.540
	MHP	-0.529	1.410	-3.565	2.292
	ANAP	1.392	0.931	-0.323	3.560
n=483			Log marginal likelihood = -732		

Table 2.7
Joint Model of the 2002 Election in Turkey

use simulation to determine the LNE of the empirical joint models, again showing non-convergence. This allows us to obtain information about activist support for the parties.

2.2.1 The 2002 Election

Figure 2.3 shows the smoothed estimate of the voter ideal points in 2002. This distribution gives the 2 by 2 voter covariance matrix, with an electoral variance on the first axis (religion) estimated to be 1.18 while the electoral variance on the second axis (nationalism) was 1.15. The total electoral variance was $\sigma^2 = 2.33$, with an electoral standard deviation of $\sigma = 1.52$ The covariance between the two axes was equal to 0.74.

Thus the voter covariance matrix is

$$\nabla_0 = \begin{bmatrix} 1.18 & 0.74 \\ 0.74 & 1.15 \end{bmatrix}$$

with $trace(\nabla_0) = 2.33$.

The eigenvalues of this matrix are 1.9, with major eigenvector $(+1.0, +0.97)$ and 0.43, with minor eigenvector $(-0.97, +1.0)$. The major eigenvector corresponds to the *principal electoral axis*, aligned at approximately 45 degrees to the religion axis.

For the pure spatial model $\mathbb{M}(\mathbf{\Lambda},\beta)$, the β coefficient was 1.52, The valence terms are estimated in contrast with the valence of the DYP, and the the party with the lowest relative valence is ANAP with $\Lambda_{ANAP} = -0.31$. By definition, $\Lambda_{DYP} = 0$. The vector of relative valences is then

$$\begin{aligned} & (\Lambda_{ANAP}, \Lambda_{MHP}, \Lambda_{DYP}, \Lambda_{HADEP}, \Lambda_{AKP}, \Lambda_{CHP}) \\ = \quad & (-0.31, -0.12, 0.0, 0.43, 0.78, 1.33). \end{aligned}$$

When all parties are at the origin, the probability, ρ_{ANAP}, that a voter chooses ANAP, in the model $\mathbb{M}(\boldsymbol{\Lambda},\beta)$, is independent of the voter. Appendix A, equation (7), shows that this is given by

$$\begin{aligned} & \frac{\exp(-0.31)}{\exp(-0.31) + \exp(-0.12) + \exp(0.0) + \exp(0.43) + \exp(0.78) + \exp(1.33)} \\ = \quad & [1 + \exp(0.19) + \exp(0.31) + \exp(0.74) + \exp(1.09) + \exp(1.164)]^{-1} \end{aligned}$$

$$\begin{aligned} = \quad & [1 + 1.2 + 1.36 + 2.09 + 2.97 + 3.2]^{-1} \\ = \quad & 0.08. \end{aligned}$$

Below, we show that the 95% confidence interval on ρ_{ANAP} is $[0, 05, 0.11]$, which includes the actual vote share (5.13%) in 2002.

Appendix A shows that the Hessian of the vote share function of ANAP, when all parties are at the origin, is given by the characteristic matrix of ANAP:

$$
\begin{aligned}
C_{ANAP} &= 2\beta(1-2\rho_{ANAP})\nabla_0 - I \\
&= 2\times(1.52)\times[(1-(2\times 0.08)]\nabla_0 - I \\
&= (2.55)\begin{bmatrix} 1.18 & 0.74 \\ 0.74 & 1.15 \end{bmatrix} - I \\
&= \begin{bmatrix} 2.01 & 1.88 \\ 1.88 & 1.93 \end{bmatrix}.
\end{aligned}
$$

Moreover, the convergence coefficient,

$$c = 2\beta(1-2\rho_{ANAP})trace(\nabla_0) = 2.55\times 2.33 = 5.94.$$

This greatly exceeds the upper bound of $+2.0$ for convergence to the electoral origin. The major eigenvalue for the ANAP characteristic matrix is $+3.85$, with eigenvector $(+1.0, +0.98)$, while the minor eigenvalue is $+0.09$, with orthogonal, minor eigenvector $(-0.98, +1.0)$. The eigenvectors of this Hessian are almost perfectly aligned with the principal and minor components, or axes, of the electoral distribution.

Although the electoral origin satisfies the first order condition for local equilibrium, it follows from a standard result that the electoral origin is a *minimum* of the vote share function of ANAP, when the other parties are at the same position. On both principal and minor axes, the vote share of ANAP increases as it moves away from the electoral origin, but because the major eigenvalue is much larger than the minor one, we can expect that the AKP (as well the other parties) in equilibrium to adopt positions along a single eigenvector. Figures 2.3 and 2.4 present two LNE obtained from simulation of the pure spatial model. These are:

$$\mathbf{z}_1 = \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x : rel & 0.16 & -0.69 & 0.40 & -0.50 & 0.47 & 0.23 \\ y : nat & 0.17 & -0.77 & 0.41 & -0.57 & 0.45 & 0.26 \end{bmatrix}.$$

$$\mathbf{z}_2 = \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x : rel & 0.17 & 0.43 & -0.65 & -0.51 & 0.47 & 0.22 \\ y : nat & 0.18 & 0.43 & -0.72 & -0.56 & 0.45 & 0.25 \end{bmatrix}.$$

Note that all the positions in these two LNE lie close to the principal axis given by the eigenvector $(1.0, 1.0)$. The higher valence parties, the AKP and CHP lie closer to the origin, while the lower valence parties tend to be further from the origin.

In contrast, the estimated positions of the parties for 2002 in Figure 2.2 are:

$$\mathbf{z}^* = \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x : rel & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y : nat & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix}.$$

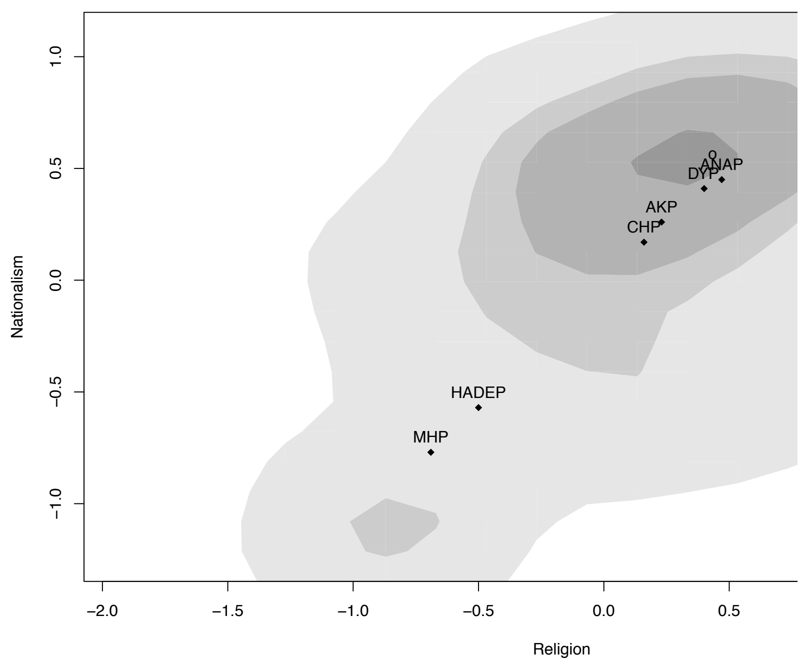

Figure 2.4. A Local Nash Equilibrium for the pure spatial model in 2002

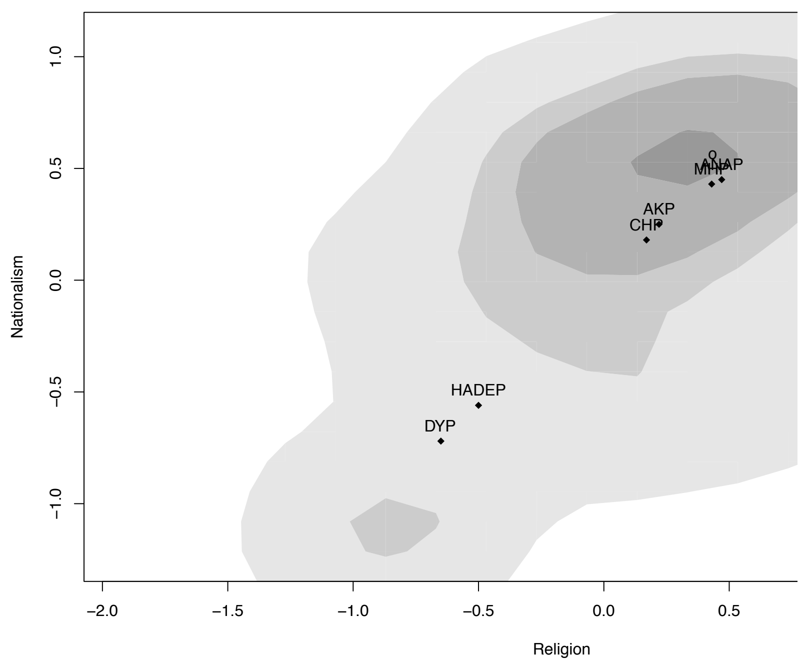

Figure 2.5. A Local Nash Equilibrium for the pure spatial model in 2002

The equilibrium positions of the CHP and MHP, particularly, are very far from their estimated positions.

Errors in the models

Appendix A shows that the standard error on Λ_{ANAP} is $h = 0.19$, so

$$\begin{aligned}\rho_{ANAP}(\Lambda_{ANAP} + h) &= \rho_{ANAP}(\Lambda_{ANAP}) + h\frac{d\rho_{Anap}}{d\Lambda} \\ &= \rho_{ANAP}(\Lambda_{_{ANAP}}) + h\rho_{_{ANAP}}(1 - \rho_{_{ANAP}}).\end{aligned}$$

This gives a standard error of 0.014 and a 95% confidence interval on ρ_{ANAP} of $[0.05, 0.11]$. Since the standard error on β is 0.12, giving a confidence interval on β of approximately $[1.28, 1.76]$, the standard error on c is 0.27. Using the lower bound on β and upper bound on ρ_{ANAP} gives an estimate for the 95% confidence interval on c of $[4.65, 7.38]$, so we can assert that, with very high probability, the convergence coefficient exceeds 4.0. Another way of interpreting this observation is that even if we use the upper estimate of the relative valence for ANAP, and the lower bound on β, then the joint origin will still be a minimum of the vote share function for ANAP.

We now repeat the analysis for the election of 1999.

2.2.2 The 1999 Election

The empirical model presented in Table 2.4 estimated the electoral variance on the first axis (religion) to be 1.20 while on the second axis (nationalism) the electoral variance, σ^2, was 1.14, giving a total electoral variance, σ^2, of 2.34, with the covariance between the two axes equal to +0.78.

The electoral covariance matrix is the 2 by 2 matrix

$$\nabla_0 = \begin{bmatrix} 1.20 & 0.78 \\ 0.78 & 1.14 \end{bmatrix}.$$

For the model ,the β coefficient was 0.375, while the party with the lowest valence was FP with $\Lambda_{FP} = -0.16$. The vector of valences is:

$$\begin{aligned} & (\Lambda_{FP}, \Lambda_{MHP}, \Lambda_{DYP}, \Lambda_{HADEP}, \Lambda_{ANAP}, \Lambda_{CHP}, \Lambda_{DSP}) \\ = \;& (-0.16, +0.66, 0.0, -0.071, +0.34, +0.73, +0.72). \end{aligned}$$

When all parties are located at the origin, the probability, ρ_{FP}, that a voter chooses FP under $\mathbb{M}(\boldsymbol{\Lambda},\beta)$ is equal to

$$\begin{aligned} & \frac{1}{[1+\exp(0.82)+\exp(0.16)+\exp(0.09)+\exp(0.5)+\exp(0.89)+\exp(0.88)]} \\ = \;& [11.27]^{-1} = 0.08. \end{aligned}$$

The standard error on Λ_{FP} is 0.175, so the 95% confidence interval can be estimated to be $[[0.01, 0.15]$. The FP vote share in 1999 was 15.41%, suggesting that the pure spatial model should be extended to include sociodemographic valences.

Now $2\beta(1-2\rho_{FP}) = 2\beta\times(1-2\times(0.08)) = 2\times 0.38\times 0.84 = 0.64$, so the characteristic matrix of the FP is

$$\begin{aligned} C_{FP} &= (0.64)\begin{bmatrix} 1.20 & 0.78 \\ 0.78 & 1.14 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.24 & 0.448 \\ 0.448 & -0.27 \end{bmatrix}. \\ \text{and } c &= 0.64 \times 2.34 = 1.49. \end{aligned}$$

Although $c < 2.0$, we can compute the eigenvalues of C_{FP} to be -0.74 with minor eigenvector $(+1, -1.116)$ and $+0.23$, with major eigenvector $(+1, +0.896)$, giving a saddle-point for the FP Hessian at the joint origin. As with the 2002 election, on the basis of the pure spatial model, we again expect all parties to align along the major eigenvector, at approximately 45 degrees to the religion axis. Note, however, that the standard error on c is of order 0.22, so unlike the result for the election of 2002, we cannot assert that there is a high probability that the convergence coefficient exceeds 2. However, there is a probability exceeding 0.95 that one of the eigenvalues is positive.

In comparing the pure spatial models of the elections of 1999 and 2002, we note there is very little difference between the model predictions.

2.2.3 Extension of the model for Turkey

We now use the empirical joint model, $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta)$, in order to better model party positioning. We use this model in order to estimate the influence of party activists in a more general activist model, denoted $\mathbb{M}(\mathbf{\Lambda}, \mu, \beta)$. In the activist model,the activist functions $\mu = \{\mu_j : j \in P\}$ are presumed to be functions of party position, rather than exogenous constants. The idea behind this model is that activists provide campaign contributions to specific parties, and these contributions can be used by the parties to affect valence. For the game theoretic foundations of this model see [Grossman and Helpman, 1996, Grossman and Helpman, 1994, Grossman and Helpman, 2002]. Grossman and Helpman, 1996 also define two distinct motives for these activists:

"Contributors with an *electoral* motive intend to promote the electoral prospects of preferred candidates. Those with an *influence motive* aim to influence the politicians' policy pronouncements."

Here we use a reduced form of the activist functions, based on [Schofield, 2006], since we only need the fact that the activist contribution to party j is a differentiable function of the party's position, and positively affects the parties valence.

Theorem 1 of Appendix A shows that the first order condition for a local equilibrium, $\mathbf{z}^* = (z_1^*, ...z_p^*)$, in the activist model is given by the set of *gradient balance conditions*:

$$\frac{d\mathcal{E}_j^*}{dz_j}(z_j^*) + \frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j^*) = 0. \tag{2.1}$$

Each term, $\frac{d\mu_j}{dz_j}(z_j)$ is the *the marginal activist pull (or gradient) at* z_j, giving the marginal activist effects on party j , while the gradient term $\frac{d\mathcal{E}_j^*}{dz_j}(z_j) = \left[z_j^{el} - z_j\right]$ is the *gradient electoral pull on the party*, at z_j, pointing towards its weighted electoral mean, z_j^{el}, as defined for party j in (5) in Appendix A:

$$z_j^{el} \equiv \sum_{i=1}^{n} \varpi_{ij} x_i, \text{where } [\varpi_{ij}] = \left[\frac{[\rho_{ij} - \rho_{ij}^2]}{\sum_{k \in N}[\rho_{kj} - \rho_{kj}^2]}\right]. \tag{2.2}$$

The weighted electoral mean essentially weights voter policy preferences by the degree to which the sociodemographic valences influence the choice of the voter.

Note in particular that (2) gives the first order condition for any of the various models considered here. In particular, if the sociodemographic and activist terms are zero, then (2) reduces to $[\alpha_{ij}] = \frac{1}{n}$, and, by the obvious coordinate transformation, we obtain $z_j = 0$, for all j, as the first order condition.

The joint model, $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta)$, allows us to draw some inferences about equilibrium positions. First we note that in the joint model, the sociodemographic valences are substitutes for the relative valences. Table 2.7 shows that the only relative valence that is significantly non zero in 2002 is $\mathbf{\Lambda}_{AKP}$. A number of the sociodemographic valences are, however, very significant. [28]

Figures 2.4 gives an LNE, $\mathbf{z}_3$, obtained by simulation of the joint model, $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta)$:

[28] The Bayes factors, or differences between the log marginal likelihoods of the joint models over the pure spatial models were +5 in both years.

$$\mathbf{z}_3 = \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x:rel & 0.12 & 0.26 & 0.40 & -0.50 & -0.58 & 0.19 \\ y:nat & 0.16 & 0.38 & 0.41 & -0.51 & -0.61 & 0.24 \end{bmatrix}.$$

Again the estimated positions are:

$$\mathbf{z}^* = \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x:rel & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y:nat & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix}.$$

Comparing the joint model with the pure spatial model, we see that the equilibrium positions are slightly better predictors for HADEP, MHP and ANAP

For this joint model, Tables 2.6 and 2.7 show that the sociodemographic valences for HADEP (or DEHAP) by Kurdish voters were very high:

$$\begin{aligned} (\theta_{HADEP} \cdot \eta_{Kurd}) &= 5.9 \text{ in } 1999 \\ (\theta_{HADEP} \cdot \eta_{Kurd}) &= 6.0 \text{ in } 2002. \end{aligned}$$

Keeping the other variables at their means in 2002, then changing η_{Kurd} from non-Kurd to Kurd increases the probability of voting for HADEP from 0.013 to 0.45. The high significance level of the sociodemographic variables indicates that the joint electoral model would predict that HADEP would move close to Kurdish voters who tend to be located on the left of the religion axis, and are also anti-nationalistic. The position marked HADEP in Figure 2.2 is consistent with this inference.

The joint model also shows that Alevi voters have very high sociodemographic valences for the CHP, with

$$
\begin{aligned}
(\theta_{CHP} \cdot \eta_{Alevi}) &= 3.1 \text{ in } 1999 \\
(\theta_{CHP} \cdot \eta_{Alevi}) &= 2.6 \text{ in } 2002.
\end{aligned}
$$

The Alevis are a non-Sunni religious community, who are adherents of Shia Islam rather than Sunni, and may be viewed as supporters of "Kemalism" or the secular state. Again, with other variables at their means, changing η_{Alevi} from non-Alevi to Alevi increases the probability of voting for CHP in 2002 from 0.16 to 0.63. Thus the joint model indicates that the CHP will move to a vote maximizing position, on the left of the religious axis, again as indicated in Figure 2.2.

Conversely, for Alevi voters $\theta_{AKP} \cdot \eta_{Alevi}) = -0.25$ in 2002, and we can infer that the AKP may have right to attract Sunni voters.

In the model $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta)$, we do not consider activist terms, so this is equivalent to setting $\{\frac{d\mu_j}{dz_j}\} = 0$. We can infer from (1) that the first order balance condition will be satisfied at a vector $\mathbf{z} = (z_1, ...z_p)$ when

$$\frac{d\mathcal{E}_j^*}{dz_j} \equiv \left[z_j^{el} - z_j\right] = 0, \text{for each } j.$$

Thus we can use $\mathbf{z}_1$ as the estimator for the vector of weighted electoral means. We find that

$$\begin{aligned}
\mathbf{z}^* - \mathbf{z}_3 &= \\
&\begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x:rel & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y:nat & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix} \\
&- \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x:rel & 0.12 & 0.26 & 0.40 & -0.50 & -0.58 & 0.19 \\ y:nat & 0.16 & 0.38 & 0.41 & -0.51 & -0.61 & 0.24 \end{bmatrix}. \\
&= \begin{bmatrix} Party & CHP & MHP & DYP & HADEP & ANAP & AKP \\ x:rel & -3.2 & -0.26 & -0.40 & -1.50 & +0.38 & 0.81 \\ y:nat & -0.15 & +1.12 & 0.09 & -0.99 & +0.51 & -0.14 \end{bmatrix}
\end{aligned}$$

Assuming that this vector is an LNE with respect to the full model, $\mathbb{M}(\mathbf{\Lambda}, \mu,\beta)$ involving activists, then by (13) in the Appendix A, we can make the identification:

$$\frac{1}{2\beta}\left[\frac{d\mu_1}{dz_1},...,\frac{d\mu_p}{dz_p}\right] = \mathbf{z}^* - \mathbf{z}_3$$

Here, $\{\frac{d\mu_1}{dz_1},...,\frac{d\mu_p}{dz_p}\}$ are the marginal activist pulls at the equilibrium vector $\mathbf{z}^*$. Under the hypothesis that the joint model with activists is valid, then the difference between these two vectors gives us an estimate of the vector of marginal pulls on the parties:

The estimated activist pull on HADEP is very high, pulling the party to the left on the religion axis, and in an anti-nationalist direction on the y axis. Similarly,the estimated activist pull on the CHP is even higher on the religious axis, pulling the party in a secular direction, and we can infer that this is due to the influence of Alevi voters.

As a consequence, this asymmetry will cause Alevi activists to provide further differential support for the CHP. It is thus plausible that secular voters (on the left of the religious axis in Figures 2.1 and 2.2) would offer further support to the CHP, located close to them. This would affect the party's marginal activist pull, and induce the CHP leader to move even further left, towards its inferred equilibrium position in the full activist model.

We suggest that activist support for the AKP would move it slightly to the right on the religion axis, as well as in an anti-nationalism direction. This would result in its estimated position as in Figure 2.2.

In contrast, we might conjecture that the military provides activist support for the MHP on the nationalism axis, and this will move the party to the left in a secular direction, and north on the nationalism axis, resulting in its position in Figure 2.2.

Overall, we note that we can expect activist valence to strongly influence party positioning, and we can proxy this support to some degree using the sociodemographic variables. Notice that the sociodemographic variables are estimated at the vector $\mathbf{z}^*$, so the estimated sociodemographic valences have been influenced by activist support. The LNE obtained from the joint model is a hypothetical solution to the vote maximizing game involving the parties, based on some empirical assumptions about the underlying nature of the important sociodemographic groups in the polity.

2.3 Concluding Remarks

Recent works by Acemoglu and Robinson, 2005, Boix, 2003, and Przeworski, 2000 have explored the transition from autocratic regimes to democracy. A recent contribution by Epstein et al., 2006 has emphasized the existence of the category of "partial democracies." These exhibit mixed characteristics of both democratic and autocratic regimes. In fact, Epstein et al., 2006 give Turkey as a prime illustration of the possible degree of democratic volatility of a regime. They observe that, in terms of Polity IV scores, Turkey fell from being a full democracy to an autocracy first in the mid 1960's and again in the early 1980's, and since then has hovered between partial and full democracy. [Epstein et al., 2006] also comment, on the basis of their empirical analysis, that "the determinants of the behavior of partial democracies elude our understanding." These models of democratic transitions have tended to consider a single economic axis, and to utilize the notion of a median citizen, or median kingmaker as the unique pivotal player.[29] While these models have been illuminating, we believe it necessary to consider policy spaces of higher dimension and to utilize a stochastic model so as to emphasize the aspect of uncertainty.

The analysis of Turkey in this paper indicates that both religion and nationalism define the political space.[30] The military in Turkey can be represented by a pro-nationalist position, which is also far from the religiously conservative masses and the governing party, AKP. It is this very phenomenon which means that Turkish politics cannot be understood in terms of a median voter. Modeling partial democracies would seem to require a very explicit analysis of the power of activist groups.

This paper has applied a theoretical stochastic model to present an empirical analysis of elections in Turkey, and argues that there is no evidence of a centripetal tendency towards an electoral center. Instead it suggests that activist groups will tend to be located far from the electoral center. Once the sociodemographic valences have caused the parties to move away from the center in order to gain electoral support, the influence of activists will separate the parties even further, pulling them towards

[29] See Gallego and Pitchik (2004).

[30] [Schofield and Sened, 2006] found the electoral model for Israel to be very similar to Turkey, with two electoral axes, religion and security. [Schofield and Zakharov, 2010] found nationalism to be one of the principal axes in Russia, but the second axis was defined by attitudes to capitalism/ communism, perhaps comparable to religion.

policy positions preferred by the activists. Thus simulation of the joint model with sociodemographic valence can be used to infer aspects of this activist influence.

2.4 Appendix A: Formal and Empirical Electoral Models

2.4.1 The Model with Activists

The electoral model presented here is an extension of the multiparty stochastic model of McKelvey and Patty, 2006, modified by inducing asymmetries in terms of valence. The justification for developing the model in this way is the empirical evidence that valence is a natural way to model the judgements made by voters of party leaders and candidates. There are a number of possible choices for the appropriate model for multiparty competition. The simplest one, which we first present, is that the utility function for leader j is proportional to the popular support, V_j, of the party in the election.[31] With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to obtain conditions for positions to be locally optimal. Thus we examine what we call *local pure strategy Nash equilibria* (LNE). From the definitions of these equilibria it follows that a PNE must be a LNE, but not conversely. A necessary condition for an LNE is thus a necessary condition for a PNE. A sufficient condition for an LNE is not a sufficient condition for PNE. Indeed, additional conditions of concavity or quasi-concavity are required to guarantee existence of PNE.

[31]The popular support may be identical to the vote share in a democratic election, or may be weighted by individual characteristics, such as domicile, income or ownership of land, in non-democratic polities.

The stochastic model essentially assumes that candidates cannot predict vote response precisely, but that they can estimate the effect of policy proposals on the expected vote share. In the model with valence, the stochastic element is associated with the weight given by each voter, i, to the average perceived quality or valence of each candidate. We also consider a formal model where the perceptions of the leader qualities vary across different sociodemographic groups in the society.

The data of the spatial model is a distribution, $\{x_i \in X\}_{i \in N}$, of voter ideal points for the members of the electorate, N, of size n. We assume that X is a subset of Euclidean space, of dimension w with w finite. Without loss of generality, we adopt coordinate axes so that $\frac{1}{n}\Sigma x_i = 0$. By assumption $0 \in X$, and this point is termed the *electoral mean*, or alternatively, the *electoral origin*. Each of the parties in the set $P = \{1, \ldots, j, \ldots, p\}$ chooses a policy, $z_j \in X$, to declare prior to the specific election to be modeled. Let $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ be a typical vector of party policy positions.

Given $\mathbf{z}$, each citizen, i, is described by a utility vector

$$\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \ldots, u_{ip}(x_i, z_p))$$

where

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) - \beta||x_i - z_j||^2 + \varepsilon_{j.} = u^*_{ij}(x_i, z_j) + \varepsilon_j. \tag{2.3}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable component of utility. The constant term, λ_j, is the fixed or *exogenous valence* of party j, The function $\mu_j(z_j)$ is the component of valence generated by activist contributions to agent j. We can also refer to this term as *endogenous valence.* The term β is a positive constant, called the *spatial parameter*, giving the importance of policy difference defined in terms of a metric induced from the Euclidean norm, $||\cdot||$, on X. The vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_p)$ is the stochastic error, whose multivariate cumulative distribution will be denoted by Ψ. The notation $\lambda_j + \mu_j(z_j)$ is intended to imply that this is the average valence for party j among the electorate, but the realized valence is a distributed by Ψ. The most common assumption in empirical analyses is that Ψ is the *Type I extreme value distribution* (sometimes called Gumbel). This cumulative distribution has the closed form

$$\Psi(x) = \exp\left[-\exp\left[-x\right]\right].$$

The theorems presented in this appendix are based on this assumption. This distribution assumption is the basis for much empirical work based on multinomial logit estimation [Dow and Endersby, 2004].

In empirical models, the exogenous valences are simply real numbers, estimated by the model. Since they are all finite, they can be ranked. We therefore assume that the exogenous valence vector is given by

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p) \text{ satisfies } \lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1.$$

This is a strong assumption, in that it assumes that every voter ranks the parties in this fashion. Adding sociodemographic valences, as in the body of the paper, means that this ranking over valences differs among the electorate.

Voter behavior is modeled by a probability vector. The probability that a voter i chooses party j at the vector $\mathbf{z}$ is

$$\rho_{ij}(\mathbf{z}) = \Pr[[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)], \text{ for all } l \neq j].$$

Here Pr stands for the probability operator generated by the distribution assumption on ε.

With this distribution assumption on Ψ, it follows, for each voter i, and leader j, that

$$\rho_{ij}(\mathbf{z}) = \frac{\exp[u^*_{ij}(x_i, z_j)]}{\sum\limits_{k=1}^{p} \exp u^*_{ik}(x_i, z_k)}. \tag{2.4}$$

For any voting model the *likelihood* of a model is

$$\mathbb{L} = \prod_{i \in N,\ j_i \in P} \rho_{ij_i}(\mathbf{z}),$$

where j_i is the party that i chooses. The log likelihood of the model is $\log_e(\mathbb{L})$. Clearly as $\mathbb{L}$ approaches 0 then $\log_e(\mathbb{L})$ approaches $-\infty$.

To compare two models, $\mathbb{M}_1$ and $\mathbb{M}_2$, the *Bayes Factor* is $\mathbb{L}(\mathbb{M}_1)/\mathbb{L}(\mathbb{M}_2)$ and the *log Bayes factor* of $\mathbb{M}_1$against $\mathbb{M}_2$ is $\log_e(\mathbb{L}(\mathbb{M}_1) - \log_e(\mathbb{L}(\mathbb{M}_2)$. A log Bayes factor over 5.0 for $\mathbb{M}_1$ against $\mathbb{M}_2$ is considered strong support for $\mathbb{M}_1$ (Kass and Raftery, 1995).

The *expected popular support* for leader j is

$$V_j(\mathbf{z}) \equiv \sum_{i \in N} s_{ij} \rho_{ij}(\mathbf{z}).$$

Here $\{s_{ij}\}$ are different weights that can be associated with different voters. In the case all weights are equal to $\frac{1}{n}$, we call the model *egalitarian.*

It is useful to have a formal model where voter weights differ. For example, in US Presidential elections, it is not the vote share *per se* but the *share of the electoral college total.* Voter weights in different States will therefor vary.

To present the model we now regard $\mathbf{V} = \{V_j : j \in P\}$ as a set of *vote share functions*, and identify $\mathbf{V}$ as a differentiable *profile function*, $\mathbf{V} : X^p \to \mathbb{R}^p$. We denote the egalitarian profile function as $\mathbf{V}_e$.

In this stochastic electoral model, it is assumed that each party j chooses z_j to maximize V_j, conditional on $\mathbf{z}_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_p)$.

Thus a vector $\mathbf{z}^*{=}(z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*)$ is called a *local strict Nash equilibrium* (LSNE) if each z_j strictly locally maximizes V_j, conditional on $\mathbf{z}_{-j}$, while $\mathbf{z}^*$ is a *local weak Nash equilibrium* (LNE) if each z_j weakly locally maximizes V_j, conditional on $\mathbf{z}_{-j}$. The notion of LSNE is convenient so as to avoid degeneracy problems associated with the Hessians.

In the same way the vector $\mathbf{z}^*$ is a *strict (or weak) pure strategy Nash equilibrium* (PSNE or PNE) if each party j chooses z_j to strictly (or weakly) maximize V_j on X.

Now assume that the vector $\mathbf{z}$ is fixed, and let $\rho_{ij}(\mathbf{z}) = \rho_{ij}$ be the probability that i picks j. Define the p by n matrix array of weights by

$$[\varpi_{ij}] \equiv \left[\frac{s_{ij}[\rho_{ij} - \rho_{ij}^2]}{\sum_{k \in N} s_{kj}[\rho_{kj} - \rho_{kj}^2]} \right] \tag{2.5}$$

The vector $\sum\limits_i \varpi_{ij} x_i$ is a convex combination of the set of voter ideal points and is called the *weighted electoral mean* for party j. Define

$$z_j^{el} \equiv \sum_{i=1}^{n} \varpi_{ij} x_i \text{ and } \frac{d\mathcal{E}_j^*}{dz_j}(z_j) \equiv \left[z_j^{el} - z_j \right].$$

Then the *balance equation* for z_j^* is given by the expression

$$\frac{d\mathcal{E}_j^*}{dz_j}(z_j^*) + \frac{1}{2\beta} \frac{d\mu_j}{dz_j}(z_j^*) = 0. \tag{2.6}$$

The term $\frac{d\mathcal{E}_j^*}{dz_j}(z_j)$ is the *marginal electoral pull of party* j at the point z_j and can be regarded as a gradient vector, at z_j, pointing towards the weighted electoral mean of the party. (Note that this electoral pull depends on the positions of all leaders.) When z_j is equal to the weighted electoral mean then the electoral pull is zero. The gradient vector $\frac{d\mu_j}{dz_j}(z_j)$ is called *the marginal activist pull for party* j at z_j.

When $\frac{d\mu_j}{dz_j}(z_j) = 0$, then then the balance equation reduces to setting $z_j = z_j^{el}$.

If $\mathbf{z}^* = (z_1^*, \ldots z_j^*, \ldots z_p^*)$ is such that each z_j^* satisfies the balance equation then call $\mathbf{z^*}$ a *balance solution.* The balance solution requires that the electoral and activist gradients are directly opposed, for every party leader.

The model just presented is denoted $\mathbb{M}(\lambda, \mu,\beta; \mathbf{V})$. [Schofield, 2006] proves the following theorem for this model.

Theorem 1.

Consider the electoral model $\mathbb{M}(\lambda, \mu,\beta; \mathbf{V})$ based on the distribution, Ψ, including both exogenous and activist valences, and defined by the family $\mathbf{V}$ of vote share functions.

(i) The first order condition for $\mathbf{z}^*$ to be an LSNE is that it is a balance solution.

(ii) If all activist valence functions are sufficiently concave[32], then a balance solution will be a PNE.

In the full activist model, $\mathbb{M}(\lambda, \mu,\beta; \mathbf{V})$, with valence functions $\{\mu_j\}$ that are not identically zero or constant, then it is the case that generically $\mathbf{z}_0$ cannot satisfy the first order conditions for LNE even when $\mathbf{V}$ is egalitarian. Instead the vector $\frac{d\mu_j}{dz_j}$ "points towards" the position at which the activist valence for leader j is maximized. When this marginal or gradient vector, $\frac{d\mu_j}{dz_j}$, is increased (as activist groups become more willing to contribute to leader j) then the equilibrium position is pulled away from the weighted electoral mean of the leader, and we can say the "activist effect"

[32] By this we mean that the eigenvalues of the activist functions are negative and of sufficient magnitude everywhere. That is to say, there exists $\alpha < 0$, such that all eigenvalues $< \alpha$ is sufficient to guarantee existence of a PNE.

for the leader is increased. In the case of two opposed leaders, j and k, if the activist valence functions are fixed, but the exogenous valence, λ_j, is increased, or λ_k, is decreased, then the weighted electoral mean, z_j^{el}, approaches the electoral origin. Thus the local equilibrium of leader j is pulled towards the electoral origin. We can say the "electoral effect" is increased.

2.4.2 The Egalitarian Model without Activists

In the case that the activist valence functions are identically zero, or constant, we denote the model by $\mathbb{M}(\lambda,\beta;\mathbf{V})$.The key consideration for the egalitarian model, $\mathbb{M}(\lambda,\beta;\mathbf{V}_e)$,when all voter weights are identical, is whether the electoral origin is a LSNE. For this model it can be shown that if all parties are at the same position, so $\mathbf{z}^*= (z^*, z^*, ...z^*)$ then every $\{\rho_{ij}(\mathbf{z}^*) : i \in N\}$ is independent of i, and can thus be written $\rho_j(\mathbf{z}^*)$. This implies that all α_{ij} in (5) are identical at $\mathbf{z}^*$ and equal to $\frac{1}{n}$. Thus, when there is only exogenous valence, the equation $z_j^* = \frac{1}{n}\Sigma x_i$ satisfies the balance solution for all j. By an appropriate coordinate change, we can assume $\frac{1}{n}\Sigma x_i = 0$. In this case, all marginal electoral pulls are zero at $\mathbf{z}_0 = (0, ...0)$, so $\mathbf{z}_0$ satisfies the first order conditions. However, to determine whether $\mathbf{z}_0$ is an LNE it is necessary to examine the Hessians of the vote share functions.

We first define the *electoral covariance matrix,* ∇_0, and then use ∇_0 to define the convergence coefficient of the model $\mathbb{M}(\lambda,\beta;\mathbf{V}_e)$. Let $X = \mathbb{R}^w$ be endowed with a system of coordinate axes $r = 1, \ldots, w$. For each coordinate axis let $\xi_r = (x_{1r}, x_{2r}, \ldots, x_{nr})$ be the vector of the r^{th} coordinates of the set of n voter ideal points. The scalar product of ξ_r and ξ_s is denoted $(\xi_r \centerdot \xi_s)$. Let $(\sigma_r \centerdot \sigma_s) = \frac{1}{n}(\xi_r \centerdot \xi_s)$ be the electoral covariance between the r^{th} and s^{th} axes, and σ_s^2 be the variance on the s^{th} axis.

(i) The symmetric $w \times w$ *electoral covariance matrix* about the origin is denoted ∇_0 and is defined by

$$\nabla_0 \equiv [(\sigma_r \centerdot \sigma_s)]_{s=1,\ldots,w}^{r=1,\ldots,w}.$$

(ii) The total electoral variance is $\sigma^2 \equiv \sum_{s=1}^{w} \sigma_s^2 = trace(\nabla_0)$.

(iii) At the vector $\mathbf{z}_0 = (0, \ldots, 0)$ the probability $\rho_{ij}(\mathbf{z}_0)$ that i votes for party j is independent of i, and is given by

$$\rho_j = \left[1 + \sum_{k \neq j} \exp\left[\lambda_k - \lambda_j\right]\right]^{-1}. \tag{2.7}$$

(iv) The Hessian of the egalitarian vote share function of party j at $\mathbf{z}_0$ is a positive multiple of the w by w *characteristic matrix*,

$$C_j \equiv 2\beta(1 - 2\rho_j)\nabla_0 - I. \tag{2.8}$$

(Here I is the identity matrix.)

The *convergence coefficient* of the egalitarian model, $\mathbb{M}(\lambda,\beta;\mathbf{V}_e)$, is defined to be

$$c \equiv c(\lambda,\beta;\mathbf{V}_e) \equiv 2\beta[1 - 2\rho_1]\sigma^2. \tag{2.9}$$

Theorem 2.

Consider the electoral model $\mathbb{M}(\lambda,\beta;\mathbf{V}_e)$ where all activist valence functions are zero (or constant) and $\mathbf{V}_e$ is the egalitarian party profile.

(i) The joint origin $\mathbf{z}_0 = (0,\ldots,0)$ satisfies the first order condition to be a LSNE for this model.

(ii) In the case that X is w dimensional then the necessary condition for $\mathbf{z}_0$ to be a LNE for this model is that $c(\lambda,\beta;\mathbf{V}_e) \leq w$.

(iii) In the case that X is 2 dimensional, a sufficient condition for $\mathbf{z}_0$ to be a LSNE for this model is that $c(\lambda,\beta;\mathbf{V}_e) < 1$.

The proof and some applications of Theorem 2 are given in [Schofield, 2007].

2.4.3 Empirical Models

In empirical models with exogenous valence alone it is necessary to estimate the model with respect to the valence of a baseline party, say k^*. We set $\Lambda_j = \lambda_j - \lambda_{k^*}$, and call these the relative valences. We denote this egalitarian model by $\mathbb{M}(\mathbf{\Lambda},\beta;\mathbf{V}_e)$.

At the joint origin,$\mathbf{z}_0$, we see that

$$\rho_{ij}(\mathbf{z}_0) = \frac{\exp(\lambda_j)}{\sum_{k=1}^{p}\exp(\lambda_k)} = \frac{\exp(\lambda_j - \lambda_{k^*})}{\sum_{k=1}^{p}\exp(\lambda_k - \lambda_{k^*})} = \frac{\exp(\Lambda_j)}{\sum_{k=1}^{p}\exp(\Lambda_k)} \tag{2.10}$$

is again independent of the individual, i, and can be written as ρ_j.

To estimate the standard error on ρ_j, we use Taylor's Theorem, which asserts that

$$\rho_j(\Lambda_j + h) = \rho_j(\Lambda_j) + h\frac{d\rho_j}{d\Lambda_j} = \rho_j(\Lambda_j) + h\rho_j(1-\rho_j). \tag{2.11}$$

2.4.4 Empirical Models with Sociodemographic Valences

As described in the body of the paper, in empirical applications with sociodemographic variables , we typically assume that $\mathbf{V}$ is the egalitarian party profile function,$\mathbf{V}_e$, so the model $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta; \mathbf{V}_e)$ is based on the assumption that voter utility has the form

$$u_{ij}(x_i, z_j) = \Lambda_j + (\theta_j \centerdot \eta_i) - \beta\|x_i - z_j\|^2 + \varepsilon_{j.}$$

The estimate of voter i's valence will then be $\Lambda_j + (\theta_j \centerdot \eta_i)$, so this will vary from one voter to another. A consequence of this is that, in the expression 2.5 for the weighted electoral mean, even when all parties are at the origin, then the denominator term $\{\rho_{kj}(\mathbf{z}_0) : k \in N\}$ will depend on voter k. This implies that voters will be weighted differently, and generically, $\mathbf{z}_0$ will not satisfy the first order condition for LNE. However, the joint empirical model, $\mathbb{M}(\mathbf{\Lambda}, \theta, \beta; \mathbf{V}_e)$, assumes that the sociodemographic effects are independent of party positions, and this implies$\frac{d\mu_j}{dz_j} = 0$,for all j. Using (6), we infer that the various LNE obtained by simulation of the joint model provides an estimate of a set of vectors of weighted electoral means: $\{\mathbf{z}^{el} = (z_1^{el}, ..., z_p^{el})\}$.

Assuming that the estimated party positions are given by the vector $\mathbf{z}^* = (z_1^*, .., z_p^*)$ and that this is in equilibrium with respect to the full activist model, then choosing one joint LNE, $\mathbf{z}^{el}$, gives an estimate of

$$\begin{aligned} \left[z_j^{el} - z_j^*\right] &\equiv \frac{d\mathcal{E}_j^*}{dz_j}(z_j^*) = -\frac{1}{2\beta}\left[\frac{d\mu_j}{dz_j}\right]. \qquad (2.12) \\ \text{Thus } \left[\mathbf{z}^* - \mathbf{z}^{el}\right] &= \frac{1}{2\beta}\left[\frac{d\mu_1}{dz_1}, ..., \frac{d\mu_p}{dz_p}\right]. \qquad (2.13) \end{aligned}$$

This observation suggests how the gradients of the activist valence functions may be inferred from a comparison of LNE of the joint empirical model with the estimated political configuration.

2.5 Appendix B: Estimation and Simulation Details

A Bayesian procedure is employed for the estimation of the model. In the Bayesian approach, rather than treating model parameters as fixed, we treat them as unknown and try to find a distribution for the model parameters using Bayes rule:

$$\pi(\psi|y) = \frac{f(y|\psi)\pi(\psi)}{f(y)} \tag{2.14}$$

The left side is the distribution of ψ after observing the data, so called the posterior distribution. $f(y|\psi)$ is the likelihood function, the density function for the observed data y given the parameter values ψ. $\pi(\psi)$ represents our belief about the distribution of ψ before observing the data and called the prior distribution. The term in the denominator normalizes the posterior and does not depend on ψ, $f(y) = \int f(y|\psi)\pi(\psi)$. That is why the posterior distribution is often stated in terms of proportionality:

$$\pi(\psi|y) \propto f(y|\psi)\pi(\psi) \tag{2.15}$$

However a posterior function obtained in this form is a high-dimensional object which is hardly informative about the parameters. For any practical purpose, we would want to know the marginal distributions of the parameters and possibly report the moments such as the posterior mean and posterior standard deviation. The rich set of MCMC simulation techniques are well suited for this task.

The first step is to choose a prior distribution for the model parameters, ψ. For our case we assess a multivariate Normal prior:

$$\psi \sim N(\psi_0, \Psi_0)$$

The likelihood function is then:

$$\begin{aligned} f(y|\psi) &= \prod_{i=1}^{n} \rho_{i,j} \\ &= \prod_{i=1}^{n} \frac{\exp[u_{ij}(x_i, z_j)]}{\sum_{l=1}^{p} \exp[u_{ij}(x_i, z_l)]} \end{aligned} \tag{2.16}$$

It has been stated that the asymptotic normal approximation is excellent for the multinomial logit likelihood in this form [Rossi et al., 2005]. Hence the posterior distribution becomes

$$\pi(\psi|y) \propto |H|^{\frac{1}{2}} \exp\{\frac{1}{2}(\psi - \hat{\psi})' H (\psi - \hat{\psi})\} \tag{2.17}$$

$\hat{\psi}$ is chosen to be the maximum likelihood estimator for $\hat{\psi}$ and H is chosen to be minus the Hessian of the likelihood evaluated at $\hat{\psi}$.

MCMC algorithms approximates the posterior distribution of the parameters by simulating from it. The Metropolis-Hastings (MH) algorithm is a general principle for this purpose. Independence MH is a special case of the MH algorithm. Without going into the details about why and how the MH algorithm work, we will summarize the independence MH algorithm here. Suppose we want to sample from the distribution

$f(X)$[33], where X can be scalar or vector random variable. Then the independence MH algorithm works as follows:

1. Choose a starting value for the parameters to be estimated, x.

2. Generate Y from q(Y).

3. Generate U from $U(0,1)$ — the uniform distribution. If

$$U \leq \min\left\{\frac{f(Y)q(Y)}{f(x)g(x)}, 1\right\},$$

   return Y. Otherwise, return x and go to 2.

The distribution $q(.)$ is called the proposal density. The choice of the proposal density is "often a matter of judgement" [Greenberg, 2008]. However one thing to be careful about choosing a proposal density is that the "acceptance probability" should be reasonable. Again following Greenberg, 2008, we have chosen a fat-tailed candidate distribution, the multivariate student. The candidate vectors were drawn using $\psi \propto$ MSt$(\nu, \hat{\psi}, H^{-1})$ where $\hat{\psi}$ is the maximum likelihood estimator for ψ, H^{-1} is the sample information matrix (minus the actual Hessian of the likelihood evaluated at $\hat{\psi}$) and ν is the tuning parameter which is chosen to be 6 for our case.[34]

The simulation described here is implemented via the `rmnlIndepMetrop` function in the `bayesm` package[35] written for the statistical software `R`.

[33]So in our case this is the posterior distribution given by 2.17
[34]This is chosen in the range (5-15) suggested by Rossi et al., 2005
[35]http://faculty.chicagogsb.edu/peter.rossi/research/bsm.html

3. A Stochastic Model of Clientelistic Linkage as Valence Buying

3.1 Introduction

There is a revival in the comparative politics literature in an examination of "machine politics." Political parties can attempt to affect electoral outcomes by targeting benefits to important constituencies and monitoring elections, either in the vote choice or turnout. This reveals a rich dynamic between policy choice and vote choice. Recently, scholars have constructed formal models of this type of clienteslistic linkage, clearly laying out the assumptions of the underlying mechanisms and generating comparative statics that greatly enhance our understanding of machine politics.

In this paper I do two things. First I build upon this recent wave of formal models, specifically building upon the models of [Stokes, 2005] and [Nichter, 2008]. The heart of my model relaxes an information assumption used in both models, and replaces it with what I argue is a more realistic assumption of monitoring by the "machine." Loosening this assumption not only changes the individual comparative statics of these models, it allows for a deeper integration of both models, showing the conditions under which the ability to monitory and the level of awards available to the machine can affect vote buying or turnout buying.

The key insight of this first model is that both [Stokes, 2005] and [Nichter, 2008] make the assumption the machine has a perfect ability to observe votes cast in favor of the machine, yet uncertainty about the votes cast against the machine. While this may seem like a minor assumption, I argue that not only is it unrealistic, it has major implications for the models. By loosening this assumption I document changes in the comparative statics of the results, but the implications are broader than a revision of these two papers.

My main finding shows that, contra Stokes, there is a much wider change of voters that can be bought by the machine. While Stokes argues that the machine will target weakly opposed voters, my model shows that under some conditions, a much wider range of voters can be influenced to vote for the machine. Yet this information assumption does not dramatically change Nichter's results on vote turnout.

A second important implication is that the ability to directly monitor either vote choice or turnout has different effects on machines given this loosened assumption. Counter-intuitively, an increase in the ability of a machine to monitor an individual actions may not increase vote buying in [Stokes, 2005] model, but it does increase vote buying in [Nichter, 2008] turnout model.

Our third, and perhaps most striking result, is that an increase in the rewards (resources) of the machine may not effect the probability of the machine to buy votes in either model. A machine with an imperfect ability to distinguish between voters who will cast their votes for the machine party, and those who will vote against, may be insensitive to the amount of potential rewards they can offer voters.

This last results could be especially important for comparative politics scholars exploring the relationship between natural resource rents and incumbency advantage. Countries with resource rich political machines may have an inability to use these resources to increase votes for the machine.

The second part of the paper introduces an alternative way of modeling the clientelistic linkage based on the models of spatial voting. Building upon a similar setup originally developed in [Schofield, 2004], I model the materialistic benefits distributed by the political parties as serving the purpose of increasing the valence of the party in the eyes of the receiver of the benefit. The policy choice of the political parties and their decision regarding to whom to bribe and how much are simultaneously determined in the equilibrium subject to their resource constraints. This approach opens the way to many interesting comparative statics which have implications outside the vote-buying literature, such as the controversial subject of bribing in international organizations. [Kuziemko and Werker, 2006]

3.2 The Literature

Since the 1950s, political science literature has been dominated by the "responsible party government" model. This is true both for rational choice theories starting with Downs, 1957 and for comparative approaches going back to Rokkan and Lipset, 1967. Although this model captures many dimensions in which parties' programs reflect the preferences of the voters, it ignores a quite a different type of party-voter linkage which exists even in the advanced industrial democracies: the clientelistic linkage.

In almost all such countries, this voter-party linkage is established through direct material benefits transferred to voters who trade their votes for a price. There is a wide variety among the clientelistic linkages in terms of the goods delivered by the machine parties. In many countries these are one shot consumable goods which can include money, clothes, food, liquor or coal [Callahan and McCargo, 1996]. In this sense "clientelistic accountability represents a transaction, the direct exchange of a citizens vote in return for direct payments or continuing access to employment, goods, and services." [Kitschelt and Wilkinson, 2007a]

In one of the earlier comparative studies Scott, 1972 provides a framework to examine the embeddedness of the clientelistic linkage in different types of regimes. The existence of policy-based competition shifts clientelistic linkages from local networks to a national level of hierarchical political machines [Scott, 1969]. So in a more democratic environment clientelism turns into a broker-mediated relationship rather than a face-to-face exchange [Weingrod, 1968]. The different faces of clientelism in different regime types and electoral systems is still an under-studied phenomenon.

Since the exchange between the patron and the client is not simultaneous, one important challenge in the clientelism literature is to account for the opportunistic defections of the voter. Why would not the voter just take the material benefit and votes for who ever he wants? One answer to this question is the parties monitoring the voter behavior. This is often difficult but without it, politicians run the risk of misdirecting the resources to voters those who will take the money and run. The simplest way to monitor individual voters is by violating the secrecy of the ballot

[Whyte, 1965] or giving voters the impression that this is the case [Chandra, 2007]. The monitoring can take the form of self-enforcing group equilibrium through ongoing network of social relations [Auyero, 1999]. This is an iterative process where all past behaviors generate an obligation for both the patrons and clients. Whatever the monitoring mechanism is, neither patrons nor clients are willing to describe the clientelistic relationship as a simple exchange of desirable goods, but instead interpret it in more innocent terms as an enactment of community relations and civic solidarity [Kitschelt and Wilkinson, 2007a].

Another important question in the clientelism literature is related to the types of voters targeted. One approach is based on the diminishing marginal utility of income and argues that the benefits will generate more voters with the same amount of budget if poor voters are targeted [Calvo and Murillo, 2004]. According to Cox and McCubbins, 1986, on the other hand, the crucial feature of the clientelistic linkage is that the party is more confident about how the core supporters will respond to rewards hence will target them. But if these voters are already ideologically attached to the party, why not target those who are not? Following this question Stokes, 2005 argues that the machine parties will target the "weakly opposed" voters. That is they will use their resources to buy the votes which are the cheapest in the sense that these voters need the smallest amount of reward to change sides.

The greatest obstacle in front of the studies on clientelism is the lack of comparative and historical data. There are many reasons for both voters and politicians to hide information in this context: to avoid prosecution, to avoid shame as well as to deter competitors [Kitschelt and Wilkinson, 2007b]. This problem is particularly severe for studies trying to provide a comparative perspective of linkage mechanisms in order to provide an explanation for the variation in different institutional contexts. Even studies which conduct small scale case studies have mentioned the difficulty of proving the existence of clientelism [Auyero, 1999]. Moreover the differences existing in the nature of the targeted goods and in the survey responses deepens the problems with the cross-country studies of clientelism.

3.3 The Iterative Model of Vote Buying

In a very influential paper, Stokes presents a formal model of machine politics based on a repeated game dynamic [Stokes, 2005]. The model is one of imperfect monitoring. The machine has a limited capability of observing the votes and rewards the voters depending on whether it thinks the voter voted for it or not. There are two parts of the voter's utility, the disutility caused by the ideological distance in a one dimensional policy space between himself and whom he voted for, and the materialistic benefit he receives from the machine if he gets any. The machine has a single opponent. I will mainly follow the notation introduced in that paper.

Let the ideological position of the voter in a one dimensional policy space be represented by x_i, the ideological position of the machine be represented by x_1, the ideological position of the opposition party by x_2, the midpoint between the parties by $x* = \frac{x_1+x_2}{2}$ and without loss of generality, let $x_1 < x_2$. Then the utility of a voter i is given by

$$u_i = -\frac{1}{2}(v_i - x_i) + b_i$$

where $v_i = \{x_1, x_2\}$ represents a vote for either the machine or the opposition and $b_i = \{0, b\}$ represents the value to the voter of the reward offered by the machine. In addition, let d_1 and d_2 be the measures of the distances between voter i and the parties, namely, let $d_1 = (x_i - x_1)^2/2$ and $d_2 = (x_i - x_2)^2/2$. One difference between my notation and Stokes's notation is that Stokes uses β for the discount factor and where as I use δ.

Next I define three different information environments applicable to our setting. Let c represent the event in which the machine receives a *Comply* signal —a vote for itself—and d represent the event in which the machine receives a *Defect* signal —a voter for the opposition party—. Let C and D represent the associated actual actions of the voter. The first information environment is the perfect information environment, namely the one in which the machine perfectly observes a compliance and a defection. In terms of conditional probability, we can write this as: $p(c|C) = p(d|D) = 1$; $p(d|C) = p(c|D) = 0$. The second environment is a semi-perfect (or semi-imperfect in that matter) information environment, namely the one in which the machine observes a compliance without error where as a defection is correctly observed with probability p. The conditional probability representation for this will be $p(c|C) = 1$; $p(d|C) = 0$; $p(d|D) = p$; $(c|D) = (1 - p)$. The third environment is an imperfect information environment in which the machine observes both compliance and defection with a certain probability. For simplicity we will assume that the probability of observing a vote correctly is the same for votes which are for the machine and against the machine. In other words we have $p(c|C) = p(d|D) = p;\ p(d|C) = p(c|D) = (1-p)$. Let I_o, I_1 and I_2 represent these information environments respectively.

The main theoretical finding in Stokes, 2005 on which all comparative statics is based on is the following:

> *..., the set of voters who would sell their votes in exchange for a private benefit is the set whose ideal point, x_i, satisfies*

$$x^* \leq x_i \leq x^* + \lambda(b/x_2 - x_1)$$

where $\lambda = \frac{p\beta}{(1-\beta+p\beta)}$. Stokes uses I_0 in some parts of the model and I_1 in other parts. The main argument of this part of the paper is that we should use I_2 for the whole model. This changes the stage game and this in turn changes the comparative statics results. I also show that similar critiques are applicable to a more recent paper by Nichter [Nichter, 2008] who uses the same modeling approach but frames the issue as a "turnout buying" problem rather than "vote buying".

The model starts with a one-shot game in which parties can perfectly observe individual votes, i.e starts with the I_0 environment. The basic argument is that the voter who gets a reward b from the machine will vote for the machine if and only if

$$b - (x_i - x_1)^2/2 \geq -(x_i - x_2)^2/2$$

or equivalently

$$x_i \leq x^* + (b/(x_2 - x_1)$$

The voters with $x^* < x_i < x^* + b/(x_2 - x_1)$ are referred as *Weakly opposed* and if the value of the vote to the machine, v, exceeds b, the machine and the weakly opposed voter are in a prisoner's dilemma with the following normal form: [1]

	Reward	No Reward
Comply	$-(x_i, x_1)^2/2 + b, v - b$	$-(x_i, x_1)^2/2, v$
Defect	$-(x_i, x_2)^2/2 + b, -b$	$-(x_i, x_2)^2/2, 0$

The model then proceeds to the repeated interaction and starts using the I_1 environment. For this reason, in fact, the stage game is not a normal form repeated prisoner's dilemma game but an extensive form game given in Figure 3.1.[2] Next, the condition under which the grim-trigger strategy pair constitutes an equilibrium given

[1] One conclusion is that $x^* < x_i$ part of the inequality from the observation that voters whose ideal points are closer to the machine are going to vote for the machine without a reward anyway.

[2] Using the normal form stage game instead of the one given in Figure 3.1 changes the first period payoff of the voter when he is caught.

in equation 3.3 and comparative statics from that equation are derived. There are

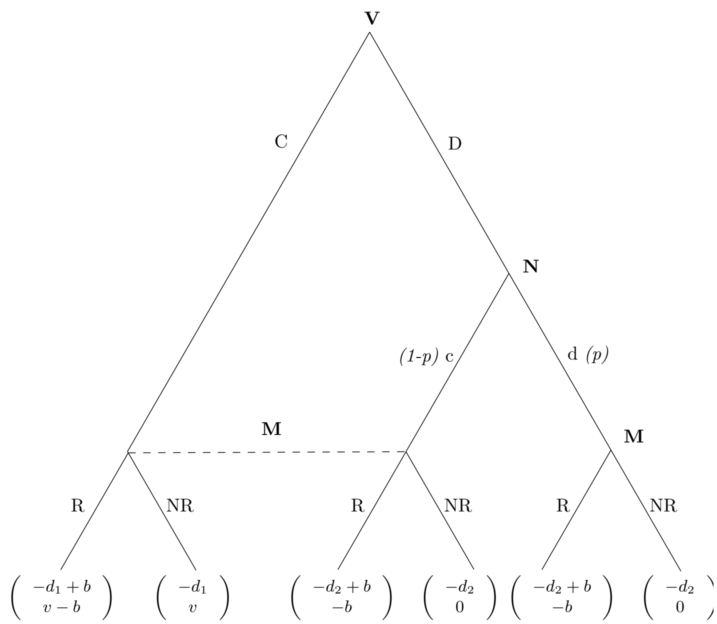

Figure 3.1. The stage game in Stokes's model

The stage game in Stokes's model. (V:Voter, N:Nature, M:Machine)

three main comparative statics derived from this model:

- The machine is most effective when it targets weakly opposed voters.
- The more accurately the machine can monitor voters, (an increase in p) the greater the potential for vote buying.
- As the value of the reward relative to the value of voting in accordance to one's policy preference increases the potential for vote buying increases.

To stress it once more, by using the I_1 assumption, we are assuming that the machine can observe a compliance action correctly with probability 1 but observes a defection correctly with probability p. I argue that there is no justification for this. Why should there be a difference in how good a machine observes a vote depending on whether the vote is for the machine or against it? After all it is *the vote* that the machine observes and the information imperfection should apply to all votes, whether it is a vote for the machine or against it. In none of the empirical studies which focus on monitoring individual votes support for such a distinction is found [Whyte, 1965, Callahan and McCargo, 1996].

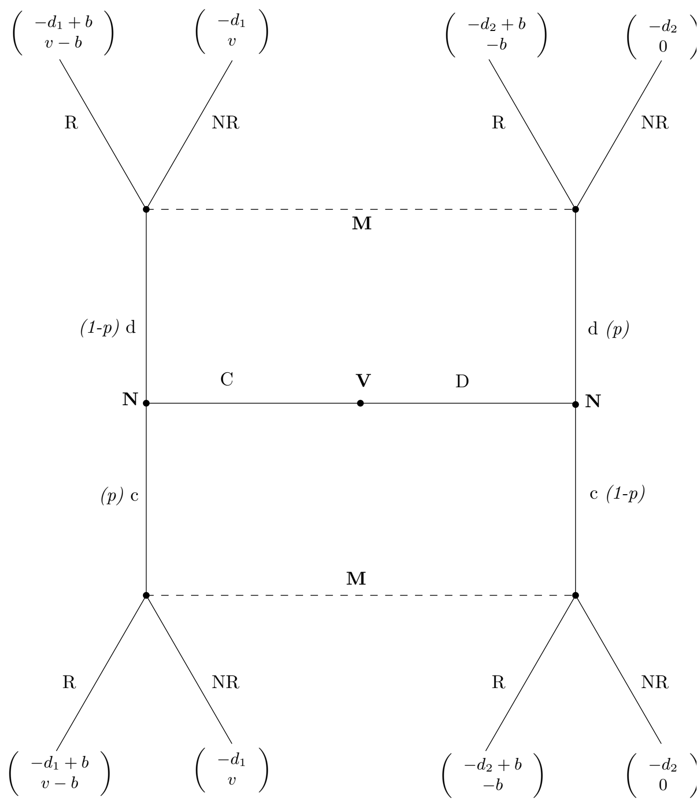

Figure 3.2. The stage game under I_2 assumption

The stage game under I_2 assumption. (V:Voter, N:Nature, M:Machine)

In the rest of this paper I will relax this assumption and resolve the model using the I_2 environment assumption throughout the model. Note that the stage game of the repeated dynamic will now be given by Figure 3.2. I will show that the comparative statics results Stokes finds do not follow. I will also discuss the similar issue in a more recent paper by Nichter who uses the same model to model turnout buying rather than vote buying and illustrate the differences I_2 environment creates in that setting.

3.3.1 Resolving the Vote Buying Model Under Imperfect Information

I start by solving the benchmark one-shot model solved under the I_0 assumption using the I_2 assumption this time. In this case the condition under which the voter will vote for the machine will be:

$$p(b - d_1) + (1 - p)(-d_1) \geq p(-d_2) + (1 - p)(b - d_2)$$

or

$$p \geq \frac{d_1 - d_2 + b}{2b}$$

Substituting for d_1 and d_2 gives us

$$x_i < \frac{2b - 4p^2 + x_1^2 - x_2^2}{2(x_1 - x_2)}$$

Depending on the values of the parameters, the right hand side of this inequality can even be greater than x_2.[3] Hence this one-shot benchmark condition does not gives us information about the voters the machine will target under I_2 assumption.

[3] Namely when $b < \frac{1}{2(4p^2 - x_1^2 + 2x_1x_2 - x_2^2)}$ and $p > \frac{1}{2(x_2 - x_1)}$

Before I solve the repeated game model under I_2 assumption, let me summarize the basic theory behind it. I start with stating one of the most powerful results of the theory of repeated games.

Proposition 1 (One-deviation property) *A strategy profile in an infinitely repeated game with a discount factor δ less than 1 is a subgame perfect equilibrium if and only if no player can increase his payoff by changing his action at the start of any subgame in which he is the first-mover, given the other players' strategies and the rest of his strategy. [Osborne, 2004]*

Hence to be able to find the equilibrium conditions for a given pair of strategies using the one-deviation property, we have to compare two utilities for each subgame and for each player: the utility the player gets from adhering to his strategy given that the other player adheres to his strategy, and the utility the player gets by deviating for one period and then sticking to his strategy given the other player sticks to his strategy. The conditions under which the first utility is greater than the second one will give us the equilibrium condition. Since an infinitely repeated game has an infinite number of sub-games, we have to identify the partition of sub-games which yield the same equilibrium conditions for our choice of strategy pairs. In our case, for the grim-trigger type strategies there are two classes of sub-games: sub-games following the initial history or a history consisting of $(Comply, Reward)$ in each period and any other subgame. We will look at sub-games of first type for the voter since it will yield the comparative statics we need. The utility from sticking to a grim-trigger type strategy will differ in the case where the machine also imperfectly monitors a vote for itself. There is always a chance that a voter is erroneously “caught” as a defector even if he has voted for the machine. This can happen at any period. Hence the total utility from sticking to your strategy will be an infinite sum which includes all such possibilities, namely, being“caught” in period 1, in period 2 and so on. If the voter is “caught” in period 1, the utility from sticking to his strategy will be

$$u_1 = (1-p)(-d_1 - d_2\delta - d_2\delta^2...) = -(1-p)(d_1 - \frac{\delta d_2}{1-\delta})$$

Similarly, being caught in later periods yield different payoffs.

Period 2:

$$u_2 = p(1-p)[(b-d_1) - d_1\delta - d_2\delta^2 - d_2\delta^3...)$$

Period 3:

$$u_3 = p^2(1-p)[(b-d_1) + (b-d_1)\delta - d_1\delta^2 - d_2\delta^3 - d_2\delta^4...)$$

So the general form of these payoffs for periods $n \geq 2$ is

$$u_n = p^n(1-p)[(b-d_1)(1+\delta+\delta^2+...\delta^{n-2}) - d_1\delta^{n-1} - d_2(\delta^n + \delta^{n+1} + ...)]$$

Hence the total payoff from adhering to the GT-type strategy is given by

$$
\begin{aligned}
u &= \sum_{n=1}^{\infty}\left[p^n(1-p)[(b-d_1)(1+\delta+\delta^2+...+\delta^{n-1})-d_1\delta^n-d_2(\delta^{n+1}+\delta^{n+2}...)\right]+u_1 \\
&= (1-p)\sum_{n=1}^{\infty}\left[p^n[b(1+\delta+\delta^2+...+\delta^{n-1})-d_1(1+\delta+\delta^2+...+\delta^n)-d_2\delta^{n+1}(1+\delta+\delta^2...)]\right]+u_1 \\
&= (1-p)\sum_{n=1}^{\infty}p^n\left[b\frac{1-\delta^n}{1-\delta}-d_1\frac{1-\delta^{n+1}}{1-\delta}-d_2\frac{\delta^{n+1}}{1-\delta}\right]+u_1 \\
&= \frac{(1-p)}{(1-\delta)}\left[b\sum_{n=1}^{\infty}p^n(1-\delta^n)-d_1\sum_{n=1}^{\infty}p^n(1-\delta^{n+1})-d_2\delta\sum_{n=1}^{\infty}(p\delta)^n\right]+u_1 \\
&= \frac{(1-p)}{(1-\delta)}\left[\left[\sum_{n=1}^{\infty}p^n\right](b-d_1)-\left[\sum_{n=1}^{\infty}(p\delta)^n\right](b-\delta d_1+\delta d_2)\right]+u_1 \\
&= \frac{(1-p)}{(1-\delta)}\left[\frac{b-d_1}{1-p}-\frac{(b+\delta d_2-\delta d_1)}{1-p\delta}\right]-(1-p)(d_1-\frac{\delta y d_2}{1-\delta}) \\
&= -\left(\frac{b(p-p\delta)-d_2(-1+p)\delta(-2+p\delta)-d_1p\left(1-(-2+p)\delta+(-1+p)\delta^2\right)}{(-1+\delta)(-1+p\delta)}\right) \qquad (3.1)
\end{aligned}
$$

If the voter deviates for one period, there are two possibilities, with probability p he will be caught and gets a payoff $-d_2$ forever or with probability $1-p$ he can avoid detection and get $b-d_2$ in the first period and get a continuation payoff of u. Thus the expected payoff from a single deviation will be given by

$$
\begin{aligned}
u' &= p(-d_2-\delta d_2-\delta^2 d_2...)+(1-p)[(b-d_2)+\delta u] \\
&= \frac{-pd_2}{1-\delta}+(1-p)[(b-d_2)+\delta u] \qquad (3.2)
\end{aligned}
$$

Hence the equilibrium condition is

$$u \geq u'$$

$$-\left(\frac{b(p-p\delta)-d_2(-1+p)\delta(-2+p\delta)-d_1 p\left(1-(-2+p)\delta+(-1+p)\delta^2\right)}{(-1+\delta)(-1+p\delta)}\right)$$
$$\geq \frac{-pd_2}{1-\delta}+(1-p)[(b-d_2)+\delta u] \quad (3.3)$$

A series of algebraic manipulation and substituting for u gives us this condition as

$$K \geq 0$$

where K is

$$\frac{\left(b(-1-2p(-1+\delta)+\delta)+(1+(-1+p)\delta)\left(d_2\left(1+p^2\delta^2-p\delta(1+\delta)\right)+d_1\left(2(-1+\delta)+p^2(-1+\delta)\delta-p\left(-1+2\delta+\delta^2\right)\right)\right)\right)}{(-1+\delta)(-1+p\delta)}$$

Although this expression is complicated we can still find the comparative statics we are interested in. We will now compare the four comparative statics derived from these two models.

1. **Targeting**

   We know that the machine is not going to target the voters who will vote for the machine even if they receive no reward, i.e, those voters with an ideal position in the interval $x^* \geq x_i$. The more interesting question is about the upper bound of the ideal positions of the voters whom the machine targets. As it is stated earlier, this upper bound was found to be $x^* + \frac{bp\delta}{(x_2-x_1)(1-\delta+p\delta)}$ in Stokes's model.

In our model, on the other hand, if p is sufficiently low and δ is sufficiently high, the machine can even target voters with ideal points $x_i > x_2$[4]. So there is significant difference between these two models in this respect. The conditions for this result are complicated. However this is not important because for us the important part is to show the possible existence of vote buying for different types of voters rather than the exact form of the bounds of these inequality conditions.

2. **Monitoring**

   One of the findings of Stokes's model is that the higher the ability of the machine of observing votes, the greater the potential for vote buying. In our model, this is not immediately clear. Depending on the values of other variables an increase in p might not increase the potential of vote buying.[5]

3. **Reward Value**

   As I stated earlier, according to Stokes's model, as the value of the reward relative to the value of voting in accordance to one's policy preference increases the potential for vote buying increases. In my model, on the other hand, it depends on the value of p. A higher reward makes the vote buying equilibrium more likely if and only if $p > \frac{1}{2}$. In other words increasing the reward does not help the machine if it cannot observe the votes with at least 1/2 probability.

[4]Namely when $2 - \sqrt{3} > p$ and $\delta > (-2 + 3p + p^2)/(4(-1 + p)p) - 1/4\sqrt{(4 - 20p + 21p^2 - 2p^3 + p^4)/((-1 + p)^2 p^2)}$

[5]One such example is when the discount factor is in a particular interval and the reward is sufficiently small. The exact cases and conditions are too complicated to present here. See the attached Mathematica output.

Moreover, the effect of an increase in the reward on the likelihood of vote buying increases with p. To put it in another way, the higher the p is the higher the effect of an increase in the reward will be.

3.3.2 The Case of Turnout Buying

In a more recent paper [Nichter, 2008] an alternative explanation for the same question is asked: "with the secret ballot, what prevents individuals from accepting rewards and then voting as they wish ?". He uses the same modeling approach but focuses on turnout buying rather than vote buying. We will use the same notation for this model.

The basic idea is that the machine will target those voters who will vote for the machine if they do, but may not vote due to a voting cost, c. The payoffs in this case is given by the following matrix[6]:

	Reward	No Reward
Comply	$-(x_i - x_1)^2/2 + b - c, v - b$	$-(x_i - x_1)^2/2 - c, v$
Defect	$b, -b$	$0, 0$

The same critique about the form of the stage game and about the assumptions about the informational environment apply to this study as well. The only thing that changes is what compliance and defection refers to in the definitions of I_0, I_1 and

[6]Note that my claim about this normal form game not being the stage game of the repeated dynamic applies for this case as well.

I_2. Compliance in this study refers to "Vote" and defection refers to "Not Vote". Similar to Stokes, Nichter assumes the I_1 information environment implicitly, i.e, he assumes that the machine can observe the ones who show up at the polls perfectly where as it might make mistakes while deciding about the ones who don't show up. If we resolve Nichter's model by relaxing the "semi-imperfect" environment assumption the equilibrium condition we have is[7]

$$u - \frac{(1-p)b}{(1-(1-p)\delta)} \geq 0$$

where u is the expected payoff from adhering to grim-trigger strategy and is equal to

$$u = \frac{(1-p)}{(1-\delta)} \left(\frac{b - d_1 - c}{(1-p)} - \frac{b + \delta d_1 + c\delta}{1 - p\delta} - (c + d_1)(1-\delta) \right)$$

Next I do the same comparative statics comparison for this model.

1. **Targeting**

   Nichter's model predicts that machines will target immobilized supporters. The turnout model under I_2 assumption makes a more clean prediction in the same direction: the machine will target voters with $x_i < x_1$.

2. **Monitoring**

   The adjusted model and Nichter's model yields the same result in this case,

[7] I am not presenting the calculations here since they are basically the same as the model above. The only difference is in the functional form of the payoffs.

the higher the observing ability of the machine, p, the more likely the turnout buying will take place.

3. **Reward Value**

   Similar to Stokes's model, Nichter's model predicts that turnout buying will become more probable if the reward gets higher. Our adjusted model on the other hand concludes that an increase in the reward will increase the chances of turnout buying if an only if $p > 1/2$. Furthermore, the effect of an increase in the reward on the likelihood of vote buying increases with p.

3.4 The Spatial Model of Valence Buying

In this section we introduce a spatial model of clientelism. The approach uses a latent utility based stochastic voting model and utilizes the concept of "valence" in order to capture the benefit the parties receive from rewarding the voters. So in terms of explaining the mechanism underlying the clientelistic linkage —why would not voters just take the money and run question—the argument is quite straightforward: the materialistic benefits distributed by the political agents—let these be parties or candidates—is one way of increasing their valences. Although less sophisticated, I argue that this is quite a natural setup.

In contrast to regular spatial voting models, there are two decision variables which the parties are trying to choose: a policy position and an allocation vector which specifies which voters to reward and how much. These two decisions are made simultaneously and this opens the way to many interesting comparative statics as we will discuss below in mor detail. Also in contrast to the clientelistic politics literature, we argue that bribing is not always limited to a single machine party the mode presented will allow each party to reward the voters[8]. Moreover, we will introduce constraints on the amount of resources available parties, possibly different from each other. By doing this we will able to introduce an interesting variation to the problem: different levels of resources available to different parties.

The concept of valence has been used to explain the non-convergence of policy positions of the parties. A rather recent stream of works by Norman Schofield and his coauthors [Miller and Schofield, 2003], [Schofield and Sened, 2006], [Schofield et al., 2009], [Schofield et al., 2011] are in this very direction. In the model underlying these papers however, valence is exogenous, i.e, it is not something that the political agents make a decision on. It is rather what is not explained by other components of the voter utility function. In the present model, the valences are the rewards given by the parties to the voters. Our model will be an extension of the setup originally developed in [Schofield, 2003].

One recent work which studies the "valence buying" problem is Ashworth and Bueno de Mesquita, 2009. Their setup is however very different than ours. Even if the level of the valence is a decision variable for the party, it is constant for all individuals. So their setup does not capture the individual level vote buying phenomenon hence

[8]Serdar Denktas, the leader of one of the two major parties in Northern Cyprus, had made the following statement after losing the elections in 2007: "The bargaining for votes started from 350 TL but then dropped. I bought votes 75 TL per voter. The other parties had more money, so they bought those votes which we could not have." (Turkish newspaper article available at http://yenisafak.com.tr/Politika/?t=27.04.2009&c=2&i=183054. Accessed on Aug 04 2012.)

do not address the main question we are interested in: "Which voters will be targeted and how much?".

3.4.1 The Model

Let N be the set of the electorate of size n and P be the set of political parties of size p. The voters are characterized by their ideal points, $\{x_i \in X\}_{i \in N}$, and the parties are characterized by their policy positions, $\{z_j \in X\}_{i \in P}$, where X is a compact convex subset of Euclidean space, $\mathbb{R}^2$.

The utility of voter i with an ideal point of x_i from voting for party j with a policy position of z_j is given by

$$u_{ij}(x_i, z_j, \lambda_j) = \lambda_{ij} - \beta ||x_i - z_j||^2 + \epsilon_j \tag{3.4}$$

where λ_{ij} is the valence of the party j as perceived by voter i, $||x_i - z_j||^2$ is the Euclidean norm, and the β is the parameter which determines the weight the voter assigns to the policy relative to the valence.

The party is trying to maximize its vote share and has two decision variables in order to do that. One is the policy position on the policy space, z_j, and the other is a "bribing vector" λ_j, which is essentially an allocation vector of size n, which specifies the amount of bribe the party is going to give to each voter subject to the budget constraint. So the strategy of a party is a pair (z_j, λ_j). Let $\mathbf{z}$ denote the vector of party positions and $\mathbf{\Lambda}$ denote the vector of allocation vectors chosen by the parties.

So the probability that voter i chooses party j is given by

$$\rho_{ij}(\mathbf{z}, \mathbf{\Lambda}) = Pr[(u_{ij}(x_i, z_j, \lambda_j) > u_{ij}(x_i, z_k, \lambda_k), \forall k \neq j] \tag{3.5}$$

We will employ the usual logit link between the stochastic latent utility and the probability of voting [Train, 2003]:

$$\begin{aligned}
\rho_{ij}(\mathbf{z}, \mathbf{\Lambda}) &= Pr\left[(\lambda_{ik} - \lambda_{ij}) - \beta(||x_i - z_k||^2 - ||x_i - z_j||^2) > \epsilon_j - \epsilon_k, \forall k \neq j\right] && (3.6)\\
\rho_{ij}(\mathbf{z}, \mathbf{\Lambda}) &= \left[1 + \sum_{k \neq j} \exp[(\lambda_{ik} - \lambda_{ij}) - \beta(||x_i - z_k||^2 - ||x_i - z_j||^2)]\right]^{-1} && (3.7)
\end{aligned}$$

The objective of the political party is to maximize its expected vote share by choosing an appropriate policy position and an allocation vector, subject to the budget constraint and given the choices of other parties:

$$\begin{aligned}
\max V_j(\mathbf{z}, \mathbf{\Lambda}) &= \frac{1}{n} \sum_{i \in N} \rho_{ij}(\mathbf{z}, \mathbf{\Lambda}) && (3.8)\\
\text{s.t} \quad & \sum \lambda_{ij} \leq B_j && (3.9)
\end{aligned}$$

where B_j is the total resources available to party j. Now we can define the equilibrium concept we are using.

Definition 1 *The vector* $(\mathbf{z}^*, \mathbf{\Lambda}^*) = [(z_1^*, \lambda_1^*), ..., (z_p^*, \lambda_p^*)] \in (X^p, \mathbb{R}^{np})$ *is a* local Nash equilibrium (LNE) *of the game of machine politics if and only if for all* $j \in P$*, there exists a neighborhood* $X_j \times \mathbb{R}_j^n$ *of* (z_j^*, λ_j^*) *in* $(X \times \mathbb{R}^n)$ *such that*

$$V_j((z_1^*, \lambda_1^*), ..., (z_j^*, \lambda_j^*), ..., (z_p^*, \lambda_p^*)) > V_j((z_1^*, \lambda_1^*), ..., (z_j, \lambda_j), ..., (z_p^*, \lambda_p^*))$$
$$\forall (z_j, \lambda_j) \in (X_j \times \mathbb{R}_j^N) - \{(z_p^*, \lambda_p^*)\}$$

We know that this equilibirium exists [Schofield, 2003]. What we are interested about this equilibrium concept is its predictions given the empirical estimates of the voter positions and budget constraints of the parties. For this purpose we move on to do the simulations of this equilibrium using empirical data.

3.4.2 Simulating the Equilibria

One advantage of using a stochastic voting model is the possibility of working with any empirical distribution of voter positions, and simulate the equilibrium predictions [Schofield and Sened, 2006]. In our case, the moving components are the voter ideal positions and relative budgets available to parties. In fact, given an estimate of voter ideal points on a policy space and the limits on the "clientelistic budgets" available to parties, we can simulate different equilibrium predictions which specifies the policy positions choices of the parties together with whom the parties would bribe by which amount for a given β.

The steps involved in the simulation procedure are as follows:

1. Estimate the empirical distribution of voter ideal positions from surveys using factor analysis.

2. Estimate the party positions using expert surveys/content analysis.

3. Estimate the model parameters using a multinomial logit model. (In this case, it is the distance coefficient, β.)

4. Use these model parameters to simulate the equilibria using a MATLAB code. (Positions and "bribing vectors".)

5. Get insights into questions like:

   (a) Which party will target which voters?

   (b) How would the possibility of bribing affect the policy position of parties?

   (c) How would budget differentials affect the policy position of parties?

I present some simulation results using data for US 2008 elections. The data and the estimates comes from an earlier work [Schofield et al., 2011]. By doing this, I do not argue that clientelistic linkages exist in the US, rather I am using it as an experiment to illustrate the empirical implications of the model. Since it is the relative sizes of the budget which is important, I will arbitrarily choose some budget sizes. Also, since the simulation takes quite a long time, I only use some random 150 voters from the original dataset.

The figures in Appendix A illustrates some equilibrium outcomes for different budget differentials[9]. Since the strategy space available to parties quite rich, the multiplicity of equilibria is inevitable. In fact these "local equilibria" are highly dependent on the initial values of the decision variables. The size of the circles around the voter ideal points are proportional to the reward they receive in the equilibrium. The red filled circle is the reward delivered by Party 1 and the blue circle is the reward delivered by Party 2. There are voters who have not received any reward, they are represented just by "x". Figure 3.3 shows the pure strategy equilibrium with no clientelistic linkage. Obviously the parties converge to the electoral mean in this case [Schofield, 2004].

Figures 3.4, 3.5 and 3.6 show the equilibria for the case where the budget difference between parties is relatively small and for different initial policy positions for the parties. (All initial reward vectors are taken as zero vectors.) There are a number of things to note here. First of all, the existence of a clientelistic linkage diverges the equilibrium party positions away from the electoral mean. The second point is that there can be multiple equilibria where different policy positions are supported by different rewarding vectors. Apart from the voters who do not receive any reward from any party, there exist voters who receive rewards from both of the parties. This might seem problematic at first sight. This however is a stochastic voting model, and the fact that one party gives more reward than the other party does not necessarily

[9]The initial party positions are $((0,0),(0,0))$ for Figures 3.3, 3.5 and 3.8; $((1,-1),(-1,1))$ for Figures 3.4, 3.6 and 3.7; and $((0,0),(0,0),(0,0))$ for Figures 3.9, 3.10 and 3.11. The spatial coefficient, β, is 0.756 in all cases.

mean that the voter is going to vote for the party, rather it means that it increases the chances of the voter voting for that party. Figures 3.7 and 3.8 show the cases for a much larger budget differential between the parties. As it can be seen, there can be an equilibrium where the party positions converge and the "rich" party makes a more homogeneous reward distribution and an equilibrium where the party positions diverge and the rich party distributes more benefits to the voters who are on the "one side of the policy space" albeit rewarding everyone. In both cases the richer party gets 98% of the votes.

Next we move on to the case where we introduce a third party. Figure 3.9 shows an equilibrium where there is no budget difference among parties. The parties are close to the electoral mean and they all have a relatively homogeneous reward allocation. Figure 3.10 illustrates the case where there is only one machine party. The machine party is almost at the electoral mean and distribute benefits to all voters. The other two parties are located at the same position, which is not surprising. The case where there are two machine parties is illustrated in figure 3.11. The machine parties are at the center where as the other party located outside (close to lower left hand corner). This observation is in line with the previous observations [Schofield et al., 2009] that since they have a comparative advantage in attracting the votes given everything else constant, the high valence parties will locate them closer to the electoral mean and the low valence parties will stay out where they can get some votes through a comparative advantage they have in ideological proximity.

3.5 Conclusion

The model developed in this paper has a potential to provide insights in problems which are quite different than the clientelistic linkage. In fact any setting in which actors are trying to maximize support by choosing a policy position and a bribing vector, hence any ”support buying” situation, will be a potential application. For an example, consider the voting game in the United Nations Security Council. It has been demonstrated that bribing is actually happening in the Security Council through different means such as foreign aid or IMF lending [Kuziemko and Werker, 2006]. In this context, one can extend the model such that the countries who are willing to be bribed also choose their policy positions strategically to get the highest benefit possible. Considering the fact that there will be domestic political constraints on the countries' choices in the international arena, this might be interesting way to look at the “two-level game” logic within the context of “large country bribing the small country.” In this case for instance, the domestic constraint on the large countries will be about the “bribing strategy” where as for the small country it will more about the degree to which the country should make compromise in return for the benefit it receives. How the institutional rules affect the equilibria would be an interesting direction to pursue.

There are many other tasks which are left for further studies. Replicating the simulation for different countries and different electoral systems to gain real insights is a crucial next step. One important theoretical extension would be to relax the exogenous budget constraint assumption. This can be done through introducing lobby groups, who provide financial support to political parties, to the model. So in this setup, while deciding which policy position to choose, the parties will try to maximize their "income" as well as capturing votes by minimizing the distance. Another related point is that in the current formulation of the model, the unused portion of the budget do not affect the utility of the party hence parties distribute all of their budgets. The model can be extended so as to incorporate this.

Also on the empirical side, the argument that the politicians are in fact treating bribing as a tool for valence buying should be supported using further evidence. There is a clear cognitive aspect of the problem related to voters. The voters might feel guilty if they received the benefit and fail to vote for that party. This might be more valid in cultures where "doing something in return of what you receive is a form integrity." So experimental studies with voters would provide useful information. Surveys with politicians is also natural candidate for this purpose.

Monitoring groups of voters then rewarding and punishing them is much more efficient than monitoring and then rewarding and punishing individuals especially where party networks are weak and the electorate is large and dispersed over wide geographical areas [Kitschelt and Wilkinson, 2007a]. This fact is further strengthened by the existence of groups with high level of cohesion such as ethnic or religious groups. In these cases the clientelistic linkage can be sustained through a self-enforcing group equilibrium where within monitoring and punishment mechanisms prevail [Auyero, 2001]. Using the model presented here to investigate this mode of citizen-politician linkage is not difficult. Theoretically, nothing would be different, just interpret voters as groups. Using the survey data we can find the median voter for each of these relevant groups (electoral districts, ethnic groups, religious minorities etc.) and run the simulation treating these median voters as the set of voters and weighting them with the respective sizes of the groups they belong to.

The relationship between the political space, including the distribution of voter ideal positions, and the possibility of bribing is another question which deserves attention. When is "corruption" more likely? Do more polarized systems lead to higher incentives of bribing? Would the lack of a salient policy dimension increases the chances of clientelistic linkages? Investigating these and similar questions might yield valuable information which connects the clientelism discussions with other important problems of political economy.

Appendix A: Simulation Results

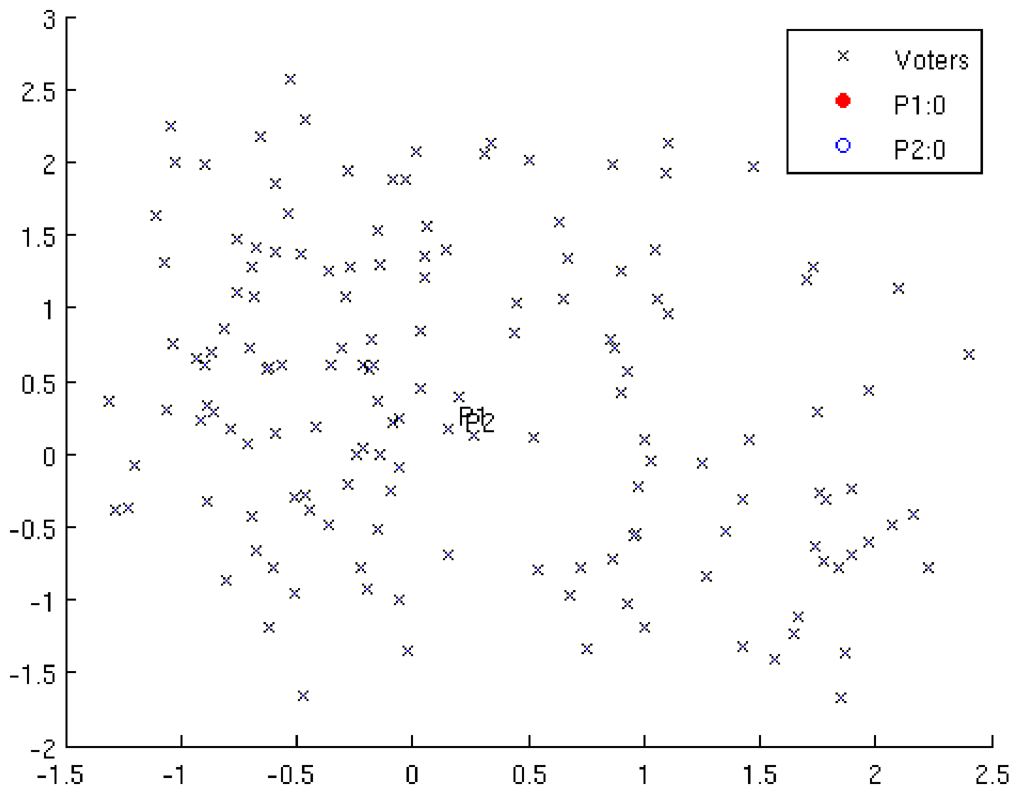

Figure 3.3. No Clientelistic Linkage

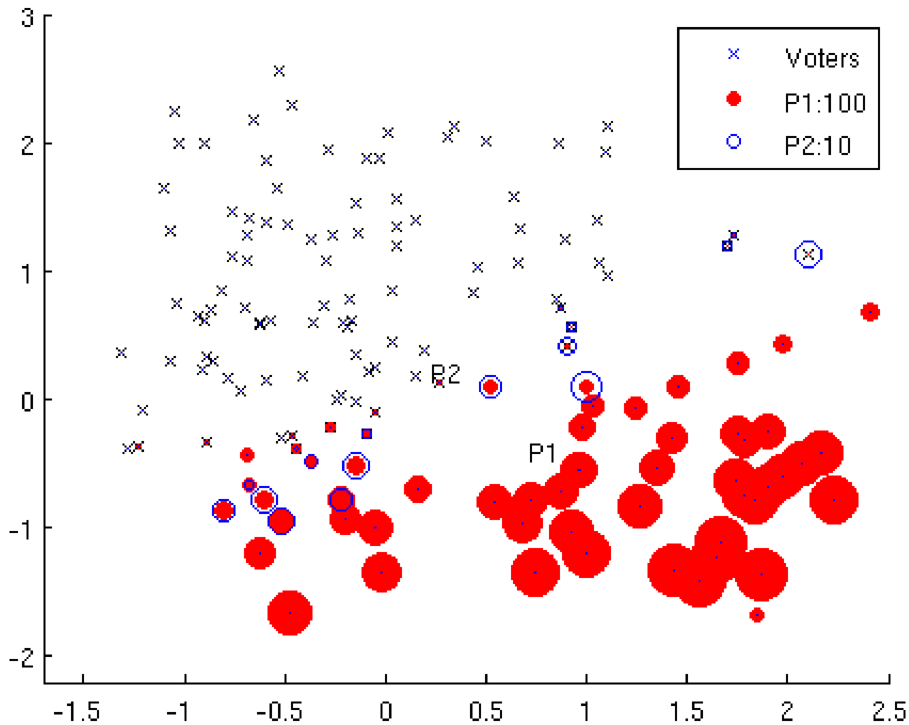

Figure 3.4. Small Budget Difference Between Parties

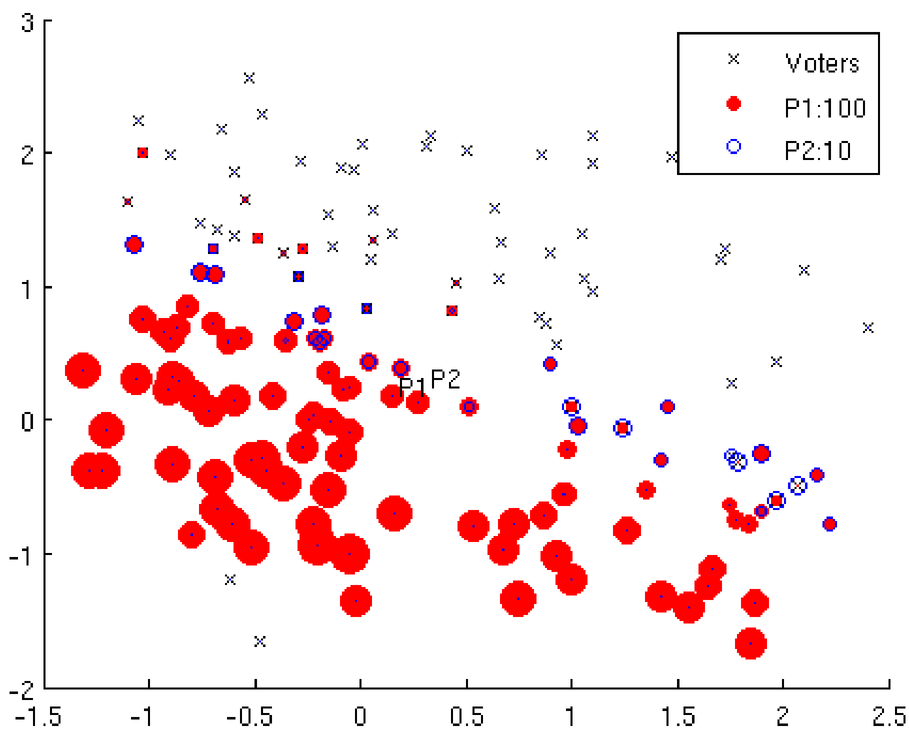

Figure 3.5. Small Budget Difference Between Parties

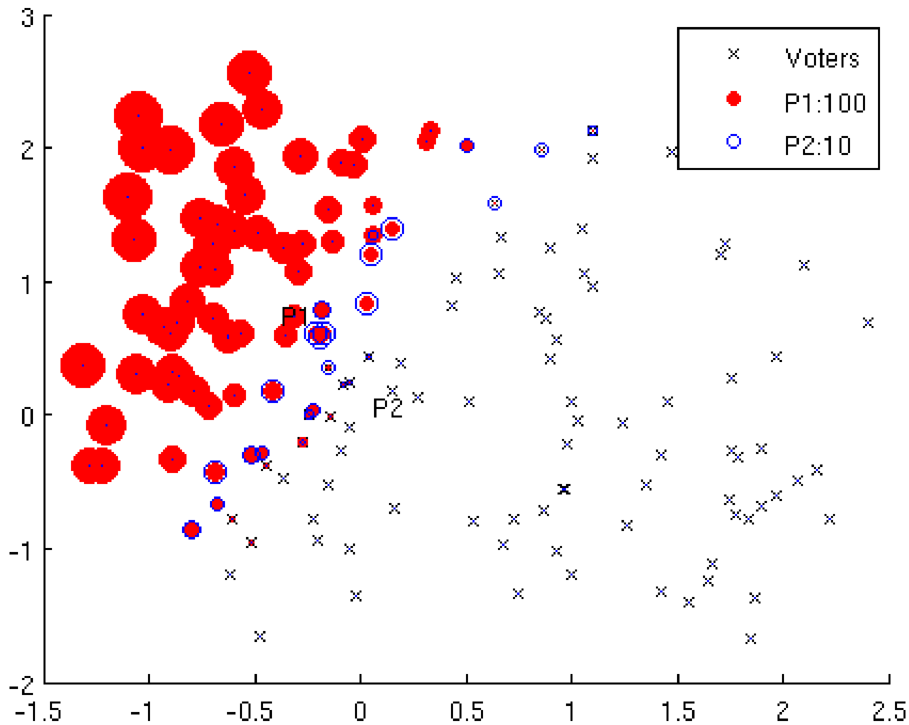

Figure 3.6. Small Budget Difference Between Parties

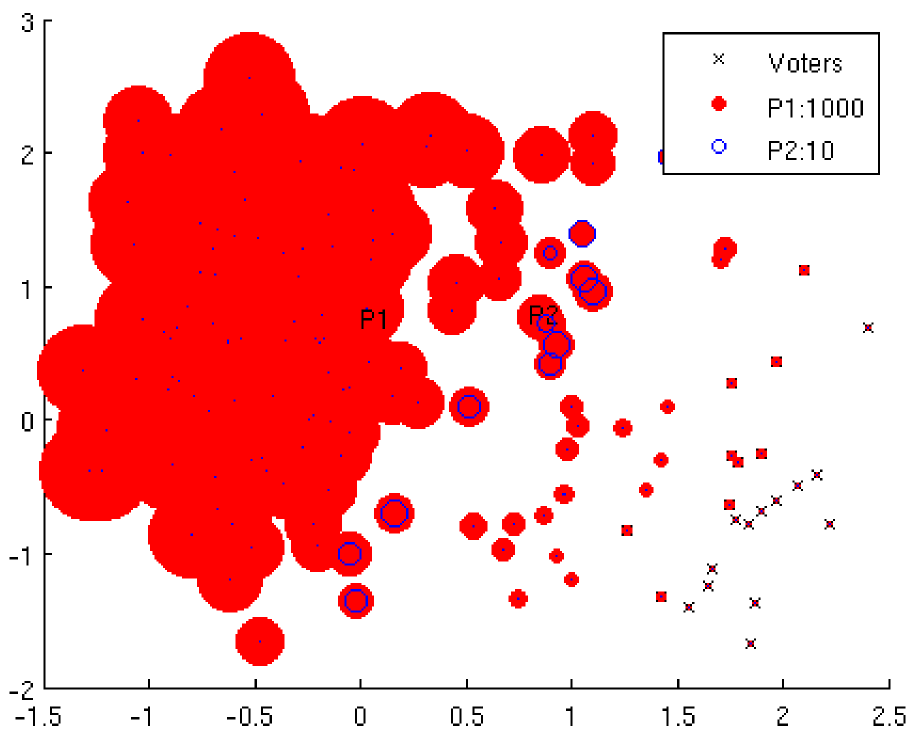

Figure 3.7. Large Budget Difference Between Parties

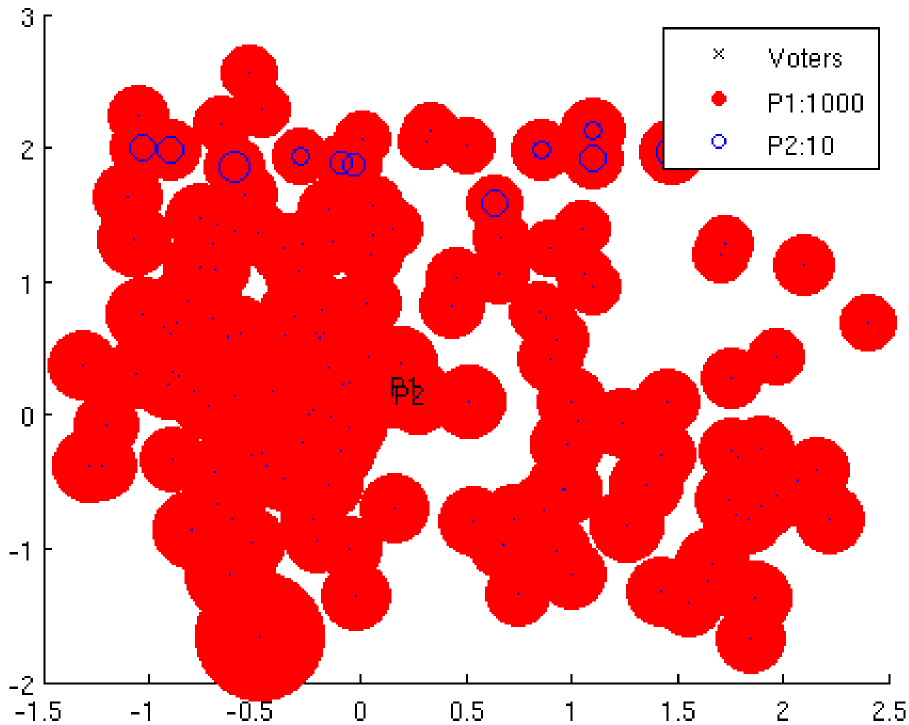

Figure 3.8. Large Budget Difference Between Parties

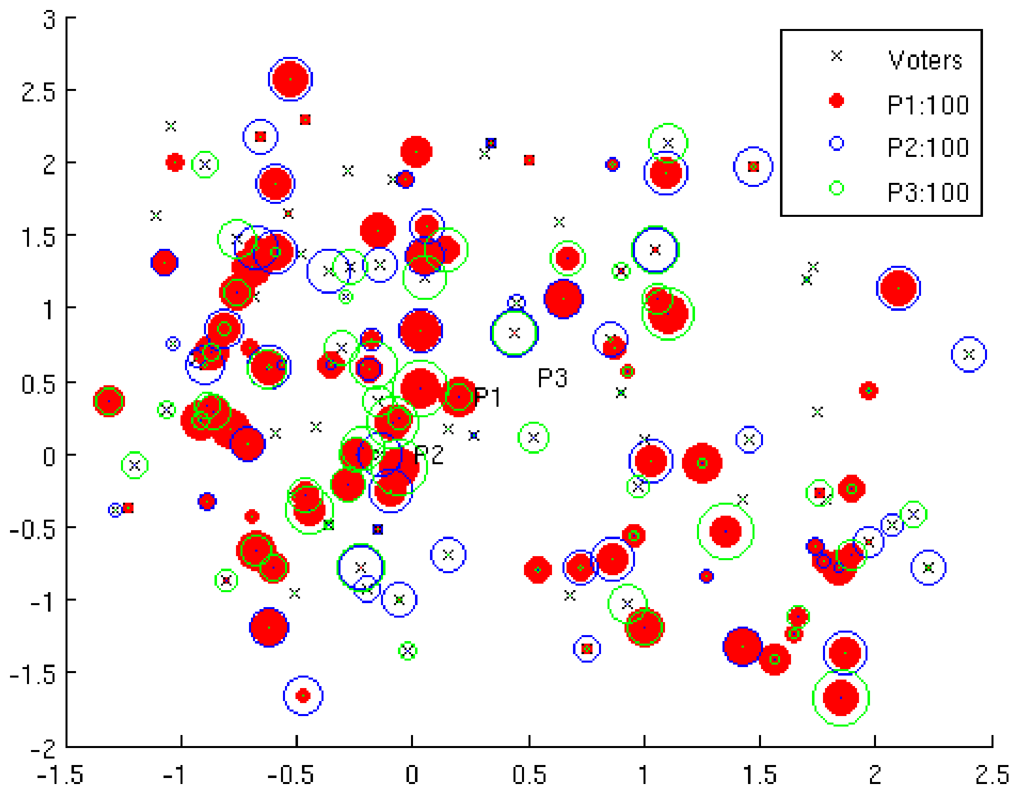

Figure 3.9. 3 Party Case: No Budget Difference Between Parties

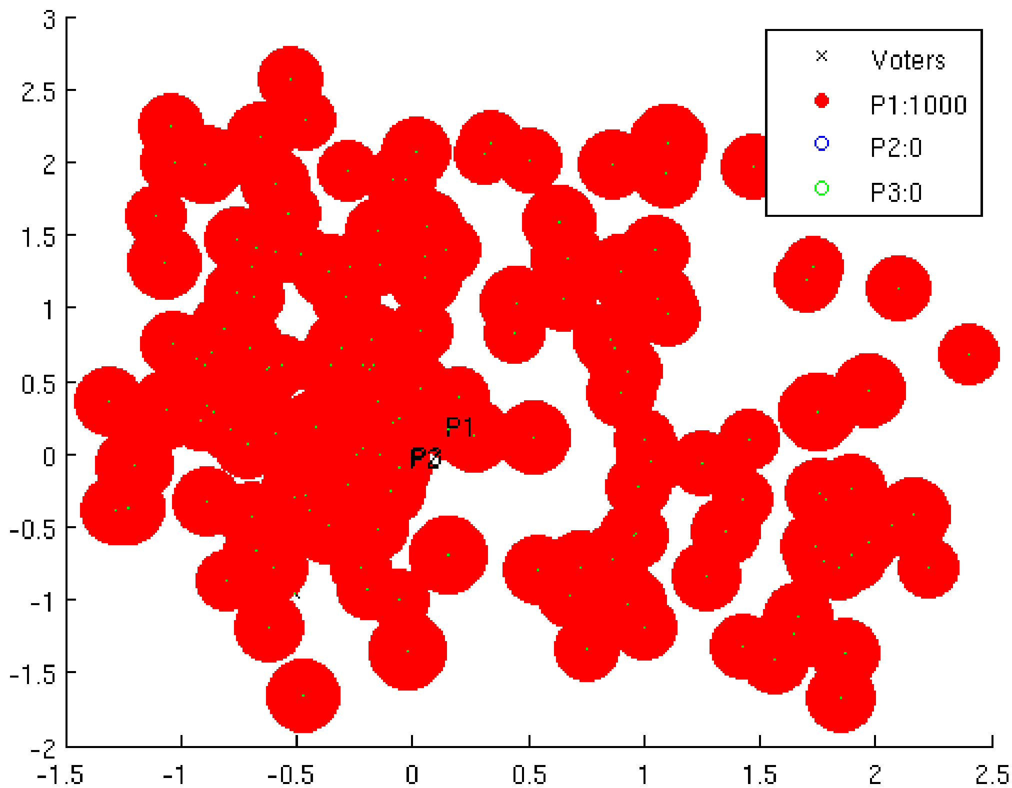

Figure 3.10. 3 Party Case: Single Rich Party Case

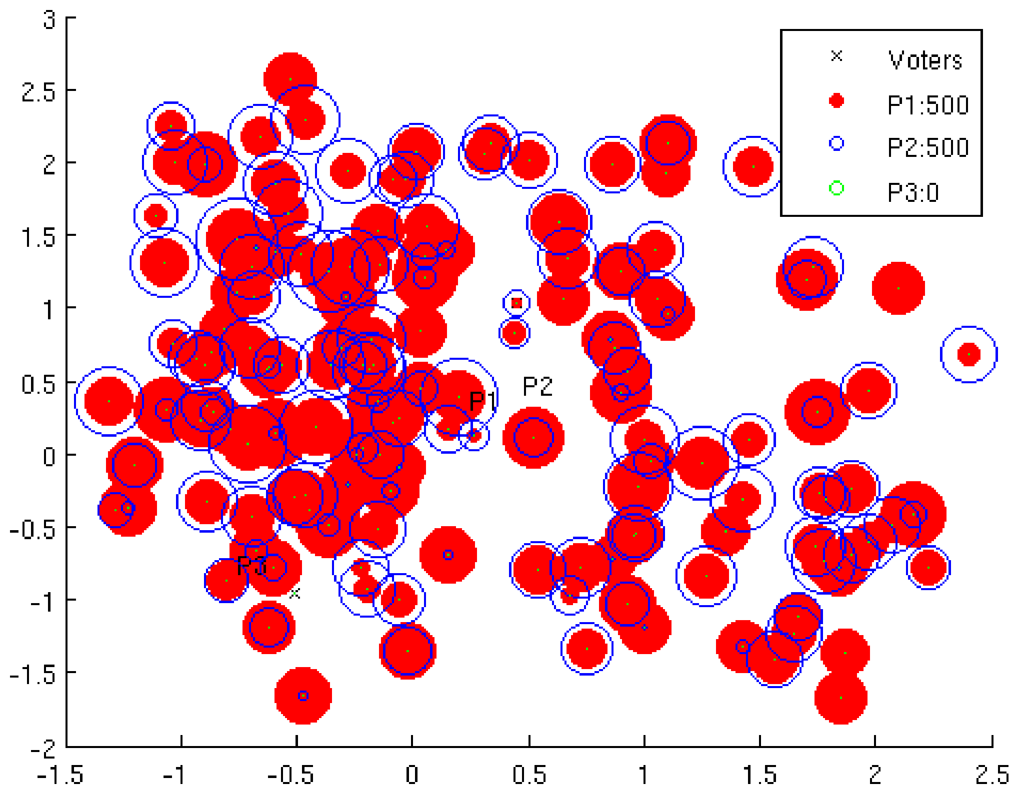

Figure 3.11. 3 Party Case: Two Rich Parties Case

Appendix B: MATLAB Codes for Simulation

```
% THIS FUNCTION FINDS THE LOCAL NASH EQUILIBRIUM BY RECURSIVELY
% CALLING THE FIND_BEST_RESPONSE FUNCTION

function f=find_local_nash(n,m,VV,beta,C,ffc,bud,tol)

% n: Number of voters
% m: Number of parties
% bud(m+2): Budgets of parties
% tol: Tolerance. The algorithm stops if the change in vote share
% is less than tol.
% VV(2xn): Voter positions
% beta(1x1): Distance coefficient
% ffc((m)x(n+2)): Initial values for decision variables.
% First two columns are policy positions of parties and the rest
% n columns are the rewards to each n voters.

max_differ=10;
differ=zeros(m,1);
shares=zeros(m,1);
while abs(max_differ)>tol
    max_differ=0;
    for k=1:m
    A0=ffc(k,:)';
    [x,fval]=find_best_response_2(n,m,k,A0,C,VV,beta,bud,ffc);
    differ(k)=max(abs(shares(k)-fval));
    if (differ(k)>tol)
        max_differ=differ(k);
    end
    ffc(k,:)=x';
    shares(k)=fval;
    end
end
f=ffc;
```

```
% THIS FUNCTION COMPUTES THE BEST RESPONSE OF A PARTY
% GIVEN OTHER PARTIES STRATEGIES

function [x,fval]=find_best_response_2(n,m,k,A0,C,VV,beta,bud,ffc)

% n: Number of voters
% m: Number of parties
% A0: Initial values
% k: Best response of which party?
% VV(2xn): Voter positions
% beta(1x1): Distance coefficient
% ffc((m)x(n+2)): Where to fix the rest of the party positions
% and bribes

A = sym ('A', [n+2,1]);

svote=0;
di2=zeros(n,m);
for i=1:n
    den=0;
    di1=(VV(1,i)-A(1))^2+(VV(2,i)-A(2))^2;
    for j=1:m
    di2(i,j)=(VV(1,i)-ffc(j,1))^2+(VV(2,i)-ffc(j,2))^2;
        if ~(j==k)
        den=den+exp(ffc(j,i+2)-beta*di2(i,j));
        end
    end
    V1=exp(A(i+2)-beta*di1)/(den+exp(A(i+2)-beta*di1)) ;
    svote=svote+V1;
end

V1=vpa(-svote/n,2);

fhc=matlabFunction(V1,'vars',{A});

Aeq=[];beq=[];lb=zeros(n+2);lb(1)=-10;lb(2)=-10;ub=[];
options=optimset('Algorithm','active-set');

[x,fval] = fmincon(fhc, A0, C, bud(k),Aeq,beq,lb,ub,[],options);
```

```
% THIS CODE PLOTS THE SIMULATION OUTPUT

n=150;
ffc = zeros(2,n+2);
ffc(:,1)=[0,0]; % First party initial position
ffc(:,2)=[0,0]; % Second party initial position
C=zeros(1,n+2)+1; C(1,1)=0; C(1,2)=0; % Constraint matrix

bud=[0,0];  %Budgets

temp=find_local_nash_2(n,m,VV,beta,C,ffc,bud,tol);

parties=temp(:,(2:3));
bribes=100*abs(temp(:,(4:n+3)));
bribes(bribes==0) = 0.00000001;
voters=VV(:,(1:n));

figure(18)
axis([-1.7 2.5 -2.2 3])
scatter(voters(1,:),voters(2,:),'x','black')
hold on
scatter(voters(1,:),voters(2,:),bribes(1,:),'filled','r')
hold on
text(parties(1,1),parties(1,2),'P1')
hold on
text(parties(2,1),parties(2,1),'P2')
hold on
scatter(voters(1,:),voters(2,:),bribes(2,:),'blue')
hold on
legend('Voters',strcat('P1:',num2str(bud(1))),
strcat('P2:',num2str(bud(2))));
```

4. Uncertainty and Institutional Stability: Microbehavioral Foundations

4.1 Introduction

As North points it out "A theory of institutional change is essential for further progress in the social sciences..." [North, 1993]. In particular, this field of research will help us better understand one of the greatest puzzles of political economy: the diverse performance of societies and economies. Developing tools for this purpose is also an important part of a bigger project of enhancing the ability of the social sciences to unpack the complexity of dynamic situations [Ostrom and Basurto, 2011]. In this paper I offer a rational choice model of conceptualizing institutional change and stability based on actors' trust in the future of the institution. The novel part of the paper is that I use stochastic discount factors as the moving component of the model, one approach which has not yet been used before in the study of institutions.

Rational choice approaches made significant contribution to the study of institutions and institutional change in the last couple of decades. The main advantage of this approach is that it demands a well-defined relationship between human behavior and institutions [Hall and Taylor, 1996]. There are different methodologies within the rational choice school of institutionalism. Calvert, 1995 distinguishes between three different views within the rational choice approach: i) institutions as features of individual preferences, ii) institutions as "rules of the game" and iii) institutions as equilibria of behavior in an underlying game.

Although this institutions as equilibria approach is very well-suited to explain the persistence of institutions, explaining institutional change remains as an important challenge. The reason is that the equilibrium, by definition, is a "no change situation" in which only shared beliefs corresponding to self-enforcing behavior can rationally prevail [Calvert, 1995]. So if institutions are equilibria of some underlying game and equilibria are steady-states, how can we explain institutional change while staying in the institutions as equilibria approach?

A new classification of the approaches in rational choice institutionalism is provided in Aoki, 2007. Aoki distinguishes between two views within the rational choice paradigm: the exogenous view and the endogenous view. The former view treats institutions as pre-determined rules outside the domain of interactions where as the latter treats those rules as something spontaneously and endogenously determined. The first approach is elaborated further and summarized as institutions as game forms and the second approach is summarized as institution as an endogenous equilibrium outcome of a game.

This chapter proposes a theory which brings together insights from both the institutions as game form view and institutions as equilibria view. I present a model of institutional change based on repeated games with stochastic discount factors and argue that this approach is very intuitive and natural in capturing the micro-behavioral foundations of institutions. To put simply, I interpret discount factors as actors' confidence in the institutions and, using a very recent result [Barlo and Urgun, 2012], I show that temporary yet arbitrarily long sequences of non-cooperative behavior is inescapable in repeated social interactions. In contrast with the earlier literature, switches between self-enforcing and self-undermining institutions are possible in the same equilibrium specification. This transition allows us to account for periods of instabilities from which institutions eventually recover.

In section 3.2 I review the different approaches in institutionalist school briefly. Later I discuss some of the models in the theory of institutional change more closely. Section 3.3 develops the main model of the essay. I conclude by making some general remarks and speculating on directions for future research.

4.2 Discussion of the Relevant Literature

One can argue that the study of institutions is as old as the study of the socio-economic systems [Schotter, 1981]. During this long history, institutions have been studied using quite distinct methodologies. Within this tradition, the "new institutionalism" emerged as an umbrella term which itself consists of different bodies of thought developed in the last four decades [March and Olsen, 2006]. There are three such main schools of thought: historical institutionalism, sociological institutionalism, and rational choice institutionalism [Hall and Taylor, 1996]. Summarizing all of the literature on institutions and institutional change will exceed the limits of a single essay so I will focus on the literature which is more closely related to the arguments developed in the present article.

Historical institutionalism is based on macro-historical research and its emphasis is on how institutions emerge from and are embedded in concrete temporal processes [Thelen, 1999]. The basic and simple idea is that the policy choices made when an institution is being formed, or when a policy is initiated, will have a continuing and largely determinate influence over the policy far into the future [Peters, 2005]. One description of this idea is "path dependency": when a government program or organization embarks upon a path there is an inertial tendency for those initial policy choices to persist [Krasner, 1984]. This does not mean that this path cannot be altered. However it requires a good deal of political and social pressure to produce that change. This logic is formalized in the "punctuated equilibrium" concept. This is rather a metaphor borrowed from Darwinian theory of evolution akin to mutation in the genes of living organisms. This approach however does not help much in understanding institutional change. What causes these "mutations"? Where exactly is that "punctuation point"? It might be a viable explanation ex post but ex ante it seems tautological. Hence explaining institutional change remains a puzzle for historical institutionalism.

The sociological work on institutions can be traced back to Max Weber and Emile Durkheim. One problem with this literature is that the distinction between institutions ans organizations are not clear. In fact the methodology of sociological institutionalism can be seen as a institutional perspective on organizations [Mohr, 1982].

It is concerned with the process of creating values and cognitive frames within an organization than it is with the end state —the differences among organizations that can predict the behavior of those institutions and individuals within them [Peters, 2005]. Sociologist has a functionalist take on institutional change: institutions will find means of adapting to changes in the environment. One alternative view is that institutions will shape their environments to meet their own needs, rather than passively responding to those environments [Pfeffer and Salancik, 1978]. This view may be especially valid for political institutions which may have the power to manipulate the socio-political environment in ways that suit them. These being said, an explicit theory of institutional change does not seem to exist within the sociological institutionalism literature [Finnemore, 1996].

Among these three schools of institutional analysis, the school of rational choice institutionalism has had a growing dominance since it was first introduced more than three decades ago [Schotter, 1981]. It has become an engine of social scientific research, producing theoretical micro-foundations and deductively derived theorems which yield testable hypotheses [Shepsle, 2006]. It is sometimes identified as a more general social scientific methodology [Diermeier and Krehbiel, 2003]. Although it is usually treated as a unified body of methodologies [Green and Shapiro, 1994], there are different varieties of rational choice institutionalism. In one of the pioneering works in the field Calvert, 1995 distinguishes between three different views within the rational choice approach: i) institutions as features of individual preferences, ii) institutions as "rules of the game" and iii) institutions as equilibria of behavior in an underlying game. The first approach embeds the arguments for cooperation, observance of social norms, altruism and other relevant behavioral characteristics in the utility function of the actor [Margolis, 1984]. The "rules of the game" approach on the other hand, treats institutions as constraints on human behavior. [North, 1981, North, 1981, Shepsle and Weingast, 1987]. The third approach conceptualizes institutional arrangements as game theoretic equilibria for which no relevant actor has an incentive to deviate [Bates et al., 1998].

The usual theoretical toolkit employed in the institutions as equilibria approach is the theory of repeated games, in particular, repeated Prisoner's Dilemma game. Prisoner's Dilemma game is a simple but strikingly powerful illustration of why rational actors might end up at the inefficient equilibria while searching for what is best for them. One "solution" to this pessimistic result comes from the theory of repeated games. It can be shown that if rational actors play the Prisoner's Dilemma repeatedly, and if the actors value the future payoffs sufficiently enough, it is very well possible that they end up at the cooperative/efficient outcome [Axelrod, 1981]. It this cooperative outcome which is defined as "the institution" in the institutions as equilibria approach [Greif and Laitin, 2004].

One important conceptual remark which is crucial for our purposes here is that, although it has been called in that way right from the beginning [Schotter, 1981], what is referred by institutions as equilibria approach is really "institutions as the outcomes of equilibria" [Shepsle, 2006]. To see this more clearly one should note the difference between strategies and outcomes in game theory. Strategies in repeated game theory should define "complete and contingent plans" [Osborne, 2004]. This means that strategies are objects which specify which actions to choose at every stage of the game. And equilibria are particular strategy profiles, meaning a vector which is composed of a single strategy for every player, from which no player has an incentive to deviate. On the other hand, outcomes are the payoff profiles which are determined by those strategies. The focus of the institutions as equilibria approach is not on the strategies themselves, the focus is on the existence of strategies which will support the cooperative equilibrium.

Aoki, 2007 focuses on exactly this conceptual difference and distinguishes between two views within the rational choice paradigm: the exogenous view and the endogenous view. Based on this distinction he proposes the following definition for an institution:

> An institution is self-sustaining, salient patterns of social interactions, as represented by meaningful rules that every agent knows and are incorporated as agents shared beliefs about how the game is played and to be played.

The exogenous view treats institutions as pre-determined rules outside the domain of interactions where as the endogenous view treats those rules as something spontaneously and endogenously determined.[1] The first approach is elaborated further

[1]One can note that this distinction is similar to distinction between institutions as rules of the game and institutions as equilibria of the game categories provided in Calvert, 1995

and summarized as institution as a game form and the second approach is labeled as institution as an endogenous equilibrium outcome of a game.

In this study I will be treating institutions as game forms as in Aoki's categorization but I will be using the repeated Prisoner's Dilemma as the technical toolkit. I will now comparatively discuss the theories of institutional change in the literature.

4.2.1 Institutions: Rules or Equilibria ?

Before moving on, let me summarize the main concepts of game theory which I will be referring to throughout the discussion.

Any strategic game can be characterized by four pieces of information: i) a set of players, ii) sets of strategies available to the players, iii) an outcome function which assigns an outcome for every strategy profile and iv) player's preferences over these outcomes. Usually iii and iv are merged together under the name of a payoff function, a function which assigns payoffs to each player for every possible strategy profile and players prefer the highest payoff.

Equilibrium concepts are developed to predict the possible strategy profile outcomes-hence payoff distributions—of a given game. Nash equilibrium for instance assumes that all players will choose the strategy which is a best response to what other players choose. If this is an extensive form game, then subgame perfect Nash equilibrium concept requires that each player chooses a strategy which is a best response to every other player at each point in the game where the player gets to move. One important point to note is that if there is any uncertainty in the game, besides actions, any equilibrium specification has to include the beliefs of the players about that uncertainty under which such actions will be chosen by the players. Even in the perfect information case the reason why we expect Nash equilibrium to occur is that we assume that each player believes that other players are going to play their best responses for sure. In other words all game theoretic equilibrium concepts specify a belief structure, whether it is explicit or implicit.

One recent rational choice theoretic approach to institutional change has been developed in Greif and Laitin, 2004. This 'endogenous theory of institutional change' has been elaborated further and supported by historical narrative in Greif, 2006. Since it is the theoretically closest work, I will discuss the framework developed in the latter in more detail in this section. Throughout the discussion I will also compare and contrast it to the ideas of North on institutional change [North, 1981, North, 1990, North, 1993].

Similar to Calvert, 1995's categorization mentioned above, Greif makes a distinction between three approaches to institutional analysis: institutions as rules, institutions as equilibria and as shared beliefs motivating equilibrium play. He places North's approach into the first category.

> In the game-theoretic approach, institutions are considered as either equilibria..., the shared beliefs motivating equilibrium play ..., or the rules of the game [Greif, 2006].

His approach is an extension of the institutions as equilibria approach:

> Institutions are not game-theoretic equilibria, games are not the basic unit of institutional analysis, and game theory does not provide us with a theory of institutions. Indeed, the key to advancing institutional analysis by using game theory is precisely to recognize the difference between game-theoretic equilibrium analysis and institutional analysis.

His extension relies on how he extends the definition of and institution. He views institutions as *"interrelated systems as rules, beliefs, norms, and organizations that together generate a regularity of (social) behavior"*. To stress the difference he makes his approach and North's approach, for Greif, what causes the regularities of human behavior is not a monolithic entity such as a rule but a system of interrelated elements. One of the many examples he gives for this institutions as systems model is the following:

Rule: Rules of the road.

Organizations: Departments of motor vehicles and law enforcement officials.

Beliefs and Internalized Norms: Beliefs that other drivers and law enforcement officials will behave in a particular way.

Implied Regularity of Behavior: Driving according to the rules.

When you see the green light you will not stop. The reason you do not stop is your belief about how others will behave: you believe that the ones who see the red light *will* stop and if not they will be punished. Why will the others stop at the red light ? Because they believe that the ones who see the green light *will not.* Moreover there is a chance that a law enforcement official will catch you doing that and you will be punished. Hence in this game of traffic lights, the equilibrium is stopping at the red light and not stooping at the green light and the equilibrium outcome is driving according to the rules. This equilibrium is enforced by the beliefs of the players.

One important point to note in this model is the role that the organizations play. What they do is to constrain the set of beliefs and actions of individuals. This set of restricted action space can be seen as the set of rules, and the set of restricted beliefs as the beliefs and internalized norms of Greif's setup. Hence if we compare this setup with a strategic game form summarized above, the organization is the only part which is not explicitly stated in a game form. However since every equilibrium concept has to specify a consistent set of beliefs which support the equilibrium payoff profile, one can think of organizations as they are embedded in the black box of belief formation process.[2]

So if we employ this approach, any change in any of the first three components of this system, namely rules, organizations or beliefs, will be an institutional change. For instance, even if rules are the same, if you live in a corrupted country and if you believe that law enforcement officials will not punish the ones who do not stop at red lights then that means you are in a different institution. Accordingly, your incentive to stop at the red lights will decrease and you will at least slow down at the green light rather than proceeding without thinking because you know that others might not stop at the red light. Hence a different institutional system which results in a different equilibrium.

[2] Indeed there is a vast literature in game theory on belief formation and belief refinements.

However if you live in a country where rules are such that you can pass at the green light and pass slowly at the red light, this will be another different institutional system and the equilibrium *outcome* will also be different. But note that the implied regularity of the behavior, namely driving according to the rules, will be the same. In other words, in Greif's setup implied regularity of behavior and the implied equilibrium outcome are two different things.

Although this approach offers a wider and unified perspective it also brings about problems in terms of theoretical analysis. According to this model, in a social context there is almost nothing but institutions. This implies that there is no room for exogenous change and all changes have to be endogenous. The problem is similar to the identification problem in statistics. It is not possible to include everything in your model because it is not possible to make inferences out of models with zero degrees of freedom. The relevant problem for the purposes of this chapter appears in modeling institutional change or stability. As discussed above, we cannot be referring to a single entity by an institution. If we want to model a change using the broad definition above, we need to incorporate all possible changes that can occur in that institutional system. However it is not clear how this broad definition of an institution is used in the way Greif models institutional change and reinforcement. His formal model of reinforcement is based on a commonly used repeated interaction setting where rules or organizations are not explicitly incorporated. As we will see in more detail in the next section, the survival of an institution—hence a no change situation—is represented by the equilibrium in which all players cooperate and reinforcement mechanism is modeled as positive shocks to payoffs. To sum up, I argue that Greifs model of institutional change does not use his own extended definition of an institution but the common limited definition which treats institutions as the equilibrium outcome of an underlying repeated game.

Now let's turn to North and try to understand what he means by "institutions are the rules of the game". If we are talking about the rules of strategic game, which component of the game form do we mean by this ? Do institutions stand for the set of available strategies ? Or do they determine the payoff function ? Can institutions sometimes determine the set of players ?

Think of the criminal law as an institution. Supposedly, it is rule of the crime game. For instance, say, if one convicts a murder, the law judges for execution. So being executed is a "payoff" for the strategy you have chosen. Hence can we say in this case the institution shows itself as the payoff function ? But where are the other players ? Is this really a strategic game ?

When we look at North's theory closer, we see that what he means by "the rules of the game" is basically an agreement or a contract among players rather than some notion related to a strategic game.

> The process of institutional change can be described as follows. A change in relative prices leads one or both parties to an exchange, whether it is political or economic, to perceive that either or both could do better with *an altered agreement or contract.* An attempt will be made to renegotiate the contract.

Hence the strategic part of North's theory starts with the bargaining process about choosing the new agreement or contract, i.e, a new institution. The context for which an institution serves as the rules of the "game" —"an exchange whether it is political or economic"—and the actual strategic game in which these are rules are determined are treated as two distinct parts of the social environment. When North uses the phrase "institutions are rules of the game", he is not necessarily referring to a strategic game.[3] The institutional change takes place only after at least one party realizes that he can do better with a new contract. After that the new institution emerges as the *equilibrium* of the "renegotiation game". In other words, the new institution which emerges as an equilibrium of a bargaining game serves as the rules of the game for the corresponding human interaction setting.

In summary, even if Greif and North start with different definitions of an institution, if we look at the underlying mechanisms more carefully, there does not seem to be a structural difference in the way they model the change and persistence. However one difference is that in North's setup, the game in which new institutions are determined and the context in which institutions constrain behavior are distinct. In Greif's setup on the other hand, there exists a single game in which institutions persist as a self-enforcing equilibrium—or change as a self-undermining equilibrium—and it is this equilibrium which also characterizes the nature of the interaction under investigation. This is the point which leads to the concept of an "endogenous institution" and "endogenous institutional " change.

[3]It might be the case that the first "game" is actually a strategic game.

4.2.2 Change: Endogenous or exogenous ?

For North,

> The agent of change is the individual entrepreneur responding to the incentives embodied in the institutional framework. The sources of change are changing relative prices or preferences [North, 1990].

In other words, the engine of change in North's framework is the purposive maximizing behavior of economic organizations. Organizations with sufficient bargaining strength might find it profitable to renegotiate the existing "contracts" to change them a way that favors them.

Since North is mainly interested in the effects of institution on economic life an emphasize on the relative prices is understandable. However he argues that relative price changes have broader effects on all institutions:

> To the non-economist (and perhaps for some economists as well), putting such weight on changing relative prices may be hard to understand. But relative price changes alter the incentives of individuals in human interaction, and the only other source of such change is a change in tastes [North, 1990].

Thus if relative price changes are the sources of change, then any such change that results from the maximizing efforts of entrepreneurs can be regarded as being endogenous to the institutional change. Consider the R&D decisions of an industry. If the technological developments lead to a new production system which does not need as much as labor the old one, i.e relative prices of labor decrease, than that industry might have an incentive to renegotiate with the labor union with a greater bargaining power. A war on the other hand, will be an exogenous factor which will increase the relative price of the labor and the effects will be opposite in this case.

So North's theory allows for endogenous changes as well as exogenous ones. I want to reemphasize a point that I talked about in the previous section: the interaction/exchange part of the analytical framework in which institutions serve as the "rules of the game" and the bargaining game in which institutions are determined are distinct in his framework. Hence this endogeneity is structurally different than how Greif approaches endogeneity, i.e, a single game where institutions are both equilibrium outcome and the constraints on behavior.

On the other hand, if institutions are equilibria of a game, as in Greif's case, in which both interactions take place and the the survival of the institutions are determined, then how do institutions change? The game-theoretic definition of an equilibrium is by definition implies a "no-change" situation. Hence endogenous institutional change seems contradictory. Greif introduces a new paradigm of endogenous institutional change around two concepts: quasi-parameters and institutional reinforcement. Parameters are exogenous to the model and variables are endogenous. Quasi-parameters are a third type; they are parameters in the short-run and variables in the long run. So regarding to self-enforceability, they are parameters, regarding to change in the long-run, they are variables. One can think of quasi-parameters as pots in which marginal changes accumulate. A certain level of accumulation is needed for it to be effective in changing the game form at hand. Maybe the most important point about the distinction among parameters, variables and quasi parameters is that it is observational, it is not possible to theoretically argue that something is one of the three. I will now discuss the formal model Greif develops for institutional reinforcement.

The technical toolkit Greif uses to formalize his theory is the theory of repeated games. Take an infinitely repeated prisoner's dilemma game with the following stage game payoffs

		P2	
		C	C
P1	C	b_t,b_t	$-k$,b_t+e
	D	b_t+e,$-k$	0,0

where b_0, k, e > 0, and players share a common discount factor $\delta \in (0,1)$. The model has four parameters: δ, b_0, k, and e. b_t, is a quasi-parameter, since it can be affected by the institution in place. The institution we are interested in is the one generating cooperation, that is, stage-game play of (c,c).

Cooperation has a positive (negative, neutral) reinforcement if play of (c,c) in period t implies that $b_{t+1} - b_t > (<,=)0$. Standard repeated prisoner's dilemma models take cooperation to have neutral reinforcement. To simplify the analysis, we assume that the change in cooperation payoffs under any reinforcement mechanism is fixed over time, i.e, for all t, $b_{t+1} - b_t = \epsilon$ with $\epsilon > (<,=)0$ under positive (negative, neutral) reinforcement.

In this model it can be shown that the cooperation institution is self-enforcing over a larger range of discount factors under positive reinforcement than under neutral reinforcement. Moreover, under negative reinforcement, cooperation is not a self-enforcing institution. The first proof is simple, just find the subgame perfect Nash equilibrium conditions for different each reinforcement mechanisms and compare. The second proof is even more straightforward, apply backward induction given that payoffs from mutual cooperation decrease by ϵ every period if players have cooperated in previous periods.

What does Greif's theory tell us? Under positive reinforcement it is relatively easier[4] to have the cooperative outcome and this is the case of institutional persistence. Under negative reinforcement, on the other hand, the institution is not self-enforcing hence will not survive. The collapse of this cooperative outcome, hence the collapse of the institution is what Greif interprets as an institutional change.

An institution is however, either self-enforcing, self-undermining or neutral. So there is no transition from a self-enforcing institution to self-undermining or vice versa. This means that, given an institution, it is either going to live forever or collapse for sure. In fact, it is this transition which we should be interested in if we are interested in institutional change. Without doing it, it is as if we are explaining stability with stability and change with change.

[4]In comparison to the neutral reinforcement, i.e., the standard repeated Prisoner's Dilemma

4.2.3 Modeling transitions: Acemoglu and Robinson

Acemoglu and Robinson's work [Acemoglu and Robinson, 2005] emerged as a significant contribution to the discussion of the role of institutions on economic and political development. Although it is not explicitly about institutions or institutional change, I believe their implicit treatment of institutional change is worth mentioning as a side note in contrast to other models of change.

Their basic methodology is to analyze the conditions (such as distributional characteristics, class structure or economic structure of a society) under which transitions between different states of the world such as from non-democracy to democracy, a revolution in a non-democracy or a coup in democracy occur. What makes their approach structurally different than North or Greif is that the equilibria of their models are about whether a change will occur or not.

They emphasize the commitment value of institutions similar to North and Weingast, 1989. In their theory, groups (rich, poor or middle class) that have political power today can introduce political institutions that favor them. The political institutions of today in turn, regulate the allocation of political power in the future. In other words, institutions make it possible to lock in groups political power. However under certain conditions the ruling group "loses the game" and hence the institutions which favor them cannot persist and new political institutions form which favor the winner group. Their dynamic game-theoretic models provide parametric conditions under which different changes or transition will take place.

One basic notion of their theory is that institutions emerge as the equilibrium of a game where the elites choose to give citizens political power under the threat of social disorder when the costs of repression are too high. For instance, one of their propositions gives a critical level of the fraction of the income of the economy destroyed during revolution, under which there will be a revolution. Another proposition gives us the critical level of the fraction of capital and land lost during a coup under which there will be a coup. In another proposition, they give different such levels for before and after financial integration.

One might think that this approach would just give us discrete intervals for parameters for which change will occur, i.e, a sense of a punctuated equilibrium. However since all these critical levels are stated as continuous functions of other model parameters, such as the tax rate, capital level, factor prices or the production function specification, it allows us to do a rich set of comparative statics. Hence these models give us all sorts of predictions about the direction and the degree of the effect if model parameters on the change if interest (democratization, coup, revolution, financial integration etc.). For instance, if increasing the marginal product of capital decreases the critical level of the fraction of capital and land lost during a coup under which there will be a coup, that means chances of having a coup would be more difficult in countries with higher marginal product of capital.

To conclude, the set of tools developed by Acemoglu and Robinson, 2005 offers an alternative way of analyzing change: setting up a carefully parametrized dynamic strategic—usually a variant of the bargaining—game on whether a change will occur and use comparative statics analysis to make inference about the nature of the phenomenon. Unlike North or Greif's framework, the stability of an institution is not characterized by identifying a reinforced equilibrium, rather it is characterized by conditions under which the transition game does not result in a "change equilibrium". These are quite distinct methodologies.

4.3 The Model: Uncertainty and Change

The idea that arbitrary changes in expectations might influence the economy even if they are not related to fundamental variables has been a widespread observation. One can go back as early as Pigou:

> The varying expectations of business men... and nothing else, constitute the immediate cause and direct causes or antecedents of industrial fluctuations [Pigou, 1967].

Institutions are not very different in this respect. There are many "sunspots" which have contagious effects on the stability of institutions which come from outside of their domains of interactions. Extrinsic uncertainty is one such sunspot with self-fulfilling features [Cass and Shell, 1983]. The closest example is the recent financial crisis. The loss of average confidence due to some failure in the U.S mortgage system had world-wide catastrophic effects.

In this section I develop a model where uncertainty regarding to the future of an institution is the mechanism which undermines the institution. I treat institutions as game forms which define social interactions [Aoki, 2007,Shepsle, 2006]. An institution persists as long as the same game is played, and the same game is played as long as actors cooperate in that game. In this sense, similar to Greif, 2006, I interpret the occurrence "bad equilibrium outcome" as the reason for the collapse of the existing institution, hence an institutional change. However, in contrast to Greif's model, not all such outcomes need to lead to a collapse. There might be a threshold of repetitions of bad outcomes over which the current institution fails and a transition from bad outcome state to good outcome state is possible. Hence, the model captures crisis situations or "instability in institutions" which are recovered.

The mechanism I offer rests on the risk perceptions of the actors. Since the introduction of the theory of rational expectations [Muth, 1961] we know very well that uncertainty about the future of the institution is a crucial determinant of the future of that institution.

The only way to capture the uncertainty component in a theoretical model is to introduce a stochastic process to generate a variable of the model. I argue that the discount factor is a natural candidate for this. One can interpret discount factors in repeated games in two different but computationally equivalent ways. The first is the common interpretation where players discount future payoffs exponentially by some factor δ due to "time value of money". In this case, the payoffs to the players will be a finite or infinite sum of discounted stage game payoff. An alternative interpretation of the same formulation would be the following: at each stage, the players believe that they will be playing the same game next period with some probability, δ. So at each repetition of the game, players believe that there is a chance that the game will stop. The expected return of this type of a formulation will be an infinite or infinite sum which is equal to the discounted payoff of the previous formulation.

With the latter interpretation in mind, I interpret δ as a measure of the trust the players have in the institution, or equivalently a measure of lack of uncertainty regarding to the survival of the institution. So in interpretation of the survival of an institution is the continuation of the game form itself. The failure of an institution on the other hand, occurs when the bad outcome is realized for a certain period of time during the game. If the trust in the institution —the belief hold by the players that the same game form will repeat in the next period—is stochastic what can we say about the equilibrium outcome, hence about the future of that institution? We need a theory of repeated games with stochastic discounting in order to deal with questions like this.

Surprisingly the work on repeated games with stochastic discounting is quite limited. To my knowledge, Baye and Jansen, 1996 is the first work which deals with repeated games with stochastic discount factors. What they do is basically to prove a folk theorem using stochastic discount factors. Their setting however does not allow for history dependent stochastic discounting, something I need to have in order to support my arguments. Recently Barlo and Urgun, 2012 provided a more comprehensive framework to study repeated games with stochastic discount factors. I will follow the setup developed in that paper.

Let $G = (N, (A_i, u_i)_{i \in N})$ be a normal form game where N is the set of players and A_i is the set of actions available to player i. Also let $u_i : A \to \mathbb{R}$ denote player i's payoff function where $A = \pi_{i \in N} A_i$ and $A_{-i} = \pi_{j \neq} A_i$. In every period $t \in \mathbb{N}_0$, a random variable, $\mathbf{d}_t$, is determined and this sequence forms a stochastic process $\{\mathbf{d}_t\}_{t \in \mathbb{N}_0}$. The supergame is defined for a given stochastic process with the initial discount factor being $\hat{\delta} = r\mathbf{d}_0$ and $r \in (0, 1]$. Then a k-stage history is given by $h_k = ((a_o, d_1), ..., (a_{k-1}, d_k))$ where d_t is a realization of $\mathbf{d}_t$.

The players have complete information. This means, at every t each player observes all the previous action profiles and all the discount factor shocks including those realized in period t. To stress it once again, players observe the current period's discount factor before making a move. All in all, this is a perturbation of the standard repeated game setup.

In order to get an idea on what the strategies look like in a repeated game with stochastic discounting consider the following generic form of 2-player repeated Prisoner's Dilemma. Let $b, c > 0, b > 1, b < c + 2$.

P2

		C	D
P1	C	1,1	$-c$,b
	D	b,$-c$	0,0

Let $\pi_{(0)}$ be given by the repetitions of (C, C), $\pi_{(P)}$ the repetitions of (D, D), and $\pi_{(NE_q)}$ be also the repetitions of (D, D). Then the strategy profile in which both player play the following strategy is an equilibrium: If the shocks have all been such that the resulting oneperiod discount factors were greater or equal to δ_* (i) play (C, C) until some player unilaterally deviates from $\pi_{(0)}$ and (ii) play $\pi_{(P)}$ if there was a single player deviation, and (iii) continue playing $\pi^{(j)}$, $j = 0, P$ if there were either no deviations or multi-player deviations; and if there is a period in which the resulting oneperiod discount factors were strictly less than δ^* play $\pi_{(NE_q)}$, repetitions of (D, D), for the rest of the game.

Below is a list of assumptions needed for the result I am interested in:

Assumption 1 *The stage game has at least one pure strategy equilibrium.*

Assumption 2 *In every period, the players use the most up to date information regarding to the state of the world.*

Assumption 3 *The stochastic process $\{\boldsymbol{d}_t\}_{t\in\mathbb{N}_0}$ satisfies the following:*

1. *Markov property*

2. *Martingale property*

3. *The state space Ω of $\{\boldsymbol{d}_t\}_t$ is a subset of $(0, 1)$*

4. *Given Ω, the set of ergodic states, Ω^E, is dense in Ω*

5. *For any $\epsilon > 0$, there exists $\tau \geq t$ with $Pr[\boldsymbol{d}_\tau < \epsilon|\mathbb{F}_t] > 0$ where $\mathbb{F}_t$ is the information available at time t.*

6. *For any given state $\omega \in \Omega \subseteq (0, 1)$, the set of states $\omega' \in \Omega$ that are reachable from ω in a single period and satisfying $\omega < \omega'$, denoted by $R(\omega)$, is finite. Moreover, for any $\omega, \omega' \in \Omega$ with $\omega' \geq \omega$, sup $R(\omega') \geq$ sup $R(\omega)$*

7. $\boldsymbol{d}_0$ *is non-stochastic.*

Parts 1 and 2 of Assumption 3 simply implies that the expectations about the future are equal to the current value and do not depend on anything else. Part 5 is an important assumption and implies that there are states arbitrarily close to 0 and such states can be reached with positive probability yet arbitrarily small probability in the long-run. Part 6 is the standard bounded increments requirement. [5]

Theorem 1 (Barlo and Urgun (2012)) *If assumptions 1,2 and 3 hold, then, for every* $K \in \mathbb{N}$, *for every* $\delta \in (0,1)$, *for every discounting process* $\{d_t\}$ *with* $d_0 = \delta$, *for every subgame perfect strategy profile* f *of the repeated game with stochastic discounting; there exists* T *which is almost surely in* $K \in \mathbb{N}_{\not{K}}$, *and the probability of* $\pi^\tau(f)$ *being a Nash equilibrium action profile of the stage game conditional on the information available at* s, *equals 1, for all* $s = T, T+1, ...T+K$ *and for all* $\tau = s, s+1, ..., T+K$.

The theorem basically says that, under assumptions 1,2 and 3, finite but arbitrarily long consecutive repetitions of Nash profile of the stage game will almost surely happen in finite time no matter which subgame perfect equilibrium strategy is considered and how high the initial common discount factor is. Moreover the prescription of Nash behavior occurs whenever the current discount factor, which serves as the expectation for the rest of the game, is sufficiently small, hence confidence level is sufficiently low.

[5] Are these assumptions on the stochastic process too restrictive? One can note that the famous Polya's urn process and standard random walk satisfy all these requirements.

For the case of repeated Prisoner's Dilemma, this will mean that the non-cooperative outcome [6] will certainly be observed no matter which subgame perfect equilibrium is considered when the realized value of the discount factor is sufficiently small. Moreover, this repetition of non-cooperative outcome will not go on forever. Hence transitions between cooperation and non-cooperation phases will take place if the game continues to be played[7] So this theorem is both good news and bad news for the persistence of cooperation. The bad news is that there will certainly be periods of non-cooperation, and the good news is that they will not last forever.

To stress it once again, in my argument, the persistence of the cooperation phase is a sufficient but not a necessary condition for the persistence of the institution. Since institution is the game form, it might persist even if players do not cooperate. However, the elongation of this non-cooperation outcome will undermine the institution, and depending on the durability of the institution it might collapse leading way to a new institution or recover and go back to cooperation outcome phase, which is guaranteed to occur in finite time.

4.4 Concluding Remarks

On 15 July 1979, President Carter, in his address to the Americans about the energy crisis, was drawing attention to "a subject even more serious than energy or

[6]Payoff profile of $(0, 0)$ as a result of the strategy profile (D, D) in our example above.

[7]Note that this transition is not a change in the equilibrium, it is indeed an equilibrium which is induced by more sophisticated strategies than the usual punishing strategies such as tit-for-tat or grim trigger. The main reason is that the histories not only include the action profiles but the realizations of the common stochastic discounting factor.

inflation...a fundamental threat to American democracy." That threat was a "crisis of confidence" [Lipset and Schneider, 1983]. It turns out that President Carter was correct to fear. Almost two decades later, when the insolvency of the investment bank Lehmann Brothers triggered the worst financial and economic crisis since the 1930s, scholars who were familiar with the notion of systemic trust were making predictions that the biggest damage caused by the crisis is most likely that to citizens' systemic trust.

The crucial characteristic of systemic risk is that the instability caused or exacerbated by idiosyncratic condition can lead to a cascading failure through interlinkages and interdependencies [De Bandt and Hartmann, 2000]. The loss of average confidence in one part of the system can bring down the whole system. The point of these arguments is that a random exogenous shock is one of the main ingredients of institutional stability in a world of interconnected institutions as we line in today.

In this chapter I offered a theory of institutional change based on the dynamics of actors' trust in the institution to capture these dynamics. I argued that institutions are game forms and the change in the game form occurs when the cooperation fails within that institution for sometime. I showed that under some assumptions on the stochastic process, the non-cooperation phase is inevitable in repeated Prisoner's Dilemma game with stochastic confidence levels in the institutions. I believe that this is a strong way of conceptualizing the external shocks on behavioral beliefs which trigger institutional change.

Developing models of socio-economic systems which involve underlying microbehavioral foundations is an important task for the social scientists. We need to pay attention to identify the conditions and processes which are likely to improve the efficiency of society by facilitating coordinated action. Designing institutions which are sturdy against uncertainties is becoming more important in a world where systemic risks are becoming more and more common. The model in this chapter showed that in a world of uncertainities, the cooperation between actors are going to fail sooner or later and it might never recover. What are then the characteristics of those institutions which seem to be sturdy against uncertainties? How can we design institutions which are resistant against the shocks of confidence losses? This is a direction which remains to be investigated furher both theoretically and empirically.

REFERENCES

[Acemoglu and Robinson, 2005] Acemoglu, D. and Robinson, J. A. (2005). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

[Adams and Merril III, 1999] Adams, J. and Merril III, S. (1999). Modeling party strategies and policy representation in multiparty elections: why are strategies so extreme? *American Journal of Political Science*, pages 765–791.

[Aldrich, 1983] Aldrich, J. (1983). A downsian spatial model with party activism. *The American Political Science Review*, pages 974–990.

[Alvarez and Nagler, 1995] Alvarez, R. and Nagler, J. (1995). Economics, issues and the perot candidacy: Voter choice in the 1992 presidential election. *American Journal of Political Science*, pages 714–744.

[Aoki, 2007] Aoki, M. (2007). Endogenizing institutions and institutional changes. *Journal of Institutional Economics*, 3(1):1–31.

[Ashworth and Bueno de Mesquita, 2009] Ashworth, S. and Bueno de Mesquita, E. (2009). Elections with platform and valence competition. *Games and Economic Behavior*, 67(1):191–216.

[Austen-Smith and Banks, 2000] Austen-Smith, D. and Banks, J. (2000). *Positive political theory I: collective preference*. Univ of Michigan Pr.

[Auyero, 1999] Auyero, J. (1999). From the client's point (s) of view: How poor people perceive and evaluate political clientelism. *Theory and Society*, 28(2):297–334.

[Auyero, 2001] Auyero, J. (2001). *Poor peopleís politics. Peronist survival networks and the legacy of Evita.* Duke University Press.

[Axelrod, 1981] Axelrod, R. (1981). *The evolution of cooperation.* Basic Books.

[Banks and Duggan, 2005] Banks, J. and Duggan, J. (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In Austen-Smith, D. and Duggan, J., editors, *Social choice and strategic decisions.* Springer.

[Banks and Duggan, 2006] Banks, J. and Duggan, J. (2006). A social choice lemma on voting over lotteries with applications to a class of dynamic games. *Social Choice and Welfare*, 26(2):285–304.

[Barlo and Urgun, 2012] Barlo, M. and Urgun, C. (2012). Stochastic discounting in repeated games: Awaiting the almost inevitable. *Unpublished Manuscript.*

[Bates et al., 1998] Bates, R., Greif, A., Levi, M., Rosenthal, J., and Weingast, B. (1998). *Analytical Narratives.* Princeton University Press Princeton.

[Baye and Jansen, 1996] Baye, M. and Jansen, D. (1996). Repeated games with stochastic discounting. *Economica*, pages 531–541.

[Bélanger and Meguid, 2008] Bélanger, É. and Meguid, B. (2008). Issue salience, issue ownership, and issue-based vote choice. *Electoral Studies*, 27(3):477–491.

[Benoit and Laver, 2006] Benoit, K. and Laver, M. (2006). *Party policy in modern democracies*. Taylor & Francis.

[Boix, 2003] Boix, C. (2003). *Democracy and redistribution*. Cambridge Univ Pr.

[Budge and Farlie, 1983] Budge, I. and Farlie, D. (1983). *Explaining and predicting elections: Issue effects and party strategies in twenty-three democracies*. G. Allen and Unwin.

[Callahan and McCargo, 1996] Callahan, W. and McCargo, D. (1996). Vote-buying in thailand's northeast: The july 1995 general election. *Asian Survey*, 36(4):376–392.

[Calvert, 1995] Calvert, R. (1995). Rational actors, equilibrium, and social institutions. In Sened, I. and Knight, J., editors, *Explaining Social Institutions*, pages 57–95. Ann Arbor: University of Michigan Press.

[Calvo and Murillo, 2004] Calvo, E. and Murillo, M. (2004). Who delivers? partisan clients in the argentine electoral market. *American Journal of Political Science*, 48(4):742–757.

[Carkoglu and Hinich, 2006] Carkoglu, A. and Hinich, M. (2006). A spatial analysis of turkish party preferences. *Electoral Studies*, 25(2):369–392.

[Cass and Shell, 1983] Cass, D. and Shell, K. (1983). Do sunspots matter? *The Journal of Political Economy*, pages 193–227.

[Chandra, 2007] Chandra, K. (2007). Counting heads: A theory of voter and elite behaviour in patronage-democracies. In Kitschelt, H. and Wilkinson, S., editors, *Patrons, Clients and Policies. Patterns of Democratic Accountability and Political Competition*, pages 84–109. Cambridge University Press.

[Clark, 2009] Clark, M. (2009). Valence and electoral outcomes in western europe, 1976-1998. *Electoral Studies*, 28(1):111–122.

[Cox and McCubbins, 1986] Cox, G. and McCubbins, M. (1986). Electoral politics as a redistributive game. *Journal of Politics*, 48(2):370–389.

[De Bandt and Hartmann, 2000] De Bandt, O. and Hartmann, P. (2000). Systemic risk: A survey. *ECB Working Paper No. 35.*

[Diermeier and Krehbiel, 2003] Diermeier, D. and Krehbiel, K. (2003). Institutionalism as a methodology. *Journal of Theoretical Politics*, 15(2):123–144.

[Dow and Endersby, 2004] Dow, J. and Endersby, J. (2004). Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies*, 23(1):107–122.

[Downs, 1957] Downs, A. (1957). An economic theory of political action in a democracy. *The Journal of Political Economy*, 65(2):135–150.

[Enelow and Hinich, 1982] Enelow, J. and Hinich, M. (1982). Nonspatial candidate characteristics and electoral competition. *The Journal of Politics*, 44(1):115–130.

[Epstein et al., 2006] Epstein, D., Bates, R., Goldstone, J., Kristensen, I., and O'Halloran, S. (2006). Democratic transitions. *American journal of political science*, 50(3):551–569.

[Finnemore, 1996] Finnemore, M. (1996). Norms, culture, and world politics: insights from sociology's institutionalism. *International organization*, 50(02):325–347.

[Fiorina, 1978] Fiorina, M. (1978). Economic retrospective voting in american national elections: A micro-analysis. *American Journal of Political Science*, pages 426–443.

[Green and Shapiro, 1994] Green, D. and Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. Cambridge Univ Press.

[Greenberg, 2008] Greenberg, E. (2008). *Introduction to Bayesian Econometrics*. Cambridge University Press: Cambridge, UK.

[Greif, 2006] Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.

[Greif and Laitin, 2004] Greif, A. and Laitin, D. (2004). A theory of endogenous institutional change. *American Political Science Review*, 98(04):633–652.

[Grose, 2005] Grose, C. (2005). Do legislators use 'pork'projects to deviate from constituents' interests? valence advantages and position-taking in congress. *Typescript. Vanderbilt University.*

[Grossman and Helpman, 1994] Grossman, G. and Helpman, E. (1994). Protection for sale. *American Economic Review*, 84(4).

[Grossman and Helpman, 1996] Grossman, G. and Helpman, E. (1996). Electoral competition and special interest politics. *The Review of Economic Studies*, 63(2):265–286.

[Grossman and Helpman, 2002] Grossman, G. and Helpman, E. (2002). *Special interest politics*. The MIT Press.

[Hall and Taylor, 1996] Hall, P. and Taylor, R. (1996). Political science and the three new institutionalisms. *Political studies*, 44(5):936–957.

[Hinich, 1977] Hinich, M. (1977). Equilibrium in spatial voting: The median voter result is an artifact. *Journal of Economic Theory*, 16(2):208–219.

[Hotelling, 1929] Hotelling, H. (1929). Stability in competition. *The economic journal*, 39(153):41–57.

[Kass and Raftery, 1995] Kass, R. and Raftery, A. (1995). Bayes factors. *Journal of the american statistical association*, pages 773–795.

[Kitschelt and Wilkinson, 2007a] Kitschelt, H. and Wilkinson, S. (2007a). Citizen-politician linkages: an introduction. In Kitschelt, H. and Wilkinson, S., editors, *Patrons, Clients and Policies. Patterns of Democratic Accountability and Political Competition*, pages 1–46. Cambridge University Press.

[Kitschelt and Wilkinson, 2007b] Kitschelt, H. and Wilkinson, S. (2007b). A research agenda for the study of citizen-politician linkages and democratic accountability. In Kitschelt, H. and Wilkinson, S., editors, *Patrons, Clients and Policies. Patterns of Democratic Accountability and Political Competition*, pages 322–343. Cambridge University Press.

[Krasner, 1984] Krasner, S. (1984). Approaches to the state: Alternative conceptions and historical dynamics. *Comparative politics*, 16(2):223–246.

[Kuziemko and Werker, 2006] Kuziemko, I. and Werker, E. (2006). How much is a seat on the security council worth? foreign aid and bribery at the united nations. *Journal of Political Economy*, 114(5):905–930.

[Laver and Hunt, 1992] Laver, M. and Hunt, W. (1992). *Policy and party competition*. Routledge and Kegan Paul.

[Lin et al., 1999] Lin, T., Enelow, J., and Dorussen, H. (1999). Equilibrium in multicandidate probabilistic spatial voting. *Public Choice*, 98(1):59–82.

[Lipset and Schneider, 1983] Lipset, S. and Schneider, W. (1983). The decline of confidence in american institutions. *Political Science Quarterly*, 98(3):379–402.

[March and Olsen, 2006] March, J. and Olsen, J. (2006). Elaborating the "new institutionalism". In R.A.W Thodes, S. A. B. and Rockman, B. A., editors, *Oxford Handbook of Political Institutions*, pages 3–20. Oxford University Press Oxford.

[Margolis, 1984] Margolis, H. (1984). *Selfishness, altruism, and rationality: A theory of social choice*. University of Chicago Press.

[McKelvey and Patty, 2006] McKelvey, R. and Patty, J. (2006). A theory of voting in large elections. *Games and Economic Behavior*, 57(1):155–180.

[Miller and Schofield, 2003] Miller, G. and Schofield, N. (2003). Activists and partisan realignment in the united states. *American Political Science Review*, 97(2):245–260.

[Mohr, 1982] Mohr, L. (1982). *Explaining organizational behavior*. Jossey-Bass San Francisco.

[Muth, 1961] Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica*, pages 315–335.

[Nichter, 2008] Nichter, S. (2008). Vote buying or turnout buying? machine politics and the secret ballot. *American Political Science Review*, 102(01):19–31.

[North, 1981] North, D. (1981). *Structure and change in economic history*. Norton New York.

[North, 1993] North, D. (1993). Institutional change: A framework of analysis. In Sjöstrand, S.-E., editor, *Institutional Change: Theory and Empirical Findings*, pages 35–46. M E Sharpe Inc.

[North and Weingast, 1989] North, D. and Weingast, B. (1989). Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century england. *Journal of economic history*, 49(4):803–32.

[North, 1990] North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

[Osborne, 2004] Osborne, M. (2004). *An Introduction to Game Theory*. Oxford University Press New York.

[Ostrom and Basurto, 2011] Ostrom, E. and Basurto, X. (2011). Crafting analytical tools to study institutional change. *Journal of Institutional Economics*, 7(3):317–343.

[Penn, 2009] Penn, E. (2009). A model of farsighted voting. *American Journal of Political Science*, 53(1):36–54.

[Peters, 2005] Peters, B. (2005). *Institutional theory in political science: the'new institutionalism'*. Continuum Intl Pub Group.

[Pfeffer and Salancik, 1978] Pfeffer, J. and Salancik, G. (1978). *The external control of organizations*. New York: Harper & Row.

[Pigou, 1967] Pigou, A. (1967). *Industrial fluctuations*. AM Kelley.

[Poole and Rosenthal, 1984] Poole, K. and Rosenthal, H. (1984). Us presidential elections 1968-80: A spatial analysis. *American Journal of Political Science*, pages 282–312.

[Przeworski, 2000] Przeworski, A. (2000). *Democracy and development: political institutions and well-being in the world, 1950-1990*, volume 3. Cambridge Univ Pr.

[Rokkan and Lipset, 1967] Rokkan, S. and Lipset, S. (1967). *Party systems and voter alignments: crossnational perspectives.* Free Press.

[Rossi et al., 2005] Rossi, P., Allenby, G., and McCulloch, R. (2005). *Bayesian statistics and marketing.* Wiley.

[Saari, 1997] Saari, D. (1997). The generic existence of a core for q-rules. *Economic Theory*, 9(2):219–260.

[Schofield, 2003] Schofield, N. (2003). Valence competition in the spatial stochastic model. *Journal of Theoretical Politics*, 15(4):371–383.

[Schofield, 2004] Schofield, N. (2004). Equilibrium in the spatial valence model of politics. *Journal of Theoretical Politics*, 16(4):447.

[Schofield, 2006] Schofield, N. (2006). Equilibria in the spatial stochastic model of voting with party activists. *Review of Economic Design*, 10(3):183–203.

[Schofield, 2007] Schofield, N. (2007). Political equilibria with electoral uncertainty. *Social Choice and Welfare*, 28(3):461–490.

[Schofield et al., 2011] Schofield, N., Claassen, C., Ozdemir, U., and Zakharov, A. (2011). Estimating the effects of activists in two-party and multi-party systems: comparing the united states and israel. *Social Choice and Welfare*, 36(3):483–518.

[Schofield et al., 2009] Schofield, N., Gallego, M., Ozdemir, U., and Zakharov, A. (2009). Competition for popular support: a valence model of elections in turkey. *Social Choice and Welfare*.

[Schofield and Miller, 2007] Schofield, N. and Miller, G. (2007). Elections and activist coalitions in the united states. *American Journal of Political Science*, 51(3):518–531.

[Schofield et al., 2003] Schofield, N., Miller, G., and Martin, A. (2003). Critical elections and political realignments in the usa: 1860–2000. *Political Studies*, 51(2):217–240.

[Schofield and Sened, 2006] Schofield, N. and Sened, I. (2006). *Multiparty democracy: elections and legislative politics*. Cambridge Univ Pr.

[Schofield and Zakharov, 2010] Schofield, N. and Zakharov, A. (2010). A stochastic model of the 2007 russian duma election. *Public Choice*, 142(1):177–194.

[Schotter, 1981] Schotter, A. (1981). The economic theory of social institutions. *Cambridge Books*.

[Scott, 1969] Scott, J. (1969). Corruption, machine politics, and political change. *The American Political Science Review*, 63(4):1142–1158.

[Scott, 1972] Scott, J. (1972). Patron-client politics and political change in southeast asia. *The American Political Science Review*, 66(1):91–113.

[Shepsle, 2006] Shepsle, K. (2006). Rational choice institutionalism. In Thodes, R., Binder, S. A., and Rockman, B. A., editors, *Oxford Handbook of Political Institutions*, pages 23–39. Oxford University Press Oxford.

[Shepsle and Weingast, 1987] Shepsle, K. and Weingast, B. (1987). The institutional foundations of committee power. *The American Political Science Review*, pages 85–104.

[Stokes, 1963] Stokes, D. (1963). Spatial models of party competition. *The American Political Science Review*, 57(2):368–377.

[Stokes, 1992] Stokes, D. (1992). Valence politics. In Kavanagh, D., editor, *Electoral Politics*, pages 141–64. Clarendon Press.

[Stokes, 2005] Stokes, S. (2005). Perverse accountability: A formal model of machine politics with evidence from argentina. *American Political Science Review*, 99(03):315–325.

[Thelen, 1999] Thelen, K. (1999). Historical institutionalism in comparative politics. *Annual review of political science*, 2(1):369–404.

[Train, 2003] Train, K. (2003). *Discrete choice methods with simulation*. Cambridge Univ Pr.

[Weingrod, 1968] Weingrod, A. (1968). Patrons, patronage, and political parties. *Comparative studies in Society and History*, 10(4):377–400.

[Westen, 2007] Westen, D. (2007). *The political brain: The role of emotion in deciding the fate of the nation*. Public Affairs.

[Whyte, 1965] Whyte, J. (1965). Landlord influence at elections in ireland, 1760-1885. *The English Historical Review*, 80(317):740–760.