Party Formation in Parliamentary Democracies

Selcen Çakır*

May, 2018

Abstract

The political power of elected representatives is determined by politicians' membership negotiations with parties. In parliamentary systems, party control over government functions generates club goods, which increases the value of party membership. Moreover, in party-centered parliamentary systems, getting influential positions requires party leader's approval, which gives the leader monopsonistic recruiting power. As a result, politicians may relegate their political power to leaders when membership is more rewarding than acting independently. I develop an equilibrium model of party formation in a parliamentary democracy that incorporates parties' provision of club goods, rent sharing between politicians and party leaders, as well as politicians' outside options. Politicians' rankings of parties critically depend on the size of the party as well as on their own political assets. Through their control of government functions, bigger parties can provide greater club goods but tax politicians' rents more upon entry. Because of this, politicians with more assets tend to prefer smaller parties. I structurally estimate my model for Turkey with a unique dataset of 33 parties, 2,000 politicians who gained seats in parliament, and 35,000 politicians who were on party ballot lists between 1995 and 2014. My model matches the high level of party switching (28.5%) that is characteristic of many party-centered systems. I find that Turkish parties accumulate club goods more easily than they produce rents, which leads to ever stronger party control. I also find that politicians with good labor market options are not productive in the political arena. In a counterfactual analysis, I find that members of smaller (bigger) parties are more powerful in party-centered (candidate-centered) systems.

JEL Classifications: D71, D72, J4, J410

Keywords: Clubs, Rent-Seeking, Coalitions, Particular Labor Markets, Matching Models

---

*sc3ch@virginia.edu. This paper is a concise version of my dissertation at the University of Virginia. I am indebted to my adviser Steven Stern for his guidance, encouragement, and patience. I am also grateful to my advisers Leora Friedberg and John McLaren for their guidance and encouragement. I would also like to thank Simon Anderson, Gaurab Aryal, Marina Azzimonti, Hugo Benitez-Silva, Zachary Bethune, Meta Brown, Sandra Brusco, Levent Çelik, Karim Chalak, Federico Cilberto, Jonathan Colmer, Juan Conesa, Kerem Coşar, Peter Debaere, Maxim Engers, İlhan Güner, Guillermo Hausmann Guil, William Johnson, Carol Mershon, Gabriel Mihalache, Toshihiko Mukoyama, Edgar Olsen, Warren Sanderson, Sheetal Sekhri, Aslı Şenkal, Yair Tauman, David Wiczer, Eric Young, and many other faculty members and students at the University of Virginia and seminar participants at Boğaziçi University and Stony Brook University for their helpful comments. I gratefully acknowledge the Bankard Fund for Political Economy for financial support. All errors are mine.

1 Introduction

Political parties and party leaders are at the center of political and electoral systems. How they operate matters for how power is distributed and whether a parliamentary system functions democratically or not. Across countries, there is considerable variation in how parties operate, and how much of a say parties, party leaders, and politicians have. In the closed-list electoral systems of Argentina, Israel, Italy, Spain and Turkey, for example, parties and party leaders have a lot more clout than, say, in the more candidate-centered systems of the United States, Canada, and other countries, and politicians will typically switch party affiliation a lot more frequently in such systems.[1] In this paper, I build a theoretical model of party formation and the effect of electoral institutions on the power distributions in parties while focusing on the interaction of party leaders and politicians in a labor search framework. I then estimate my model with a dataset I constructed for Turkey of 35,000 listed politicians, 2,000 politicians who gained access to parliament, and 33 parties between 1995 and 2014.

Party leaders in many proportional representation systems are apparently quite powerful. The leaders select candidates and order them in priority for winning seats. Hence, gaining an influential position in a party requires the party leader's approval. Moreover, party control over government functions generates valuable club goods. Club goods, such as the party's facilities or the security gained by affiliating with a strong team, are shared among party members. A politician may relegate the use of his electoral power to party authorities when the benefits of affiliating with the party exceed the rewards from acting independently.[2] Accordingly, to the extent that the electoral institutions allow party leaders to use members' political power in exchange for supplying other membership benefits, politicians' ability to act independently in a party will be limited. In political systems that yield extensive party control, the legislative activities, the appointment of the bureaucrats, and the allocation of government spending are largely determined by a few strong leaders instead of by the members of the parliament as a whole, which can severely damage democratic functioning.

Political arenas in all systems can be considered as markets for rent production, which embody the ability to influence government institutions. Political rents, such as winning a seat in parliament or a party primary election, decisive power over the use of government budget, or employing supporters in public institutions, are private and exclusive. The extent to which a party leader can use the party's club goods to exploit members' rents is determined by electoral institutions. In candidate-centered systems where voters can show a preference for a candidate, each politician produces rents with more independence. This

[1] See Nikolenyi (2016) for a comparison of party switching in Canada and Israel.

[2] Throughout the paper, I refer to a politician and a leader with male and female pronouns, respectively.

independence in rent production gives politicians more bargaining power during membership negotiations with a party leader. In a party-centered system such as a closed-list proportional representation system, on the other hand, voters can vote only for the party as a whole, which gives a party leader monopsonistic power while recruiting politicians. Accordingly, in a party-centered system, parties function as entities that produce political rents and politicians provide their parties the political assets for rent production. The party competition for rents then brings about a competition for productive politicians, just like in a labor market. Party-switching by politicians resembles workers' transition between firms, and the parties' competition over their services highlights the importance of outside options. This paper constructs an equilibrium model of party formation which incorporates these features of team production. The model can be applied to labor markets, in which a worker trades off compensating differentials and wages he earns in a firm engaged in team production. Similar to the compensating-differentials literature, in a Nash equilibrium, a small party pays higher rents to a politician to compensate for his disutility in lacking club goods (Rosen 1986, Sorkin 2015).

This paper models the political arena as a labor market in which a party is represented by a leader, who has the exogenous ability to lead a party of a certain capacity that produces political benefits, i.e., rents and club goods. A leader aggregates the assets of heterogeneous politicians to produce political benefits and seeks new members through a random matching process when there is a vacancy. Accordingly, I model the unstable party structures that are common in party-centered systems as resulting from frictions that prevent ideal matches of politicians and leaders.[3] Once brought together in pairwise matches, a leader and a politician bargain over a share of the politician's rent production in the party. The provision of club goods allows a leader to attract members who accept joining the party by receiving less rents than what they produce. In equilibrium, a leader makes acceptable offers to only the most profitable politician-types who ask for the smallest rents. As a result, depending on the primitives that describe the heterogeneous abilities of the leaders and the amounts of politicians' assets, party governance may function either as a democracy or a dictatorship. The power distribution in parties may also translate into the power distribution in parliament. For example, if one party's leader can accumulate more club goods, she can fill her party with members who join without receiving any rents. In this extreme case, only one

[3]Other possible explanations for the unstable party structures are changes in politicians' tastes and parties' vote shares over time. These explanations can be studied in a dynamic discrete choice framework (see Berkovec and Stern (1991), Keane and Wolpin (1994), among others). However, many politicians switch parties within a few weeks of winning a seat in parliament, which suggests that they may had not started politics in their most-preferred parties. Moreover, the importance of outside options and the general equilibrium effects render a matching model more suitable to study a political arena.

party is able to produce a sizeable amount of political rents, and the party's leader decides how to use those rents all by herself, which resembles a dictatorship. At the other extreme, when all leaders have similar abilities for leadership and all politicians have large assets, members across different parties would have comparable amounts of rents, which correspond to egalitarian governance.

My model builds on Burdett and Mortensen (1998) and Cahuc, Postel-Vinay, and Robin (2006) and contributes to the theory of on-the-job search in two ways.[4] First, a politician's ranking of the parties' membership values is endogenously determined by the capacity-constrained leaders' rent maximization problem. A party leader, who ranks the politicians vertically, aims to fill her party with politicians who would join the party with the smallest rent shares. However, match frictions prevent her from making membership offers to only the most profitable politician types. Instead, a leader solves her rent maximization problem by choosing the maximum rent share she is willing to offer to each politician-type. The extent to which a leader is willing to negotiate a politician's rent share determines the party's value ranking for the politicians. Pushing up the rent share schedule attracts more members to the party and results in greater rent production, but the leader shares a greater proportion of the party's increased rents with members. This trade-off jointly determines the offers a leader makes to the potential party members and the party's value ranking for each politician-type. The leaders' sorting of members differs from the sorting by firms in the theoretical assignment literature, as the latter arises due to supermodularities in production.[5]

The second contribution to search-theoretic models of labor markets is that politicians differ in their rankings of parties because a party's value has two components that have different returns to party size. Club goods are increasing in party size, so, if they were all that mattered, the payoff from joining a party would increase in party size. This is the case if a politician has very little in private assets. On the other hand, the payoff from private rents is decreasing in party size since richer parties demand a higher tax on private assets to join. If that were all that mattered, the payoff from membership would be a decreasing function of party size. The decreasing returns to party size dominate if a politician has high private assets. If a politician has intermediate levels of private assets, the payoff is a combination of the two, and the politician may rank two parties with different sizes equally. The heterogeneity in politicians' value-ranking of parties depends on their heterogeneous contributions to a party's rents. Accordingly, my model differs from the

[4]For a survey of the search-theoretic models of the labor market, see Rogerson, Shimer, and Wright (2005) and Eckstein and Van der Berg (2007).

[5]See Becker (1973), Burdett and Coles (1997), Shimer and Smith (2000), Lopes de Melo (2013), Lise, Meghir, and Robin (2016), among others.

theoretical assignment literature, where the heterogeneity in workers' most preferred firms stems from complementarities in production (see de Melo (2013) for a review).

This paper is also related to the literature on coalition formation and politicians' career choices. To my knowledge, Desposato (2006) is the first to model the benefits of party membership as the sum of a politician's rent share and the party's club goods and a party's club goods as a function of the members' resources. In his model, a politician's rent share is approximated by his ideological match with his party. My model differs by endogenizing the rent shares of the politicians and considering dynamics, outside options, and match frictions, while abstracting from ideological match. Diermeier, Keane, and Merlo (2005) study the career choices of the members of the U.S. Congress in a dynamic discrete choice framework while abstracting from party switching. Aldrich and Bianco (1992) develop a theoretical model of party affiliation in a bi-party system which finds that the politicians choose their parties strategically to maximize the probability of current and future successes. Diermeier, Eraslan, and Merlo (2003) study the effects of political institutions on government formation in parliamentary democracies. Anderson and Meagher (2012) develop a theoretical model for a two-step democracy of primary elections followed by runoffs, where a party has the twin roles of choosing candidates and promoting their prospects. This paper contributes to these literatures by studying party membership in general equilibrium, considering the outside options of both the leaders and the politicians and comparing the rewards from party membership across different systems.

I estimate the model using a data set I constructed on the Turkish political arena from 1995-2014. During the data period, 28.5% of the members of parliament switched their parties at least once, and the average party switcher switched 1.36 times. The switchers appeared on ballot lists more frequently and stayed in parliament for a longer time. These observations are consistent with the model, as they indicate that politicians switch to parties that provide a greater value to them as opposed to switching randomly.

Unlike the conventional search models which use observed wages in estimation, the estimation procedure in this paper cannot use politicians' rent shares since they are not observed. However, the high party-switching rate across parties of different sizes and politicians' heterogeneous valuations of the two types of political benefits provide the necessary information to identify the functions for production of political rents and club goods. The major identification challenge involves the distribution of unobserved heterogeneity, which enters the hazard rate nonseparably. I show that the results of Evdokimov (2010, 2011) can be applied for nonparametric identification of the distribution of unobserved heterogeneity.[6]

[6]Evdokimov also suggests a constructive estimation procedure for application of his theory. However, as most of the observations in my data have incomplete party membership spells, this estimation procedure

Using my data, I find that a party accumulates club goods more easily than it produces political rents. Interestingly, the estimates show that politicians with better labor market options are less productive to their parties. I interpret these results using the definition of political rents. Specialization in a highly-respected occupation may prevent a politician from engaging in activities to influence government institutions, such as employing one's supporters in municipalities.

In a counterfactual analysis, I study the within-party rent distribution in a system that is more candidate-centered than a closed-list democracy. I find that members of small parties earn higher rents in a party-centered system, where their assets are more productive compared to a candidate-centered system. On the other hand, members of bigger parties prefer a candidate-centered system where they can be more productive with their assets as opposed to a party-centered system. Acemoğlu and Robinson (2012), among others, suggest that inclusive institutions, including a broad distribution of power in society, support economic growth. Although this paper does not elaborate on how a politician or a leader disseminates their political rents to the general public, understanding how the members split a party's rents is a step towards understanding the distribution of political power in a democratic society. The results of this paper are complementary to the literature on politicians' selection into the political market. Caselli and Morelli (2004) and Messner and Polborn (2004) find that sufficiently high rewards from holding a public office attracts higher-quality citizens to the political market. Although this paper abstracts from politicians' selection into the political arena, its results are informative about which political systems make the political arena more attractive to good citizens.

The paper proceeds as follows. The next section provides a more detailed discussion of rent production mechanisms across different democratic systems. Section 3 develops the model, and section 4 shows that the model is nonparametrically identified. Section 5 presents the data, section 6 describes the estimation procedure, section 7 presents the results, and section 8 derives the equilibrium of a candidate-centered system for conducting the counterfactual analysis. Section 9 discusses the institutions that can improve the power distribution in parliamentary democracies and section 10 concludes.

is inapplicable. The equilibrium equations in my model allow me to undertake the duration analysis in a known-component-density mixture model framework, which I use for estimating the structural parameters.

2 Rent production across different systems

This section discusses the institutional details of different democratic systems and how they relate to this paper. The model in section 3 builds on two main institutional assumptions for a politician's membership procedure in a party. First, a party is represented by a leader, and membership in a party is possible only with her approval. Second, party members aggregate their assets to produce rents and club goods. Relaxing the first assumption requires studying coalition formation in a multilateral bargaining environment, which complicates the model tremendously as the number of agents is very large. However, the model can be adjusted for different types of rent-accumulation processes, which would allow for studying the member comparison of political parties across different systems.

Arguably, an electoral system's rent-accumulation process can be inferred from its party-centeredness. In a highly party-centered system, the political parties, not individual candidates, hold most of the influence for election campaigns and other political processes. Accordingly, to study a party-centered system, one can assume a rent-production mechanism in which party members aggregate their assets to collaboratively produce their party's rents. The model in section 3 is an example of a party-centered system, where a politician bargains with the party leader to earn a share of the party's collaboratively produced rents. In a candidate-centered system, on the other hand, individual candidates, not political parties, influence election campaigns and produce votes. Accordingly, to study such a system, one should allow for each member to produce rents on his own in a party. In this case, the bargain between a politician and a party leader would reflect a politician's trade-off between benefiting from a party's club goods in exchange for his supply of rents. In a counterfactual study, I compare the member composition of parties across systems with varying degrees of party-centeredness.[7]

3 Model

In this section, I develop an equilibrium model of coalition formation within parties in a closed-list parliamentary democracy. I model the political arena as a labor market in which a party has an identity of its own rather than a collection of like-minded politicians. A party

[7] Grofman (2005) classifies the electoral systems in their party-centeredness. According to his study, while the closed-list proportional representation system is the least candidate-centered system, single member district plurality, single transferable vote, single non-transferable vote, and cumulative voting can all be considered as the most candidate-centered electoral systems. Open-list proportional representation and mixed systems are considered to have intermediate levels of party-centeredness. Golder (2005) provides detailed information on each of these systems.

is represented by a leader, who is distinguished from other politicians by having the ability to lead a party of a certain size. A party's size, in turn, is equal to the sum of its members' resources. All members of a party combine their resources to produce political benefits, i.e., political rents and club goods. While political rents are private and exclusive, the club goods are non-exclusive to the party members.[8]

The politicians and the leaders are heterogeneous in their resources and party-leading capacities, respectively. When there is a vacancy, a leader seeks members through the random matching process of Burdett and Mortensen (1998). Once brought together pairwise by this process, the leader makes a take-it-or-leave-it-offer over a share of the politician's rent productivity should the politician join the party. If the politician accepts the offer, he utilizes both his rent share and the party's club goods, while the leader receives the rest of the politician's rent production. The random matching process can bring a leader together with either an independent politician or a member of another party. In the latter case, the two parties' leaders enter into a Bertrand competition in the membership value they offer to the politician. Therefore, although a leader has monopsony power while making an offer to a politician, she also has an incentive to renegotiate when the politician gets an offer from an outside party. The extent to which the leaders have incentives to renegotiate a politician's rent share in the party determines the value-ranking of the parties for that politician. Similar to Cahuc, Postel-Vinay, and Robin (2006), henceforth CPR, when two parties compete over the services of a politician, the politician joins the party that he ranks more highly, and he receives a membership value that equals the last value offered by the losing party.

One important difference of this paper from search-theoretic models of the labor market is that a politician's ranking of the parties' values is endogenously determined by the capacity-constrained leaders' rent maximization problem.[9] Since the rent production function aggregates all members' resources, a party leader considers all politician types as perfect substitutes, and, hence, aims at filling her party with the politicians who would join the party with the smallest rent shares. Although the politicians are vertically ranked by a leader, the match frictions prevent her from making offers to only the most profitable politician types. Instead, a leader solves her rent maximization problem by choosing the maximum rent share she is willing to pay to each politician type, which determines the party's value-ranking for the politicians. Pushing up the rent share schedule attracts more members to the party and

[8]The theory of club goods identifies the club goods as locally public goods that are excludable to non-group members (Buchanan 1965, Scotchmer 2002, Chen 2010). Dey and Flinn (2008) also consider public goods in a search environment by modeling health insurance coverage as a public good in a household and study its implications for the labor market careers of spouses.

[9]For a survey of the search-theoretic models of the labor market, see Rogerson, Shimer, and Wright (2005) and Eckstein and Van der Berg (2007).

results in greater rent production, but the leader shares a greater proportion of the party's increased rents with members. This trade-off jointly determines the take-it-or-leave-it offers a leader makes to the potential party members and the party's value ranking for each politician type.[10]

Another difference of this paper from conventional labor search models is that politicians differ in their rankings of the parties because a party's value has two components that have different returns to party size. While the maximum private rents a politician can earn in a party is decreasing, the club goods of a party is increasing in party size. Politicians with greater resources weight the first component of the party value more (as they contribute more to a party's rent production), and therefore tend to rank smaller parties, who pay a greater rent share, higher.[11] Indeed, politicians can be categorized in three groups. Low types consider all parties as "big" in the sense that an incremental increase in party size provides a greater gain in club goods than the politician's loss in terms of private rents. Throughout the section, I use the term "returns to party size" to refer to the change in a party's value for a politician with an incremental change in party size. Accordingly, there are increasing returns to party size for the low types. Similarly, the returns to party size are U-shaped and decreasing for the medium and high politician types, respectively.[12]

The section proceeds as follows. Section 3.1 sets up the model, and section 3.2 de-

[10] There are different models that find assortative sorting in a search environment in the theoretical assignment literature that builds on Becker (1973), who finds sorting due to complementarities in a frictionless environment. Burdett and Coles (1997) find that, when utility is nontransferrable, marriage occurs only between equal classes of males and females. Becker's model with transferrable utility is extended by Shimer and Smith (2000) with search frictions, Lise, Meghir, and Robin (2016) and Lopes de Melo (2013) with on-the-job search, and Eeckhout and Kircher (2010) with directed search. Note that these papers find both negative and positive assortative sorting in different examples of complementarities or substitutes in production.

[11] A politician's trade-off between club goods and private rents earned in a party resembles a worker's trade-off between nonpay and pay characteristics of a firm. The compensating differentials literature suggests that higher pay compensates for undesirable nonpay characteristics of a firm (Rosen 1986). Sorkin (2015) develops a partial equilibrium model in which a firm posts a utility that aggregates both pay and nonpay characteristics of the firm. However, the model does not allow for workers to have different valuation of these two types of firm characteristics, which results in vertical ranking of the firms by all workers. Accordingly, the model does not derive the within-firm distribution of the worker types. On the other hand, Gronberg and Reed (1994), Dey and Flinn (2005), Bonhomme and Jolivet (2009), and Aizawa and Fang (2015) study the value of specific observable amenities in a search environment and allow for the workers to have different preferences for these characteristics. In a preliminary paper by Lopes de Melo (2015), the firms are heterogeneous in both their productivities and compensating differentials. However, to my knowledge, the existing literature does not allow for the firms' and the workers' productivities to determine a firm's nonpay characteristics and the workers' preferences for these characteristics, respectively.

[12] A common feature in the theoretical assignment literature is that the complementarities in production induce the workers to be heterogeneous in their most-preferred types of the firms (see Lopes de Melo (2013) for a review). This model differs from that literature as the heterogeneity in poiticians' preferences over the parties is a result of the politicians' heterogeneous valuations of the private rents over the club goods.

scribes the matching process. Preferences and stationary decision rules of the leaders and the politicians are described in sections 3.3 and 3.4, respectively. Section 3.5 presents the value functions, and section 3.6 derives the ranges of the low, medium, and high politician types. Section 3.7 presents the closed-form solutions of the equilibrium rent shares, which are derived following CPR, and section 3.8 characterizes a Nash equilibrium in the steady-state.

3.1 The environment

Time is continuous. The politicians and the political parties are assumed to live forever.[13] There is a measure M of politicians of whom a fraction φ are independents and a fraction $(1-\varphi)$ have a party affiliation. The politicians are heterogeneous in their infinitesimal resources, denoted by z, that produce political benefits. The distribution of the politicians' resources $L(z)$ is constant over time, with density $\ell(z) > 0$ over $[0, z^{max}]$. A party is represented by a leader, who is distinguished from other politicians by having the ability to lead a party of a certain size. There is a continuum of leaders with a mass normalized to 1. The leaders are heterogeneous in the size of the party that they are capable of leading, denoted by $\tilde{x}$. The party-leading capacities of the leaders are distributed $\tilde{x} \sim \Upsilon(\tilde{x})$, which is constant over time. A party's size is given by the sum of its members' resources. It is assumed that the total assets of the politicians are sufficient for all party leaders to fill their parties up to their leading capacities,

$$\int \tilde{x} d\Upsilon(\tilde{x}) \leq \int_0^{z^{max}} z\ell(z)dz.$$

However, a leader fills her party-leading capacity only when it is profitable, which, in turn, depends on her prospects of finding members who would join the party with the smallest rent shares. Accordingly, the distribution of the party sizes can be different from the distribution of the leaders' capacities. Let x denote the size of a party, distributed $x \sim \vartheta(x)$ with $\frac{d\vartheta(x)}{dx} > 0$ on $[x_{min}, x^{max}]$. Let Z_x denote the set of politicians in type-x parties. Since a party's size is equal to the sum of its members' resources, the law of large numbers implies

$$x\frac{d\vartheta(x)}{dx} = \int_{z \in Z_x} zdz.$$

All members of a party combine their resources to produce political rents and club goods.

[13] There are 35 parties that participated in an election during the data's time span, and 24 of these parties disappeared at least for one election. Although modeling a party's dissolution seems to be empirically important, the model abstracts from this aspect for three reasons. First, not participating in an election does not mean that the party is dissolved. Many political parties continue to hold offices although they do not participate in the elections. Second, four of the parties that disappeared for an election reappeared in a following election. Third, there are very few parties to identify a party's dissolution rate.

Political rents are private and exclusive and are defined as the ability to influence government institutions in one's interest. The total seats in the parliament and the government offices controlled by the party are a few examples of the total political rents of a party. Note that one member's use of the party's rents, such as gaining a seat in the parliament, becoming a governor, or employing one's supporters in the municipalities, prevents the other members from using it. A party of size x produces rents according to $\theta(x)$, which has diminishing returns to scale, i.e., $\theta'(\cdot) > 0$ and $\theta''(\cdot) < 0$.

The club goods of a party are the benefits of belonging to a team, which are provided to all members non-exclusively. While the total seats of a party are its political rents, the pride or the security gained by becoming a member of a strong team are the benefits that are provided to all members non-exclusively. Other examples of a party's club goods include its facilities and legislative achievements. A party of size x produces club goods according to $\psi(x)$, with $\psi'(\cdot) > 0$ and $\psi''(\cdot) < 0$. While all politicians can use their resources to produce rents either as an independent or in a party, accumulation of the club goods requires leadership, and hence club goods can be produced only within a party.

3.2 Matching

A party leader searches for new members when there is a vacancy through an identical, random, time-consuming, sequential matching process in the spirit of Burdett and Mortensen (1998). This process can bring a leader together with either an independent politician or a member of another party. The distribution of the politicians' and the parties' types are common knowledge. However, a politician and a leader see each others' types, and the leader observes the politician's party membership status after being brought together pairwise by the matching process. When a match is formed, the leader makes a monopsonistic offer to the politician over a share of the politician's rent contribution to the party. When a member receives an offer from an outside party, the two parties' leaders enter into a Bertrand competition over the membership value they offer to the politician. Accordingly, although otherwise monopsonist, a leader also has an incentive to renegotiate when the politician is poached by an outside party. The degree to which the leaders have incentives to renegotiate a politician's rent share in the party determines the value-ranking of the parties for that politician. Similar to CPR, this competition resembles a sequential auction game, which results in the politician joining the party that he ranks better, and he receives a membership value that is equal to the last value offered by the losing party. Both agents' types and the politician's party membership status at the time of the match jointly detemine the take-it-or-leave-it offer a leader makes to a politician and the politician's ranking of the party, which is described in section 3.4, after presenting the agents' preferences in the next section. A match can also exogenously break up, which is described in section 3.5.

3.3 The political arena

An independent politician produces rents using his resources. The utility flow to a type-z independent politician is

$$u_0(z) = \theta(z).$$

When a type-z politician joins a type-x party, the politician's contribution to the party's rent production is $\frac{z}{x}\theta(x)$, so that the politician's productivity is proportional to his relative resources in the party.[14] The take-it-or-leave-it offer by the leader gives the politician a share $\phi \leq 1$ of his contribution to the party rents.[15] The politician also benefits from the party's club goods. Hence, the utility flow to a type-z politician who gets a share ϕ of his rent production in a type-x party is

$$u(z, \phi, x) = \phi \frac{z}{x} \theta(x) + \psi(x).$$

I temporarily assume and later show that $\frac{d}{dx}\left\{\frac{z}{x}\theta(x)\right\} < 0$ and $\frac{d^2}{dx^2}\left\{\frac{z}{x}\theta(x)\right\} > 0$.[16] Since $\psi'(x) > 0$, the benefits of party membership have two components with different returns to party size. Accordingly, a politician may receive the same utility flow in two parties with different sizes.

A leader receives all the rents that are not paid to the party members. Accordingly, when

[14]Assuming that a politician's rent contribution to a party is proportional to his relative resources is more compatible with the idea of collective "team" bargaining where the members have different bargaining powers. Although I do not model within-party bargaining of the members, in section 3.8, I find that, conditional on being in a party, the maximum rents a politician can earn in a party is linearly increasing in the politician's resources, as the party leader equates the upper-bound of the rent share a politician can earn in the party across different politician types. Accordingly, if a member has twice the resources of another member, a leader is willing to pay him twice as much rents to keep him in the party. This outcome is the asymmetric Nash bargaining solution in a frictionless environment where a politician's bargaining power is equal to his relative resources in the party (Roth 1979). Accordingly, this specification is consistent with a subgame in which the politicians multilaterally bargain in their party. Note that, to the extent of the frictions, a party leader is able to extract more out of a match surplus, which is explained in the following sections. An alternative specification of a type-z politician's rent productivity in a type-x party would be the politician's marginal productivity in the party, i.e., $\theta(x) - \theta(x - z)$, which is not compatible with the collective team bargaining.

[15]It is possible for a politican to receive negative rents in a party. This occurs when the party's club goods compensate the politician's loss in terms of the rents. For example, although the donors of a party incur some monetary cost, they obtain some ofsetting utility from increasing the party's chances of winning an election, which is a club good, even when they do not get any other benefit from party membership.

[16]A sufficient condition to achieve this is the rent production function to be of the form $\theta(x) = x^\eta$, with $0 < \eta < 1$.

a type-z politician joins a type-x party with rent share ϕ, the utility flow from this contract to the leader is

$$w(z, \phi, x) = (1 - \phi)\frac{z}{x}\theta(x).$$

This paper does not elaborate on how a politician or a leader use their rents. However, it assumes that the politicians use their ability of influencing the government institutions to maintain the support of their electorate, which is reflected in their vote shares. Accordingly, the value of voting for a type-x party for voter i, v_{ixt}, is equal to the sum of the party's rents (as members of a party use their rents to maintain the support of the electorate), the electorate's unobserved, zero-mean, stationary preference shock for the party at time t, ξ_{xt}, and an idiosyncratic taste shock, ϵ_{ixt}. Formally, the value of voting for a type-x party is

$$v_{ixt} = \theta(x) + \xi_{xt} + \epsilon_{ixt}.$$

Accordingly, the voting behavior in this paper does not explicitly take into account strategic voting, valence, or the parties' policy positions that are studied in the previous literature, and assumes that a party leader's ability of leading a party is determined exogenously.[17]

[17] The majority of the existing literature uses a spatial model to study the voters' preferences (see Coughlin (2011) for a survey). Building on Hotelling (1929) and Downs (1957), the spatial theory of voting assumes that each voter has an ideal point in a policy space and votes for the party that locates closest to their ideal points. Accordingly, given the voters' preferences, each political party chooses a point that maximizes its vote share. The simplest model considers two-party competition in a one-dimensional policy space when there is no randomness about the voter preferences. This model has a pure-strategy Nash equilibrium in which both parties locate at the median of the electoral distribution. The literature extends the spatial model by considering higher dimensional policy spaces, multi-party competition, and stochastic voter preferences. The equilibria in the extended models generally exhibit all parties converging to the electoral center (see Schofield 2007 for a review). There are also studies that incorporate nonspatial candidate characteristics such as valence into a spatial model, which yields equilibria featuring the parties diverging from the electoral mean. Intuitively, when valence is exogenous, the low-valence parties locate far from the center as they cannot compete with the stronger parties in the center (Schofield 2007). On the other hand, when the candidates can invest in valence by incuring a cost, the marginal return to valence depends on platform polarization, and, hence, the candidates may choose divergent platforms to soften valence competition (Ashworth and de Mesquita 2007). Furthermore, there are models that incorporate strategic voting into a spatial model. McKevey and Patty (2006) develop a probabilistic voting model where each voter receives a privately observed taste shock for each candidate, and includes the probability of being the pivotal voter in his expected utility of voting for a party. After proving the existence of a Bayesian equilibrium for the voters' strategies, the authors derive a Nash equilibrium of the candidates' strategies, in which all candidates converge to the policy that maximizes the expected sum of voters' utilities. Feddersen, Sened, and Wright (1990) and many others also studied models that allow a voter to not vote for his most preferred candidate when the candidate's chances of winning is low. This explanation may be especially relevant for countries that impose electoral thresholds for a party to win a seat in the parliament, such as Turkey (10%) and Israel (3.25%). In this paper, I assume that, each leader has an ideal point in the policy space which is common knowledge among the electorate. Then, given her exogenous valence, a leader's capacity is determined by the voters' distribution in the policy space. Lastly, there are other models which emphasize the connection between political rents and voting. Defining the political rents as the taxes collected for financing a public good net of its cost,

3.4 Stationary decision rules

This paper studies the stationary equilibrium and abstracts from the dynamics of the transition to the steady-state.[18] In a stationary equilibrium,

1. The distribution of the leader and the politician types, $\Upsilon(\tilde{x})$ and $L(z)$, respectively, are constant. As the flows into and outflows from a party of each politician type balances out, each party's size, and, therefore, political rents and club goods are constant, which yields a constant sampling distribution, $F(x)$. Since the benefits of affiliating with a party depend only on a politician's type, his outside option, and the party's type, a politician's decision to join or leave a party is independent of the other politicians' behaviors.[19]

2. A party leader maximizes her share of the party rents subject to filling the party. Ac-

Polo (1999) finds that, under certain conditions, electoral competition reduces the rents the politicians can extract. In a model with a similar definition of the political rents, Aldashev (2015) finds multiple equilibria in which voter turnout and political rents are jointly determined. Unlike these models, this paper assumes that the voters derive utility from a party's rents through pork-barrel projects. I assume that a party with a greater influence on government institutions has more rents for pork-barrel spending such as creating public employment, investing in infrastructure, and designing policies to please its supporters. Kunicova and Rose-Ackerman (2005) distinguishes pork-barrel from corrupt rent seeking and argues that the former brings electoral support. While this model does not elaborate on how much of a party's rents is used to please the voters, a party with greater rents has more opportunities for this.

[18]During the sample period, new parties were formed and dissolved with some frequency in Turkey. There is also a considerable variation in the existing parties' vote shares over time, which is in contrast with the stationarity assumption. However, an AR (1) regression for the district-level vote shares of the parties finds some evidence in favor of the stationarity assumption, which is explained in section 5. Hence, I assume that, the volatility in voters' preferences do not affect how a team with a certain amount of political assets can produce rents. Since the voters derive utility from a party's rents through pork-barrel and also have time-varying preferences for the parties, the variation in the vote shares are solely explained by the changing voter preferences.

[19]During the data's time-span, there are two cases where a large number of an electorally successful party's members resigned to form a new party. While the first of these attempts was electorally unsuccessful, the second one became the largest party in the parliament in the following election. In order to explain these politicians' behaviors, one can follow any of three strategies. First, by assuming away from the constant, predetermined party sizes, one can allow for a politician's decision to leave a party to induce the other members to leave the party by changing the party's size, and therefore, relative ranking (de Paula 2009). Estimation of this model would require keeping track of the member composition of each party and increase the state vector tremendously. Second, allowing for the politicians to receive correlated private shocks can explain the comovement of the politicians between the parties, which requires modeling the politicians' expectations of the future party sizes and prevents maintaining constant sizes for each party. In this case, a similar equilibrium could still be obtained by assuming that the dissolving parties are immediately replaced with a new party of the same size. Third, by incorporating social learning to allow for a politician to imitate the behavior of his predecessor rather than acting on his information alone, one can attribute the politicians' comovement across the parties to herding or an informational cascade (Banerjee 1992, Bikhchandani, Hirshleifer, and Welch 1992, Smith and Sørensen 2000, Çelen and Kariv 2004). While all of these approaches are interesting, the model abstracts from an explanation for this aspect of the data as there are very few cases that suggest correlated movement of the politicians across the parties. So, the model cannot explain the correlated movement of the politicians across the parties.

cordingly, a leader's optimal take-it-or-leave-it offer to convince a politician to join the party gives the politician the value of his outside option. However, convincing a politician to join the party may not be optimal when a leader expects to fill the vacancy with a more profitable politician. This occurs, for example, if the leader expects to match with a politician who has the same amount of assets but a worse outside option after not making an acceptable offer to the politician she is currently matched. Similarly, when a party member receives an offer from an outside party, it would not be optimal to renegotiate an acceptable offer when the leader expects to fill his vacancy with another politician of the same type with a worse outside option. Thus, a leader's offer depends on both the politician's assets and his outside option. In a stationary equilibrium, a leader's optimization problem for filling the party reduces to deciding the maximum rent share up to which she is willing to renegotiate each type of politicians' share as his outside option improves. Then, the leader offers each politician the value of his outside option as long as providing this value does not require paying a greater rent share than the maximum she has decided to pay. Given each party's size and the upper-bound of the rent share the other party leaders are willing to pay, the maximum share a leader is willing to pay to a politician anchors the overall ranking of the party values for the politician. This is because the other components of the value of membership in a party are either predetermined as a function of its size (i.e., the club goods) or pinned down by the maximum rent share the politician can earn in the party (since, given the club goods and the value of membership in other parties, it determines which parties could win the Bertrand competition over his services when he gets an outside offer, as explained below). This aspect of the model distinguishes it from CPR, who assume a firm's outside option to be zero on each match (which would arise from free entry to and exit from the search market). Accordingly, while a firm is willing to renegotiate a worker's wage up to the match surplus as his outside option improves in CPR, it is possible for a leader to not make an acceptable offer to a politician even in the presence of a match surplus in this paper. The maximum rent share a leader is willing to offer to a politician, in turn, depends on the frictions in the labor market. In a frictionless market, a leader would make acceptable offers to only the most profitable politician types that are just sufficient to fill the party. As the friction level increases, filling the same party requires making acceptable offers to less profitable types, too, due to the decreased chances of meeting the most preferred types. Given a friction level, a leader chooses these upper-bounds to attract members who are just sufficient to fill her party. These ideas can be formalized as follows. Let c denote the indicator function that is equal to 0 if the politician's outside option is to be an independent and 1 if he has an offer from a type-x' party. Since, all else constant, a leader's offer to a politician changes with his outside option, she considers each (z, cx') pair as a different politician type. Let $\phi^{l^*}(z,x)$ denote the maximum share the leader of a type-x party offers to a type-z politician and $\phi^p(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x'))$ denote the rent share that provides a type-(z, cx') politician the same value in party x as he obtains in his outside option. Note that, the only relevant characteristics of the outside offer is the outside party's size and the maximum rent share it pays to the politician, as in a stationary equilibrium, a party's value for a politician can be pinned down by these two objects. Then, a type-x party leader's monopsonistic offer to a type-(z, cx') politician, denoted by $\phi^l(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x'))$, solves

$$\begin{aligned}
&\underbrace{V(z, \phi^l(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x')), \phi^{l^*}(z,x), x)}_{\substack{\text{the take-it-or-leave-it offer of a type-}x\text{ leader} \\ \text{to a type-}(z,cx')\text{ politician}}} \\
&= min\Big\{ \underbrace{V(z, \phi^p(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x')), \phi^{l^*}(z,x), x)}_{\substack{\text{the minimum value a type-}(z,cx')\text{ politician needs to} \\ \text{receive to join party }x}}, \underbrace{V(z, \phi^{l^*}(z,x), \phi^{l^*}(z,x), x)}_{\substack{\text{the maximum value }z\text{ can} \\ \text{receive in a type-}x\text{ party}}} \Big\}.
\end{aligned} \tag{3.1}$$

Equation 3.1 states that, when a type-x party leader meets a type-(z, cx') politician, the take-it-or-leave-it offer she makes to the politician is equal to the minimum value that can convince him to join the party, $V(z, \phi^p(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x')), \phi^{l^*}(z,x), x)$. When the minimum value that convinces the type-(z, cx') politician to accept the offer requires paying him a greater share than the maximum share the leader is willing to pay, $\phi^{l^*}(z,x)$, the leader instead offers $V(z, \phi^{l^*}(z,x), \phi^{l^*}(z,x), x)$, which he turns down.

Specifically, let $V_0(z)$ denote the lifetime utility of being an independent to a type-z politician. Then, the minimum rent share that convinces the politician to join the party, $\phi^p(z, x, 0, \phi^{l^*}(z,x), 0)$ solves $V(z, \phi^p(z, x, 0, \phi^{l^*}(z,x), 0), \phi^{l^*}(z,x), x) = V_0(z)$. Accordingly, an equivalent expression of a leader's stationary decision rule for a $(z, 0)$ match is

$$\underbrace{V(z, \phi^l(z, x, 0, \phi^{l^*}(z,x), 0), \phi^{l^*}(z,x), x)}_{\substack{\text{the take-it-or-leave-it offer of a type-}x\text{ leader} \\ \text{to a type-}(z,0)\text{ politician}}} = min\Big\{ \underbrace{V_0(z)}_{\substack{\text{value of }z\text{'s} \\ \text{outside option}}}, \underbrace{V(z, \phi^{l^*}(z,x), \phi^{l^*}(z,x), x)}_{\substack{\text{maximum value }z\text{ can} \\ \text{receive in a type-}x\text{ party}}} \Big\}. \tag{3.2}$$

When $V_0(z) \leq V(z, \phi^{l^*}(z,x), \phi^{l^*}(z,x), x)$, the politician joins the party with share $\phi^l(\cdot)$, which provides him the same value as being an independent. On the other hand, when $V_0(z) > V(z, \phi^{l^*}(z,x), \phi^{l^*}(z,x), x)$, the leader offers the politician membership with share $\phi^{l^*}(z,x)$, which he turns down.

Now, suppose that the politician is a member of a type-x' party, so that the leaders of type-x and x' parties enter into a Bertrand competition over his services. Given the maximum rent share a type-x' party leader is willing to pay, $\phi^{l^*}(z,x')$, the minimum rent share that convinces the politician to accept the offer of the type-x party solves $V(z,\phi^p(\cdot),\phi^{l^*}(z,x),x) = V(z,\phi^{l^*}(z,x'),\phi^{l^*}(z,x'),x')$. Accordingly, an equivalent expression to equation 3.1 is

$$\underbrace{V(z,\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')),\phi^{l^*}(z,x),x)}_{\substack{\text{the take-it-or-leave-it offer of a type-}x\text{ party}\\ \text{leader to a type-}(z,x')\text{ politician}}}$$
$$= min\Big\{\underbrace{V(z,\phi^{l^*}(z,x'),\phi^{l^*}(z,x'),x')}_{\substack{\text{the minimum value a type-}(z,x')\text{ politician needs}\\ \text{to receive to join party }x}}, \underbrace{V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)}_{\substack{\text{the maximum value }z\\ \text{can receive in a type-}x\text{ party}}}\Big\}. \tag{3.3}$$

When $V(z,\phi^{l^*}(z,x'),\phi^{l^*}(z,x'),x') \leq V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)$, the politician joins party x with share $\phi^l(\cdot)$. When $V(z,\phi^{l^*}(z,x'),\phi^{l^*}(z,x'),x') > V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)$, the leader offers him membership with share $\phi^{l^*}(z,x)$, which he turns down. Note that equations 3.2 and 3.3 are special cases of equation 3.1.

In a stationary equilibrium, the maximum rent share a leader offers to each politician type takes into account her prospects of filling the party. Let $\Phi^{l^*}(z,x)$ denote the distribution of the upper bound of the rent shares the leaders are willing to offer to a type-z politician. A politician type's density in a party depends on this distribution as it anchors the overall ranking of the parties by the politician given the party sizes. Let $\mu_{z,x'|x}(z,x'|x,\Phi^{l^*}(z,x))$ and $\mu_{z,0|x}(z,0|x,\Phi^{l^*}(z,x))$ denote the equilibrium densities of type (z,x') and $(z,0)$ politicians in a type-x party, respectively, and $g_{z|x}(z|x,\Phi^{l^*}(z,x))$ denote the overall density of type-z politicians in the party. Then, the formal statement of a type-x party leader's optimization problem is

$$\max_{\substack{\phi^{l^*}(z,x),\\ \phi^l(\cdot),\\ \forall z,cx'}} \int_0^{z^{max}}\int_{x_{min}}^{x^{max}} \underbrace{(1-\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')))\frac{z}{x}\theta(x)}_{\substack{\text{utility flow to a type-}x\text{ leader when a type-}(z,x')\\ \text{politician joins the party with share }\phi^l(\cdot)}} \times \underbrace{\mu_{z,x'|x}(z,x'|x,\Phi^{l^*}(z,x))}_{\substack{\text{equilibrium density of type-}(z,x')\\ \text{politicians in a type-}x\text{ party}}}\, dx'dz$$
$$+\int_0^{z^{max}} \underbrace{(1-\phi^l(z,x,0,\phi^{l^*}(z,x),0))\frac{z}{x}\theta(x)}_{\substack{\text{utility flow to a type-}x\text{ leader when a type-}(z,0)\\ \text{politician joins the party with share }\phi^l(\cdot)}} \underbrace{\mu_{z,0|x}(z,0|x,\Phi^{l^*}(z,x))}_{\substack{\text{equilibrium density of type-}(z,0)\\ \text{politicians in a type-}x\text{ party}}}\, dz$$
$$\text{subject to } \underbrace{x}_{\text{party size}} = \underbrace{\int_0^{z^{max}} z g_{z|x}(z|x,\Phi^{l^*}(z,x))dz,}_{\text{equilibrium party size}} \tag{3.4}$$

which is solved by her stationary decision rules given in equations 3.2 and 3.3, $\phi^{l^*}(z,x)$ and $\phi^l(z,x,cx',\phi^{l^*}(z,x),c\phi^{l^*}(z,x'))$, $\forall z,cx'$.

3. I conjecture that, in equilibrium, the maximum private rents a politician can earn in a party, $\phi^{l^*}(z,x)\frac{z}{x}\theta(x)$, is decreasing in party size.

4. The distribution across the parties of the upper bound of the rent share a politician can earn in a party determines the value-ranking of the parties for him. Recall that the benefits of party membership have two components that have different returns to party size: while the maximum rents a politician can earn in a party is decreasing, the club goods a politician accesses in a party is increasing in party size. Section 3.6 shows that this feature of the model divides the continuous types of politicians into three categories, named *low, medium*, and *high*. A politician's preference ordering of the parties by their size is increasing, U-shaped, and decreasing for the low, medium, and high types, respectively. Since a politician's stationary decision rule for switching a party is determined by his preference ordering of the parties, the low (high)-type politicians switch only to the bigger (smaller) parties, while a medium-type politician may switch to either a smaller or a bigger party.

5. Given the decision rules of the party leaders, an independent medium-type politician behaves according to his own stationary decision rule. Due to the U-shaped returns to party size, two parties with different sizes may have the same value for a medium-type independent politician. Let x_{a0} and x_{b0} denote the types of the smaller and the bigger parties, respectively, that make a type-z politician equally well-off as being an independent. Note that, these thresholds exist only when the politician's value of being an independent is strictly greater than the lowest point of his U-shaped returns to party size, denoted by $x_0(z)$. When these thresholds exist, the politician is strictly better-off in all parties that are smaller than x_{a0} and bigger than x_{b0}, compared to being in the threshold-type parties. Therefore, the stationary decision rule of an independent medium-type politician is to join party x' if $x' \in \{[x_{min}, x_{a0}] \cup [x_{b0}, x^{max}]\}$ when he receives an offer. These threshold values solve

$$V_0(z) = V(z,\phi^{l^*}(z,x_{a0}),\phi^{l^*}(z,x_{a0}),x_{a0}) = V(z,\phi^{l^*}(z,x_{b0}),\phi^{l^*}(z,x_{b0}),x_{b0}), \tag{3.5}$$

where $V_0(z)$ and $V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)$ are the values of being an independent and being a member of a type-x party with share $\phi^{l^*}(z,x)$ for a type-z politician, respectively. Accordingly, $x_{a0}(\cdot)$ and $x_{b0}(\cdot)$ are functions of z, x_{a0}, $\phi^{l^*}(z,x_{a0})$ and z, x_{b0}, $\phi^{l^*}(z,x_{b0})$, respectively. Note that, when $V_0(z)$ is either tangent to or lower than the minimum point of a medium-type politician's U-shaped returns to party size, the equalities defined in equation 3.5 no longer hold. In this case, a medium-type independent politician's stationary decision rule is to join any party.

6. The low and the high politician types' stationary decision rules when independent are constructed similarly to that of a medium politician type. A low-type independent politician joins party x' if $x' \in [x_{b0}, x^{max}]$, where $x_{b0}(\cdot)$ solves

$$V_0(z) = V(z, \phi^{l^*}(z, x_{b0}), \phi^{l^*}(z, x_{b0}), x_{b0}).$$

Note that, if a low-type politician's value of being an independent is lower than the maximum value he can receive in the lowest party type, x_{min}, then $x_{b0} = x_{min}$. Similarly, a high-type independent politician joins party x' when $x' \in [x_{min}, x_{a0}]$, where $x_{a0}(\cdot)$ solves

$$V_0(z) = V(z, \phi^{l^*}(z, x_{a0}), \phi^{l^*}(z, x_{a0}), x_{a0}).$$

If the value of being an independent is greater than the value a high-type politician receives in the smallest party, i.e., when $V_0(z) > V(z, \phi^{l^*}(z, x_{min}), \phi^{l^*}(z, x_{min}), x_{min})$, the politician does not join any party.

7. Given the decision rules of the party leaders, each politician with a party membership behaves according to his own stationary decision rule. Similar to that of an independent politician, the stationary decision rule of a medium type-z politician in a type-x party is to switch his party if he gets an offer from party x' such that $x' \in \{[x_{min}, x_a) \cup (x_b, x^{max}]\}$, where $x_a(\cdot)$ and $x_b(\cdot)$ solve

$$V(z, \phi^{l^*}(z, x), \phi^{l^*}(z, x), x) = V(z, \phi^{l^*}(z, x_a), \phi^{l^*}(z, x_a), x_a) = V(z, \phi^{l^*}(z, x_b), \phi^{l^*}(z, x_b), x_b), \tag{3.6}$$

and, hence, $x_a(\cdot)$ and $x_b(\cdot)$ are functions of z, x, x_a, $\phi^{l^*}(z, x)$, $\phi^{l^*}(z, x_a)$, and z, x, x_b, $\phi^{l^*}(z, x)$,$\phi^{l^*}(z, x_b)$, respectively. Similarly, a low-type politician in a type-x party switches to party x' when $x' \in (x, x^{max}]$ and a high-type politician in party x switches to party x' when $x' \in [x_{min}, x)$.

3.5 The value functions

In this section, I present the closed forms of the value functions of the politicians and the party leaders using their stationary decision rules. All agents in the model discount the time at rate ρ. An independent politician receives an offer from a political party at rate λ. Given his stationary decision rule in equation 3.5, the lifetime utility of an independent type-z politician is

$$\begin{aligned}
\underbrace{V_0(z)}_{\substack{\text{value of being}\\ \text{an independent}}} &= \underbrace{\tau\theta(z)}_{\text{flow payoff}} + \underbrace{\frac{1}{1+\rho\tau}}_{\text{discounter}} \Big\{ \underbrace{\tau\lambda}_{\text{offer}} \Big[\underbrace{\int_{x_{min}}^{x_{a0}(\cdot)} V(z, \phi^l(z, 0, m, \phi^{l^*}(z, m), 0), m) dF(m)}_{\text{join a small party}} \\
&\quad + \underbrace{\int_{x_{b0}(\cdot)}^{x_{max}} V(z, \phi^l(z, 0, m, \phi^{l^*}(z, m), 0), m) dF(m)}_{\text{join a big party}} + \underbrace{\int_{x_{a0}()}^{x_{b0}(\cdot)} V_0(z) dF(m)}_{\text{reject offer}} \Big] \\
&\quad + \underbrace{(1 - \tau\lambda) V_0(z)}_{\text{no offer}} + o(\tau) \Big\}. \qquad (3.7)
\end{aligned}$$

Reading from left to right, a type-z independent politician receives value $V_0(z)$. This value consists of a flow payoff and a continuation value that he receives for an infinitely small period of time τ, plus a term $o(\tau)$ with the property that $\lim_{\tau \to 0} \frac{o(\tau)}{\tau} = 0$. While the flow payoff is equal to the politician's own rent production, the continuation value, which the politician discounts at rate ρ, weights the expected value of randomly matching with a party and not matching with any party. When the politician matches with a party, he either accepts or rejects the leader's take-it-or-leave-it offer following his stationary decision rule. When the politician joins a type-x' party, he receives a value of $V(z, \phi^l(z, 0, x', \phi^{l^*}(z, x'), 0), x')$, which depends on the leader's take-it-or-leave-it offer $\phi^l(z, 0, x', \phi^{l^*}(z, x'), 0)$. When the politician either rejects an offer or does not receive any offer, he continues to receive the value of being an independent.

Substituting the leaders' stationary decision rules for independent politicians given in equation 3.2 into equation 3.7 and taking the limits as $\tau \to 0$, a politician's lifetime utility of being an independent in a stationary equilibrium solves as

$$V_0(z) = \frac{1+\rho}{\rho}\theta(z). \qquad (3.8)$$

A party member receives offers from outside parties at rate λ. The stationary decision rules described in section 3.4 determine the ranges of party sizes that induce him to switch his party. However, when the politician is not paid the maximum share that his party pays to the politician's type, an offer from an outside party may cause an increase in the politician's rent share without inducing him to switch his party. Suppose that a type-z politician earns a share $\phi < \phi^{l^*}(z, x)$ in a type-x party. The politician's share in the party increases when he gets an offer from a party of type x' such that $x' \in \{[x_a(\cdot), q_a(\cdot)] \cup [q_b(\cdot), x_b(\cdot)]\}$. The threshold values for having a share improvement in the party, $q_a(\cdot)$ and $q_b(\cdot)$, solve

$$V(z,\phi,\phi^{l^*}(z,x),x) \;=\; V(z,\phi^{l^*}(z,q_a),\phi^{l^*}(z,q_a),q_a) = V(z,\phi^{l^*}(z,q_b),\phi^{l^*}(z,q_b),q_b),$$

and, hence, $q_a(\cdot)$ and $q_b(\cdot)$ are functions of x, ϕ, $\phi^{l^*}(z,x)$, q_a, $\phi^{l^*}(z,q_a)$, and x, ϕ, $\phi^{l^*}(z,x)$, q_b, $\phi^{l^*}(z,q_b)$, respectively.

Although a politician's party affiliation may end endogenously when the politician gets an offer from a party with a greater match surplus, the model also allows for exogenous break up of a match, which occurs at rate δ. When a match breaks exogenously, the politician becomes an independent and receives the utility flow associated with that state, $V_0(z)$. Given the politician's stationary decision rule and the ranges of the parties that cause a share improvement, the lifetime utility of a type-z politician in a type-x party when his share of his rent production in the party is ϕ is

$$\begin{aligned}
&\underbrace{V(z,\phi,\phi^{l^*}(z,x),x)}_{\text{value of } z \text{ in } x \text{ with share } \phi} \\
&= \tau\Big[\underbrace{\phi}_{\text{share}} \quad \underbrace{\frac{z}{x}\theta(x)}_{z\text{'s rent production in } x} \quad + \quad \underbrace{\psi(x)}_{\text{x's club goods}}\Big] + \underbrace{\frac{1}{1+\rho\tau}}_{\text{discounter}}\Big\{ \quad \underbrace{\tau\delta V_0(z)}_{\text{exogenous match break-up}} \\
&+ \underbrace{\tau\lambda}_{\text{offer}}\Big[\underbrace{\int_{x_{min}}^{x_a(\cdot)} V(z,\phi^l(\cdot),\phi^{l^*}(z,m),m)dF(m) + \int_{x_b(\cdot)}^{x_{max}} V(z,\phi^l(\cdot),\phi^{l^*}(z,m),m)dF(m)}_{z \text{ switches to either a smaller or a bigger party}} \\
&+ \underbrace{\int_{x_a(\cdot)}^{q_a(\cdot)} V(z,\phi^l(\cdot),\phi^{l^*}(z,x),x)dF(m) + \int_{q_b(\cdot)}^{x_b(\cdot)} V(z,\phi^l(\cdot),\phi^{l^*}(z,x),x)dF(m)}_{z\text{'s share in party } x \text{ increases}}\Big] \\
&+ \underbrace{\int_{q_a(\cdot)}^{q_b(\cdot)} V(z,\phi,\phi^{l^*}(z,x),x)dF(m)}_{z \text{ does not report the offer}} + \underbrace{(1-\tau\lambda-\tau\delta)V(z,\phi,\phi^{l^*}(z,x),x)}_{\text{no offer, no exogenous match break-up}} + o(\tau)\Big\}.
\end{aligned} \tag{3.9}$$

Reading from left to right, a type-z politician in a type-x party with share ϕ has value $V(z,\phi,\phi^{l^*}(z,x),x)$. This value consists of a flow payoff and a continuation value that he receives for an infinitely small time period τ, plus a term $o(\tau)$. The flow payoff is the sum of the politician's share of his rent production in the party and the party's club goods. The continuation value, which he discounts at rate ρ, weights the expected value of three mutually exclusive possibilities. When the match breaks up exogenously, the politician receives the value of being an independent. If the politician gets an offer from a party, which occurs at rate λ, he follows his stationary decision rules to either accept or reject the offer. The politician does not report a rejected offer if the offer does not improve his rent share in the party. Lastly, when the politician neither gets an offer nor the match breaks up exogenously, he continues to receive the value of being a member of party x with share ϕ.

Next, I present the value function of a party leader. The lifetime utility flow to a type-x leader when a type-(z, cx') politician gets a share ϕ of his rent contribution to the party is

$$
\begin{aligned}
&\underbrace{W(z, cx', x, \phi, \phi^{l^*}(z, x), c\phi^{l^*}(z, x'))}_{\substack{\text{value of a type-}x\text{ leader when a type }(z,x')\\ \text{member is paid }\phi}} \\
&= \underbrace{\tau(1-\phi)\frac{z}{x}\theta(x)}_{\text{flow payoff}} + \underbrace{\frac{1}{1+\tau\rho}}_{\text{discounter}} \Big\{ \underbrace{\tau\delta W_0(z, \phi^l(\cdot), \phi^{l^*}(z, x), x)}_{\text{exogenous match break up}} \\
&+ \underbrace{\tau\lambda}_{\substack{\text{politician gets}\\ \text{an offer}}} \Big[\underbrace{\int_{q_a(\cdot)}^{q_b(\cdot)} W(z, cx', x, \phi, \phi^{l^*}(z, x), c\phi^{l^*}(z, x'))dF(m)}_{\text{politician does not report the offer}} \\
&+ \underbrace{\int_{x_{min}}^{x_a(\cdot)} W_0(z, \phi^l(\cdot), \phi^{l^*}(z, x), x)dF(m) + \int_{x_b(\cdot)}^{x^{max}} W_0(z, \phi^l(\cdot), \phi^{l^*}(z, x), x)dF(m)}_{\text{politician switches to }m\text{, leader receives her reservation value on a }(z,m)\text{ match}} \\
&+ \underbrace{\int_{x_a(\cdot)}^{q_a(\cdot)} W(z, m, x, \phi^l(\cdot), \phi^{l^*}(z, x), \phi^{l^*}(z, m))dF(m) + \int_{q_b(\cdot)}^{x_b(\cdot)} W(z, m, x, \phi^l(\cdot), \phi^{l^*}(z, m))dF(m)}_{\text{politician's share increases, leader receives the value of recruiting a type }(z,m)\text{ politician}} \Big] \\
&+ \underbrace{(1-\tau\lambda-\tau\delta)W(z, cx', x, \phi, \phi^{l^*}(z, x), c\phi^{l^*}(z, x'))}_{\text{no offer, no match break-up}} + o(\tau) \Big\}. \qquad (3.10)
\end{aligned}
$$

Reading from left to right, a type-x party leader's value of having a type-(z, cx') politician as a member with share ϕ is $W(z, cx', \phi, x)$. This value function depends on x' directly as the politician's outside option determines the leader's outside option, $W_0(\cdot)$, for a type-(z, cx') politician. In other words, from the point of view of a party leader, each (z, cx') pair constitutes a different politician type. In a stationary equilibrium, where the inflows of any type of politician equals its outflows, a leader's outside option for a type-(z, cx') politician is the value of recruiting another type-(z, cx') politician. The value the leader obtains from this contract is the sum of a flow payoff and a continuation value that she receives for an infinitely small period of time τ, plus a term $o(\tau)$. The flow payoff is the leader's share of the politician's rent contribution to the party. The continuation value, which is discounted at rate ρ, is the weighted sum of the expected values of three mutually exclusive possibilities: exogenous match break-up, the politician getting an outside offer, and neither of these events occurring.

3.6 The low, medium, and high types of politicians

In this section, I show that the continuous distribution of the politician types are divided into three categories in their rankings of the parties. Substituting the stationary decision rules of the party leaders in equation 3.3 into the lifetime utility of a type-z politician in a type-x party in equation 3.9 and taking the limits results in

$$\begin{aligned}&[\rho+\delta+\lambda\bar{F}(q_b(\cdot)+\lambda F(q_a(\cdot))]V(z,\phi,\phi^{l^*}(z,x),x)\\&\quad=\phi\frac{z}{x}\theta(x)+\psi(x)+\delta V_0(z)+\lambda[F(x_a(z,x))+\bar{F}(x_b(z,x))]V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)\\&\quad+\lambda\int_{q_b(\cdot)}^{x_b(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m)+\lambda\int_{x_a(\cdot)}^{q_a(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m).\end{aligned} \tag{3.11}$$

Imposing $\phi=\phi^{l^*}(z,x)$ in equation 3.11, which has the properties that $q_b(\cdot)=x_b(\cdot)$ and $q_a(\cdot)=x_a(\cdot)$, results in the maximum value a type-z politician can earn in a type-x party is

$$V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x) \quad = \quad \frac{\phi^{l^*}(z,x)\frac{z}{x}\theta(x)+\psi(x)+\delta V_0(z)}{\rho+\delta}, \tag{3.12}$$

with derivative,

$$\frac{dV(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)}{dx}=\frac{1}{\rho+\delta}\Big[z\underbrace{\frac{d}{dx}\Big\{\phi^{l^*}(z,x)\frac{\theta(x)}{x}\Big\}}_{<0}+\underbrace{\psi'(x)}_{>0}\Big]. \tag{3.13}$$

Equation 3.13 characterizes the returns to party size for a type-z politician. The maximum value a politician can earn in a type-x party, $V(z,\phi^{l^*}(z,x),x)$, has two components that have different returns to party size: while the upper bound of the rents a politician can earn in a party is decreasing, the club goods a politician accesses in a party is increasing in party size. This feature of the model divides the continuous distribution of the politician types in three categories in terms of their ranking of the parties, named *low, medium*, and *high*.

The low types of politicians consider all parties as "big" and rank the bigger parties better because their resources are low enough that their loss in terms of the private rents is always dominated by their gain in terms of the club goods as party size increases. For all politician types such that $z\leq\underline{z}$, the second term in equation 3.13 dominates the first term in all parties, and hence, the returns to party size are increasing on $[x_{min}, x^{max}]$. The threshold type $\underline{z}$ that separates the low types from the medium types receives the same marginal utility from the club goods and the upper bound of the private rents when he is a member of the minimum party type, x_{min}. All higher politician types value the private rents more, and, therefore, have $\frac{dV(z,\phi^{l^*}(z,x_{min}),\phi^{l^*}(z,x_{min}),x_{min})}{dx} < 0$.[20] Accordingly, the threshold type $\underline{z}$ solves

$$\underline{z}\frac{d}{dx}\left\{\phi^{l^*}(\underline{z}, x_{min})\frac{\theta(x_{min})}{x_{min}}\right\} = -\psi'(x_{min}).$$

Similarly, for all politician types such that $z \geq \bar{z}$, the first term in equation 3.13 dominates the second term in all parties, and hence, the returns to party size is decreasing on $[x_{min}, x^{max}]$. The politicians whose assets are within the range $[\bar{z}, z^{max}]$ are named as "high" because their resources are high enough that their loss in terms of the private rents is never dominated by their gain in terms of the club goods as party size increases. The politicians in this group see all parties as "small," and rank the smaller parties better. The threshold type $\bar{z}$ that separates the high types from the medium types receives the same marginal utility from the party's club goods and the upper-bound of the private rents he can receive in the biggest party type, x^{max}. All higher politician types value the private rents more, and, therefore, have $\frac{dV(z,\phi^{l^*}(z,x^{max}),x^{max})}{dx} < 0$. Accordingly, the threshold type $\underline{z}$ solves

$$\underline{z}\frac{d}{dx}\left\{\phi^{l^*}(\underline{z}, x^{max})\frac{\theta(x^{max})}{x^{max}}\right\} = -\psi'(x^{max}).$$

Finally, the politicians with assets within range $(\underline{z}, \bar{z})$ have a wiggled returns to party size. Note that since $\psi'(x) > 0$ and $\psi''(x) < 0$, a sufficient condition for the medium-type politicians to have a U-shaped returns to party size is that $\frac{d}{dx}\left(\frac{\phi^{l^*}(z,x)\theta(x)}{x}\right) < 0$ and $\frac{d^2}{dx^2}\left(\frac{\phi^{l^*}(z,x)\theta(x)}{x}\right) > 0$. I temporarily assume and later show that this condition holds.

Let $x_0(z)$ denote the lowest point of a medium type-z politician's U-shaped returns to party size. The politician considers all parties that are smaller than $x_0(z)$ as "small" because over this range, the loss in private rents dominates the gain in club goods as party size increases. Similarly, he considers all x such that $x > x_0(z)$ as "big."

3.7 The share equation

In this section, I present the closed-form solutions of the rent share a politician earns in a party for both cases of the politician's initial party membership status. These equations,

[20]Note that, since the support of x is $[x_{min}, x^{max}]$, at x_{min} (at x^{max}), only the right (the left) derivative exists.

which are derived in Appendix A following CPR, reflect an option value effect, i.e., the politicians forgo today's rents in expectation of higher rents in the future. However, the resulting share equation is different from CPR in two aspects. First, while all workers rank the firms vertically in CPR, the politicians have heterogeneous ranking of the parties in this paper. Accordingly, the option value effect includes the possibilities of having a rent share improvement that would arise from receiving offers from both the smaller and the bigger parties. Second, the authors allow a worker to get a positive share of the match surplus when s/he has bargaining power. In contrast, since a party leader formulates her stationary decision rules to extract the entire match surplus in any match, the resulting share equation is independent of a politician's bargaining power in this paper.

When a type-z politician joins a type-x party from the pool of independents, the rent share the party he earns in the party, $\phi^l(z, x, 0, \phi^{l^*}(z, x), 0)$, solves

$$\begin{aligned}\phi^l(z, x, 0, \phi^{l^*}(z, x), 0) = \frac{x}{z\theta(x)}\Big\{&[\rho V_0(z) - \psi(x)] \\ &- \lambda \int_{x_{b0}(\cdot)}^{x} \underbrace{\frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm}}_{>0} \bar{F}(m)dm \\ &+ \lambda \int_{x_a(\cdot)}^{x_{a0}(\cdot)} \underbrace{\frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm}}_{<0} F(m)dm\Big\}. \end{aligned} \tag{3.14}$$

Now suppose that a type-(z, x') politician joins party x. The rent share he earns in the party, $\phi^l(z, x, x', \phi^{l^*}(z, x), \phi^{l^*}(z, x'))$, solves

$$\begin{aligned}\phi^l(z, x&, x', \phi^{l^*}(z, x), \phi^{l^*}(z, x')) \\ &= \frac{x}{z\theta(x)}\Big\{\phi^{l^*}(z, x')\frac{z\theta(x')}{x'} + [\psi(x') - \psi(x)] \\ &- \lambda \int_{q_b(\cdot)}^{x_b(\cdot)} \underbrace{\frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm}}_{>0} \bar{F}(m)dm \\ &+ \lambda \int_{x_a(\cdot)}^{q_a(\cdot)} \underbrace{\frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm}}_{<0} F(m)dm\Big\}. \end{aligned} \tag{3.15}$$

The first (second) integral term in equations 3.14 and 3.15 reflect the effect of getting an offer from a bigger (smaller) outside party on the rent share. As the returns to party size is increasing (decreasing) over this range, the integrand is positive (negative). Accordingly, assuming that $\lambda > 0$, the rent share which convinces a politician to stay in the party is decreasing in the rate of offer arrival. The politician is willing to accept a smaller rent share today in expectation of higher future rents, which is the option value effect.

3.8 Steady-state equilibrium

In this section, I first define the conditions for a steady-state equilibrium, then, assuming the existence of an equilibrium, I characterize the properties of an equilibrium.[21] Let

$$\Omega^{Pr} = \left[\begin{array}{ccccccc} \Upsilon(\tilde{x}) & \ell(z) & \theta(x) & \psi(x) & \rho & \delta & \lambda \end{array} \right],$$

$$\Omega^{P} = \left[\begin{array}{cccc} x_{a0}(\cdot) & x_{b0}(\cdot) & x_{a}(\cdot) & x_{b}(\cdot) \end{array} \right],$$

$$\Omega^{L} = \left[\begin{array}{cc} \phi^{l^*}(z,x) & \phi^{l}(z,x,cx',\phi^{l^*}(z,x),c\phi^{l^*}(z,x')) \end{array} \right],$$

and

$$\Omega^{E} = \left[\begin{array}{c} \Gamma_{\phi,z,x}(\phi,z,x) \end{array} \right],$$

be the the sets of the primitives, the politicians' and the leaders' policy functions, and the equilibrium distributions, respectively.

Definition 1. In a stationary equilibrium, for each $x \sim F(x)$ and $z \sim L(z)$, the following conditions hold,

(i) Given Ω^{Pr}, Ω^{P}, Ω^{L}, and Ω^{E}, the decision rules Ω^{L} of a party leader defined in equation 3.1 maximize the value of each match to the party leader, which leads to $W(\cdot)$ in equation 3.10.

(ii) Given Ω^{Pr}, Ω^{P}, Ω^{L}, and Ω^{E}, each politician, when independent, maximizes the lifetime utility of being an independent by following the decision rule ($x_{a0}(\cdot)$ and $x_{b0}(\cdot)$) to join a party defined in equation 3.5, which leads to $V_0(z)$ in equation 3.7.

(iii) Given Ω^{Pr}, Ω^{P}, Ω^{L}, and Ω^{E},each politician, when member of a party, maximizes the lifetime utility of being a member of that party by following the stationary decision rule ($x_a(\cdot)$ and $x_b(\cdot)$) defined in equation 3.6, which leads to $V(\cdot)$ in equation 3.9.

(iv) Given Ω^{Pr}, Ω^{P}, Ω^{L}, and Ω^{E}, the within-party share distribution of politicians, $\Gamma_{\phi,z,x}(\phi,z,x)$, yields a sampling distribution, $F(x)$, and a density of each politician type in each party, $\mu_{z,cx',x}(z,cx',x|\Phi^{l^*}(z,x))$ and $g(z,x|\Phi^{l^*}(z,x))$.

(v) Given Ω^{Pr}, Ω^{P}, and Ω^{L}, the within-party share distribution, $\Gamma_{\phi,z,x}(\phi,z,x)$, and the ratio of independent politicians, $\varphi_z, \forall z$, are constant.

[21] It is on the research agenda to investigate the existence and the uniqueness of an equilibrium.

Next, I assume that a steady-state equilibrium that is consistent with the above definition exists. Given the distribution of the leaders' stationary decision rules, $\Phi^{l^*}(z,x)$, Appendix B adjusts the steps taken in CPR for the possibility of a U-shaped returns to party size to derive the following steady-state conditions.

- *The proportion of independent politicians is*

$$\varphi_z = \frac{\delta}{\delta + \lambda[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]}. \tag{3.16}$$

- *The joint density of type-z politicians in type-x parties is*

$$g(z,x|\Phi^{l^*}(z,x)) = \frac{\delta(\delta+\lambda)}{[\delta + \lambda[F(x_a(\cdot)) + \bar{F}(x_b(\cdot))]]^2}\tilde{\ell}(z)f(x), \tag{3.17}$$

where

$$\tilde{\ell}(z) = \frac{\ell(z)}{\delta + \lambda[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]} \tag{3.18}$$

is defined to be *the effective density* of type-z politicians, as it weights the density of a politician type by its demand from the parties.

- *The joint density of type-$(z, q_b(\cdot))$ politicians in type-x parties is*

$$\mu_{z,q_b(\cdot),x}(z,q_b(\cdot),x|\Phi^{l^*}(z,x)) = 2\frac{\delta(\delta+\lambda)\lambda f(x)f(q_b(\cdot))\tilde{\ell}(z)}{[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]]^3}. \tag{3.19}$$

- *The joint density of type-$(z,0)$ politicians in type-x parties is*

$$\mu_{z,0,x}(z,0,x|\Phi^{l^*}(z,x)) = \frac{\delta}{[\delta + \lambda[F(x_{a0}(z) + \bar{F}(x_{b0}(z))]]}\tilde{\ell}(z)f(x). \tag{3.20}$$

- *The within-party share distribution of type-z politicians in type-x parties is*

$$\Gamma_{\phi|z,x}(\phi|z,x,\Phi^{l^*}(z,x)) = \Big(\frac{\delta + \lambda[\bar{F}(x_b(\cdot)) + F(x_a(\cdot))]}{\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]}\Big)^2. \tag{3.21}$$

- *Equilibrium party size is*

$$x \;=\; \int_0^{z^{max}} zg(z,x|\Phi^{l^*}(x))dz. \tag{3.22}$$

Equation 3.16 is slightly different from its counterpart in CPR. As the authors take a firm's outside option to be zero independent of the match productivity, a firm is willing to renegotiate each worker's share of the match surplus up to the match productivity as the worker's outside option improves. This results in an equilibrium in which a worker accepts any firm's offer, and hence a constant rate of unemployment across different worker types. On the other hand, in this paper, the proportion of independent politicians varies over the politician types because the leaders sort their members.

Equation 3.17 is identical to its counterpart in CPR only for the low types of politicians, as these types switch only to the bigger parties, i.e., they have $x_a(\cdot) = x_{min}$.

Note that type-$(z, 0)$ and type-(z, x') politicians' densities in a type-x party (equations 19 and 20) are independent of the

Theorem 1 states that there is no Nash equilibrium in which a party leader does not negotiate a member's rent share in the party up to the match surplus.

Theorem 1. *Let $\underline{z}(x)$ be the least profitable politician type a type-x party leader needs to hire to fill her party when $\phi^{l^*}(z,x) = 1, \forall z$. Let $\Pi(z, \phi^{l^*}(z,x), x)$ denote the total profitability of type-z politicians to a type-x leader when the maximum rent share a type-z politician can earn in a type-x party is $\phi^{l^*}(z,x)$. There is no Nash equilibrium in which $\phi^{l^*}(z,x) \neq 1$ $\forall z$ such that $\Pi(z, \phi^{l^*}(z,x), x) > \Pi(\underline{z}(x), \phi^{l^*}(\underline{z}(x), x), x)$, and $\phi^{l^*}(z,x) \leq \phi^{l}(z, x, 0, \phi^{l^*}(z,x), 0)$ for all other z.*

Proof. A party leader solves her rent-maximization problem in equation 3.4 by following the stationary decision rule in equation 3.1, i.e., she chooses $\phi^l(z, x, cx', \phi^{l^*}(z,x), c\phi^{l^*}(z,x'))$ to give the politician the value he receives in his outside option as long as providing this value does not require paying him a greater rent share than the maximum the leader has decided to pay, $\phi^{l^*}(z,x)$. Equation 3.17 shows that, the total density of type-z politicians in type-x parties, $g(z, x|\Phi^{l^*}(z,x))$, depends on $\phi^{l^*}(z,x)$ through a type-z politician's party-switching thresholds, $x_a(\cdot)$ and $x_b(\cdot)$. However, as long a politician ranks a type-x party better than the type-x' party, the density of type-(z, x') and type-$(z, 0)$ politicians in a type-x party does not change with $\phi^{l^*}(z,x)$ (equations 3.19 and 3.20).

One can write the total profitability of type-z politicians to a type-x party leader as

$$\begin{aligned} \Pi(z, \phi^{l^*}(z,x), x) &= \frac{z\theta(x)}{x}\Big\{ \int_{x_a(\cdot)}^{x_b(\cdot)} (1 - \phi^l(z, x, x', \phi^{l^*}(z,x), c\phi^{l^*}(z,x')))\mu_{z,x'|z}(z, x'|x, \Phi^{l^*}(z,x))dx' \\ &+ (1 - \phi^l(z, x, 0, \phi^{l^*}(z,x), 0))\mu_{z,0|}(z, 0|x, \Phi^{l^*}(z,x))\Big\}, \end{aligned} \tag{3.23}$$

with $\frac{d\Pi(z,\phi^{l^*}(z,x),x)}{d\phi^{l^*}(z,x)} > 0$ as long as $\phi^{l^*}(z,x) \leq 1$ because we have that $\frac{d\mu_{z,x'|x}(z,x'|x,\Phi^{l^*}(z,x))}{d\phi^{l^*}(z,x))} = 0$, $\frac{d\mu_{z,0|x}(z,0|x,\Phi^{l^*}(z,x))}{d\phi^{l^*}(z,x))} = 0$, and $\frac{d\phi^l(\cdot)}{d\phi^{l^*}(z,x))} < 0$.

Suppose that, given all other leaders' stationary decision rules, a type-x party leader follows the rule $\phi^{l^*}(z,x)$. As the party leader seeks the most profitable politician types, it must be the case that the party leader makes acceptable offers to only the most profitable politician types whose densities in the party are just sufficient to satisfy the constraint in equation 3.4. Let $\underline{z}(x)$ be the least profitable politician type that the leader has to hire to fill her party under the rule she follows. Consider a type-z politician such that $\Pi(z,\phi^{l^*}(z,x),x) > \Pi(\underline{z}(x),\phi^{l^*}(\underline{z}(x),x),x)$ and suppose that $\phi^{l^*}(z,x) \neq 1$. This cannot be an equilibrium, as setting $\phi^{l^*}(z,x) = 1$ would make a type-z politician even more profitable. The Nash equilibria is setting $\phi^{l^*}(z,x) = 1$, $\forall z$ such that $\Pi(z,\phi^{l^*}(z,x),x) > \Pi(\underline{z}(x),\phi^{l^*}(\underline{z}(x),x),x)$, and $\phi^{l^*}(z,x) \leq \phi^p(z,x,cx',\phi^{l^*}(z,x),c\phi^{l^*}(z,x'))$ for all other z, where $\phi^p(\cdot)$ is the minimum rent share that convinces a type-(z,cx') politician to join a type-x party and $\underline{z}(x)$ denote the least profitable type that needs to be hired to meet the constraint in equation 3.4 when the leader sets $\phi^{l^*}(z,x) = 1$ for all members of the party.

□

4 Identification

This section discusses identification of the primitives of the model presented in section 3. For each party k, the observable time-variant variables are the party's vote share in each of eighty-five districts and each of five elections, $\{\nu_{kct}\}_{t=1,c=1}^{5\ \ 85}$. The dataset also includes S time-invariant party characteristics, $\{\tilde{\varnothing}_{ks}\}_{s=1}^{S}$. For each politician i, the observable variables are the M characteristics of his private assets, $\{y_{im}\}_{m=1}^{M}$, the duration of each of I spells with no party affiliation, $\{t_{0i\iota}\}_{\iota=0}^{I}$, and the duration of each of J spells with a party affiliation, $\{t_{ij}\}_{j=0}^{J}$.[22] Note that, when a politician competes with the political parties in an election by running as an independent candidate, he is considered as a party of one (with no club goods).

The primitives of the model are the distributions of the leaders' exogenous capacities, $\Upsilon(\tilde{x})$, the politicians' private assets, $L(z)$, and the voters' time-varying preferences for K parties, $\Xi(\xi_{1t},\xi_{2t},...,\xi_{Kt})$, the rent production function, $\theta(\cdot)$, the club goods production function, $\psi(\cdot)$, and the rates of offer arrival, λ, exogenous separation, δ, and discounting, ρ. A politician's private assets have observed and unobserved components. Hence, the primitives include the distribution of the unobserved heterogeneity in politicians' assets, $H(\epsilon)$, and the contribution of the observed characteristics to a politician's assets, $\{\beta_m\}_{m=1}^{M}$. In addition to the primitives, estimation of the model requires identification of each party's type, $\{x_k\}_{k=1}^{K}$,

[22]Note that, the panel is not balanced as many parties and politicians disappear in some of the elections.

and the equilibrium distribution of the parties from which the politicians sample offers, $F(\cdot)$.

Appendix D shows identification of the structural parameters and functions. Section D.1 describes how the model maps the transition parameters (λ, δ) to the duration of party membership as these parameters pin down the rate at which a politician leaves a party. Section D.2 discusses the application of Evdokimov (2011) for the nonparametric identification of the time-invariant hazard of leaving a party. The derivatives of the hazard of leaving a party with respect to the observed politician characteristics, in turn, identify the contribution of these characteristics to a politician's assets. Since a party aggregates all members' resources, using the equilibrium density of politician types in a given party (which depends on the hazard of leaving a party, the transition parameters, and the density of politician types), one can calculate each party's assets, which yields the sampling distribution from which the politicians draw membership offers, $F(x)$.

Given each party's assets, the rent production function, $\theta(\cdot)$, and the distribution of the voters' time-varying preferences for a party, $\Xi(\xi_{1t}, \xi_{2t}, ..., \xi_{Kt})$, are mapped into the vote shares via Hotz and Miller (1982) inversion of the voters' choice probabilities, as described in section D.3. Having identified the rent production function and given each agent's type, the club goods production function, $\psi(\cdot)$, is identified from the conditional likelihood of observing a party affiliation duration. Intuitively, since the richer politicians value the private rents more than the club goods, given a party's rents, the variation in the hazard of leaving the party across different politicians identifies the party's club goods. Similarly, having identified the political benefit production functions and given each agent's type, the discount rate can be identified from the unconditional likelihood of joining a party from the pool of independents. This is because the value of being an independent is equal to the discounted value of the politician's rent production on his own, and, hence, the hazard of joining a party is determined by the discount rate when all the other relevant objects are given, as described in section D.5. A party leader fills her party up to her entire capacity only when it is profitable. Although it is possible to recover each party's type as explained in Section D.2, since a party's total assets does not have to equal its leader's asset-accumulation capacity, the distribution of the leaders' exogenous capacities, $\Upsilon(\cdot)$, is not identified.

5 Data

Estimation of the model presented in Section 3 requires data on politicians' characteristics, party choices, the lengths of each party membership spell, and the details of the electoral environment under which each politician made his party choices. I gather these data for politicians who appeared in a party's ballot list which participated in an election in Turkey during 1995-2014. There are 33 parties and 35,648 politicians of whom about 1,900 won a seat in parliament in the sample. The Official Gazette of Turkey issues each candidate's occupation, education level, electoral district, and the ranking in his party's list about two months before an election.[23] If a politician wins a seat, a more detailed resume is published in parliament's website.[24] Moreover, the archives of the daily newspapers provide the exact date at which a member of the parliament, henceforth MP, switches his party. About 4% of the entire sample and 28% of the MPs switched a party at least once during their political careers. The low party-switching rate in the entire sample can be due to censoring in the data. A large majority of the politicians in the sample appeared in a party's ballot lists only once, partly because many parties participated in only one election. As the elections are party-centered, politicians who ran for an election but do not win a seat rarely appear in the media. Therefore, it is not possible to observe their party choices late in their careers after running for office.

The section proceeds as follows. Section 5.1 describes Turkey's electoral institutions during the 1995-2014 period. Section 5.2 presents the characteristics and the party choices of the politicians who appeared on a party's ballot list at least once during this period.

5.1 Electoral environment

This section describes the electoral institutions of Turkey and discusses the validity of the stationary assumption during the data period. Turkey uses a closed-list proportional representation system to distribute 550 parliamentary seats to the parties. Each party lists its candidates, in order of priority, for each of the 81 electoral districts before an election. Each voter observes the ballot lists, and s/he can vote for a party as a whole rather than for individual candidates. In order to win a seat in parliament, a party has to gain at least 10% of the national votes. The seats are distributed to the parties that clear the electoral threshold via the d'Hondt method. According to the d'Hondt method, when there is no electoral threshold, a party's share of the total seats approaches to its vote share as the number of seats to be allocated approaches to infinity.

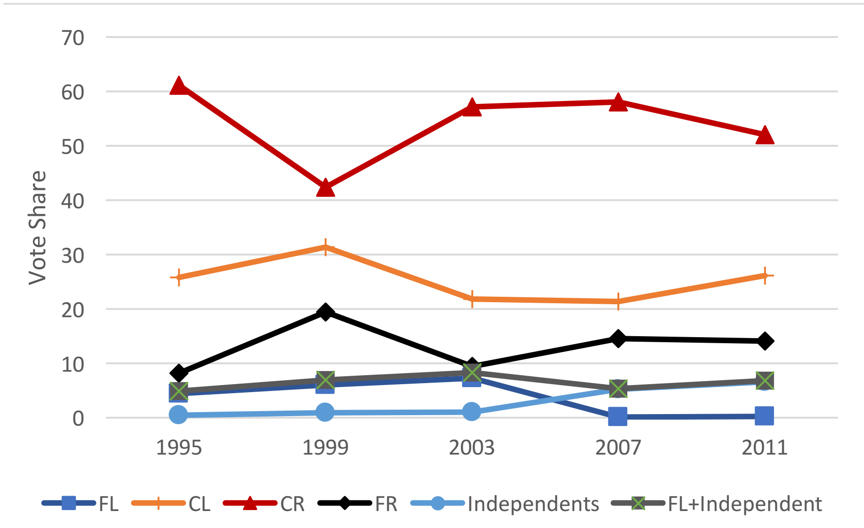
During 1995-2014, 5 elections were held to distribute parliamentary seats to the parties.[25] Thirty-three parties competed in these elections, but only 3 participated in all 5. The number of parties that participated in a given election ranged from 13 to 21. Due to the electoral

[23] http://www.resmigazete.gov.tr/

[24] https://www.tbmm.gov.tr/TBMM_Album.htm

[25] The MPs are elected for a four-year term in an election. However, the prime minister can call an early election, or an early election can result following a no-confidence vote.

Figure 1: Total vote shares of categories of the political spectrum



threshold, at most five parties gained seats in the parliament in each electoral term.

Figure 1 presents the vote share of each category of the political spectrum (far left FL, center left CL, center right CR, and far right FR) in each of five elections.[26] The figure shows that the total vote share of the center-right parties exceeded that of all other categories in each of the elections. The center-left parties had the second highest total vote share and were followed by the far-right parties. The total vote share of the far-left parties decreased significantly in 2007 and 2011 because the members of the major far-left party ran as independent candidates in these elections. Therefore, the sum of the vote shares of the far-left parties and the independents converged to the total vote share of the independent candidates in these years.

During the data period, in contrast to the smaller parties' steady (and low) vote shares, the bigger parties' vote shares were highly volatile over time. Recall that the model in section 3 assumes constant party types and explains the variation in parties' vote shares over time solely with time-varying voter preferences. The stationarity assumption is crucial for the implications of the model, and, therefore, a large volatility in vote shares may cast doubt on the applicability of the model to the Turkish political arena. To have a better understanding of the validity of the stationarity assumption, I run an AR (1) regression for the district-level vote shares of the parties. The correlation coefficient estimate from this regression is 0.945 and significant at the 0.01 level. This high persistence provides some evidence in favor of the stationary assumption, i.e., the current vote share is a good predictor of the next period's

[26]These data are obtained from the Supreme Election Committee (www.ysk.gov.tr) and Turkish Statistical Institute (www.turkstat.gov.tr).

vote share. Moreover, this result is robust to various specifications. However, when I run the AR (1) regression analysis separately for each category of the political spectrum, I find that the center-right and the far-left parties' vote shares are highly volatile, with correlation coefficient estimates of 1.009 and 1.149, respectively. However, the stationarity assumption holds both when each party's type is constant over time and when a party is replaced by a new party of the same type. Despite the highly volatile vote shares of the parties in these categories, as new parties are formed and dissolved with some frequency in these categories of the political spectrum, the stationarity assumption may also hold by replacement of a party with a similar successor.

Note that, during this period, there were five cases where the Constitutional Court closed a party. In each of these cases, members of the party formed a new party after the court's decision. Accordingly, I consider the successor party as the same as the original party.[27] Moreover, there was one case where a party changed its name, from the Socialist Power Party (SIP) to the Communist Party of Turkey (TKP). Since the party did not declare any changes in its policies, I consider the subsequent party to be the same as the original party.

5.2 Politicians

This section summarizes the data on politician's characteristics and party choices. Section 5.2.1 describes the construction and the limitations of the data. The characteristics and the political careers of the politicians are summarized in sections 5.2.2 and 5.2.3, respectively.

[27] The center-right Welfare Party (RP) was closed by the Constitutional Court in 1998 for violating the secularist articles of the constitution. The members of the party immediately formed the Virtue Party (FP), which, in turn, was also banned from politics in 2001 for the same reason. After the closure of the FP, the former members were separated into two groups. The first group called themselves *the traditionalists*, and formed the Felicity Party (SP) after the closure of the SP. The second group, calling themselves *the reformists*, formed the Justice and Development Party (AKP). Accordingly, I consider the RP, the FP, and the SP as the same parties, but consider the AKP as a new party. Similarly, the far-left People's Democracy Party (HADEP) was closed by the Constitutional Court for becoming the focal point of activities against the unity of the state. The members of the party immediately formed the Democratic People's Party (DEHAP), which, in turn, was closed by the Constitutional Court in 2005 on the same grounds. DEHAP later merged with the Democratic Society Party (DTP), which, in turn, was closed by the Constitutional Court in 2009 for similar reasons, and reemerged as the Peace and Democracy Party (BDP). In 2011 elections, 35 members of the BDP ran as independent candidates and got elected into the parliament. These politicians later joined a newly-formed far-left party, HDP, which declared changes in party's traditional policies. Accordingly, while I count the HADEP, the DEHAP, the DTP, and the BDP as the same parties, I consider the HDP as a new party.

5.2.1 Construction and limitations of data

This section explains the construction of data and its limitations arising from the limited information about politicians who ran for an election but never won a seat in parliament. The Official Gazette of Turkey publishes each party's ballot list for each of eighty-one electoral districts about two months before an election. The Gazette states each politician's name, ranking on the ballot lists, occupation, and education level. If a politician becomes a member of parliament, henceforth MP, a more detailed resume can be found in the parliament's website. Moreover, the exact date at which an MP switches a party can be found in the archives of the daily newspapers. Accordingly, the biographical and party membership information for MPs is rich and subject to minimal error.

Unlike the MPs, most of the politicians who never gain a seat in the parliament do not appear in the media, and, thus, the information about them is limited to what is provided by the Official Gazette. I do not use information, such as birth date and birth city, that is only available for MPs and not for other politicians.

If a given name appeared multiple times on the ballot lists, I required that at least two of the observables, i.e., party, electoral district, occupation, or education level, are the same and the other observables are not very different to identify those politicians as the same. Accordingly, it is possible that different individuals were counted as the same if their observable characteristics were highly similar. It is also possible that the same politician was considered as a different individual in different years due to spelling mistakes in their names.

The biographical information provided by the Official Gazette is not detailed. While the occupation of a politician is usually described by a broad categorization, their education level is categorized as either primary, secondary, or college. For a small fraction of politicians who were listed as candidates in different years, the stated education differed, sometimes decreasing or increasing. In those cases, if the politician's occupation, such as being a lawyer, implied an education level, I adjusted his education accordingly. Otherwise, I used the most frequently occurring education level if he was listed more than twice or the latest education level if he was listed twice, as record-keeping technology may have improved over time.

Politicians' names indicate their gender. For the gender-neutral names, which are not many, I assumed that the politician was male, as a great majority of the politicians in the sample are male.

If a politician stated different occupations in different years, then I treated each occupation as the truth for that year and assumed that a politician acquired the skills of all stated occupations up to that year.[28] I followed the 2010 Standard Occupational Classification (SOC) system of the Bureau of Labor Statistics to classify politicians into occupational categories and constructed several additional occupational categories considering the specific features of the labor market studied in this paper and data limitations. There are 142 politicians whose either occupational or educational information is missing, which are dropped from the estimation sample.

Finally, as the politicians who never gain a seat in the parliament rarely appear in the media, the information related to their party choices is limited. As studied in the next subsection, a large majority of these politicians were listed as a candidate in only one election. Although not reappearing in their parties' ballot lists in consecutive elections does not mean that their party affiliations ended, only 4% of these politicians reappeared in their parties' ballot lists after disappearing for an election. This is partly because some of the parties competed in only one election. I assume that if a politician disappeared from data, his membership ended at some date between the date of the election he lost and the date of the consecutive election. Accordingly, these politicians' membership durations are interval-censored. If a politician appeared in a different party's ballot list in a consecutive election, the exact duration of his party membership spell is still unknown, as the data tells only that he switched his party between the two election dates. Accordingly, these politicians' membership durations are also interval-censored. The total number of uncensored, interval-censored, and right-censored party membership spells in the sample are 849, 27,123, and 7,809, respectively.

5.2.2 Characteristics of politicians

This section summarizes the data on politicians' characteristics. Table 1 shows the mean values of the observed characteristics for the entire sample as well as separately for the MPs, and compares the party switchers to the whole sample. As all of the observables are dummy variables, the standard errors are not reported.

[28]It is possible to assume different skill aggregation mechanisms. For example, one may think of changing an occupation as an indicator of not being good at the initial occupation. It is also possible to think that the occupational skills that are not actively used depreciate over time. However, in this paper, it is assumed that an occupation's contribution to political assets is either through social skills or networking that are beneficial in everyday life. For example, a politician who has improved his communication skills by working as a teacher continues to use these skills when he later becomes a businessperson. Similarly, an engineer might lose his technical abilities after changing his occupation, but he continues to benefit from the analytical skills he gained by obtaining an engineering degree in his daily life, which is a political asset. Moreover, as estimation allows for unobserved heterogeneity in politicians' assets, it is possible to capture the degree at which the politicians differ in their unobservable characteristics that cause them to acquire different amounts of skills from the same occupation.

Table 1: Summary statistics for politician characteristics

	All politicians		MPs	
	Whole Sample	**Switchers**	**Whole Sample**	**Switchers**
Characteristics	**Mean**	**Mean**	**Mean**	**Mean**
Female	0.156	0.067	0.077	0.054
College	0.498	0.696	0.907	0.876
Production	0.089	0.050	0.006	0.013
Architecture and engineering	0.087	0.146	0.183	0.168
Education, training, and library	0.072	0.095	0.147	0.129
Legal	0.058	0.112	0.159	0.161
Business and financial operations	0.055	0.053	0.048	0.046
Health practitioners and technical	0.049	0.081	0.109	0.105
Management	0.033	0.046	0.043	0.050
Life, physical, and social sciences	0.031	0.076	0.091	0.098
Arts, design, entertainment, sports, and media	0.030	0.059	0.039	0.055
Governance	0.017	0.035	0.066	0.070
Business	0.333	0.316	0.176	0.209
Other	0.175	0.145	0.178	0.153
Political Career				
# of party switches	0.047	1.169	0.385	1.364
# of winning a seat	0.077	0.630	1.438	1.691
# of participating in an election	1.194	2.388	2.067	2.624
N	35,648	1,449	1,912	540

Notes: Although there are a total of twenty observable characteristics, this table combines seven characteristics in the *other* category.

Although there are about 15 parties competing in an election and most of the political parties that participate in an election show candidates for each of 550 parliamentary seats, there are only 35,648 politicians in the sample, as some of these politicians entered the ballot lists multiple times. There are only 1,449 party switchers in the entire sample and 540 among the MPs. Note that most of the politicians appeared on the ballot lists only once, as indicated by an average number of participating in an election that is slightly above 1. Accordingly, the information related to the political careers of these politicians is limited to a very short duration, and it is not possible to observe whether they switched a party late in their careers, after running for office. The average number of times someone switches a party is about three times higher for the party switchers both among the MPs and in the entire sample. Finally, in the entire sample, a politician appears on a ballot list an average of 1.194 times. This number is twice as high for the party switchers. Similarly, party switchers are more likely to win a seat in parliament. If it is the case that winning a seat is valuable, these observations are consistent with the theoretical model, as they imply that a politician switches to a party that provides a better membership value to him.

According to the table, the large majority of politicians are male. Female politicians make up 15.6% of the entire sample and only 7.7% of MPs. Although about half of the politicians do not have a college degree, 90% of MPs have one. Party switchers are also much more likely to have a college degree than are non-switchers.

The occupational categories in Table 1 are not mutually exclusive as some of the politicians reported different occupations in different election years. The politicians who have occupational backgrounds in education, law, healthcare, management, sciences, community and social service, and governance are more likely to win a seat in the parliament as they make up a greater fraction of the MPs compared to the entire sample.

The politicians who have occupations in production, law, management, sciences, media, and governance make up a greater proportion of the switchers than the entire sample both among all politicians and the MPs. On the other hand, the politicians who have occupations in engineering and education make up a greater proportion of the switchers in the entire sample but not among the MPs.

Table 2: The observed matrix of party switches

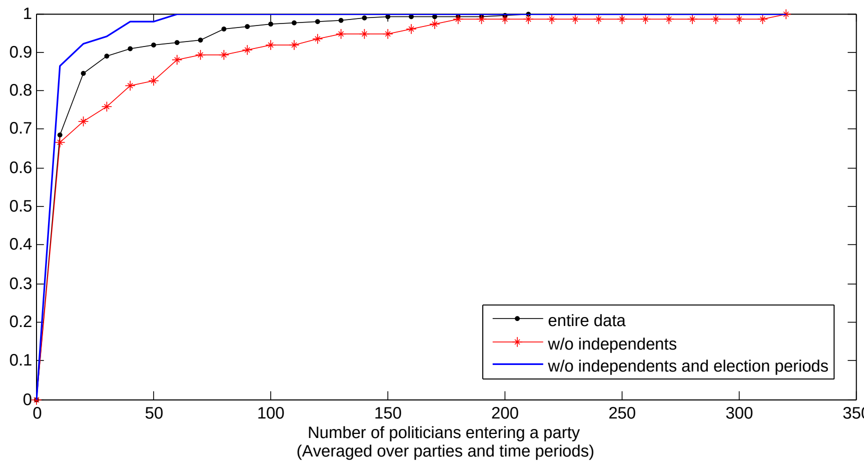
To:	**Stay**	**New FL**	**New CL**	**New CR**	**New FR**	**New Indep.**
From:						
Old FL	8,546	125	13	13	0	25
Old CL	9,762	20	184	63	23	57
Old CR	24,076	4	51	689	76	114
Old FR	6,318	1	6	35	28	25
Old Indep.	1,261	20	33	69	19	-

Notes: The cell (x, y) presents the total number of switches from a party in category x to a party in category y.

5.2.3 Party switching

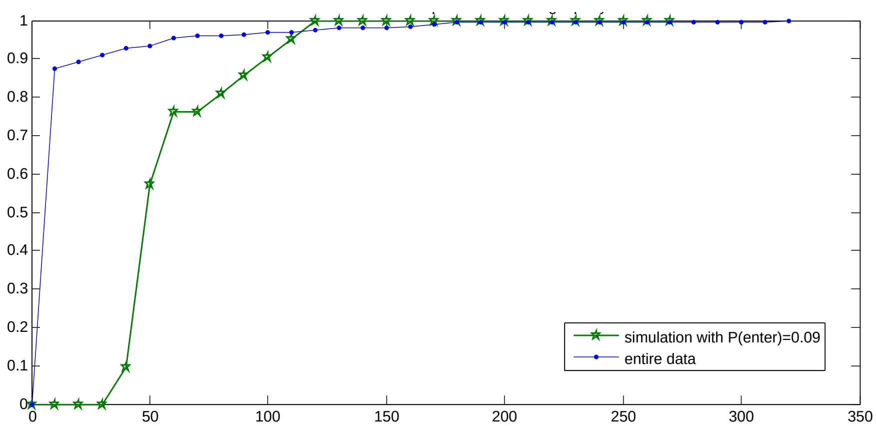
This section presents a more detailed analysis of party switching by politicians and investigates the validity of the assumption of uncorrelated party switching decisions by politicians. Table 2 shows the observed matrix of switches across the categories of the political spectrum, where each politician's party choice from one year to the next is counted as a separate observation. According to the table, the politicians who switched parties present examples of switches both within and across all categories of the political spectrum; however, the majority of the switches occurred between the parties of the same category.

Figure 2: The CDF of the number of MPs entering a party



Notes: This figure plots the CDF of the number of MPs entering a party, defined as $F(s^E) = \frac{\frac{1}{KT} 1[0 < \sum_{i \in k} s^E_{ikt} \leq s^E]}{\sum_{kt} 1[0 < \sum_{i \in k} s^E_{ikt}]}$, where p^E_{kt} is the proportion of politicians entering party k at time t.

Figure 3: The simulated CDF of the number of people entering a party



Notes: This figure compares the empirical and the simulated CDFs of the number of MPs entering a party. The simulation is implemented using a binomial distribution with 550 trials (number of seats in parliament) and a success probability of 0.09 (average party switching rate across years).

The model presented in section 3 studies the stationary equilibrium where each party's size is constant. This implies that, given the primitives, the value of a party to a politician depends only on the party's and the politician's types, and not on other politicians' behaviors. In the data, on the other hand, there are examples of both correlated and uncorrelated party switching decisions by politicians.[29] To gain an insight about the correlation in politicians' party-switching decisions, Figure 2 plots the cumulative distribution of the number of politicians entering a party. In the entire MP data, the probability that about 10 or less politicians will enter a new party is 0.90. This probability increases smoothly to 1 for about 60 or less politicians entering a new party, and, thus, presents some evidence that it is unlikely to observe a large number of politicians entering a party simultaneously. Note that this CDF averages all years and parties, and considers each independent politician as a party on his own. When independent politicians are excluded from the sample, the CDF shifts down, which indicates that the behaviors of politicians who are not independents have a higher correlation. However, when we exclude the election years together with the independent politicians, the CDF shifts up and imply only very little correlated behavior. This has two reasons. First, many switches occur in elections years. Second, many new politicians enter the sample in election years, whose membership initiation cannot be correlated with existing members' behaviors. To gain a better insight on the correlation in politicians' behaviors, Figure 3 compares the empirical CDF with its simulated counterpart. The simulation is implemented using a Binomial distribution with 550 trials (equal to the number of seats in parliament) and a success probability of 0.09 (equal to the average party switching rate across years among MPs). The comparison of the CDFs show that, if the politicians behaved completely randomly, we would observe more politicians entering a new party simultaneously conditional on less than about 110 politician entering a new party. On the other hand, the probability that more than 110 hundred politicians will enter a new party simultaneously is greater in the data than the simulation. The higher probability of a large number of politicians entering a new party in the data compared to the simulation may be due to either herding or a large number of politicians entering the sample in the election

[29]There are two cases where a large number of the members of a party resigned simultaneously to form a new party. Both of these events occurred during the 1999-2002 electoral term. As discussed in section 5.1, after the Constitutional Court closed the Virtue Party (FP), forty-seven members of the party formed the Justice and Development Party (AKP) in August 2001, which formed a majority government in the following three electoral terms. Similarly, seventy-five members of the Democratic Peace Party (DSP) resigned to form the New Turkey Party (YTP) in July 2002. Moreover, instead of participating in the election as a party, seven members of the Great Union Party (BBP) cooperated with the Motherland Party (ANAP) in 1995 and gained seats from its candidate lists. These politicians switched back to their own party immediately after winning the election. The model cannot explain these behaviors. Lastly, some of the members of the Peace and Democracy Party (BDP) won seats as independent candidates in 2007 and 2011, and formed a group for their party in the parliament immediately after winning the election. Throughout estimation, I consider them as winning the election as a party, since they ran as independents only to surpass the electoral threshold.

years. When election years are excluded, the empirical CDF stochastically dominates the simulated CDF. Accordingly, there is some evidence that the behaviors of politicians were not highly correlated.

6 Estimation

This section describes the procedure for estimating the structural parameters and functions. The econometrician observes $\{\{y_{im}\}_{m=1}^{M}, \{t_{il}, \{s_{ilk}\}_{k=1}^{K}, d_{il}^{n}, d_{il}^{i}, d_{il}^{r}\}_{l=1}^{L_i}\}_{i=1}^{N}$, and $\{\nu_{0ct}, \nu_{kct}\}_{k=1}^{K}{}_{c=1}^{C}{}_{t=1}^{T}$ where y_{im} is the mth observable characteristic of politician i, t_{il} is the length of politician i's lth spell, s_{ilk} is a dummy variable that is equal to 1 if the lth spell of politician i is in party k, $d_{il}^{n}, d_{il}^{i}, d_{il}^{r}$ are indicator variables for the uncensored, interval-censored, and right-censored observations that are equal to 1 if the lth spell of politician i has the relevant type of censoring and 0 otherwise, ν_{0ct} is the share of the electorate who did not vote for any party in district c at time t, and ν_{kct} is party k's vote share in district c at time t. The objects to be estimated are the rates of offer arrival and exogenous separation, λ and δ, the functions for rent and club goods production, $\theta(\cdot)$ and $\psi(\cdot)$, the contribution of each of M observable characteristics to a politician's assets, $\{\beta_m\}_{m=1}^{M}$, the distribution of unobserved heterogeneity in politicians' assets, $H(\epsilon)$, each party's size, $\{x_k\}_{k=1}^{K}$, and the distribution of the voters' preferences for each party, $\{\Xi_k(\xi_{kct})\}_{k=1}^{K}$.

There are four main challenges in estimation of the structural parameters. First, the only information on the outcome of the bargaining between a politician and a party is the politician's position on the ballot list, which can be only an imperfect proxy for a politician's rent share in the party. Therefore, unlike the conventional search models which use observed wages in estimation, the estimation procedure in this paper cannot use politicians' rent shares. Second, a party's total assets, which is the sum of all members' assets, is not observed. Third, estimation of the distribution of the unobserved heterogeneity is challenging as the unobserved heterogeneity enters the hazard rate nonseparably. Evdokimov (2010, 2011) suggests a procedure for estimating this distribution following his identification strategy; however, its application requires at least one complete spell for each observation, which is not available. Fourth, some important components of a politician's assets, such as valence, are not observed.

Fortunately, the one-to-one relationship between the voters' party-specific value functions and the vote shares allows for estimating the party sizes and the distribution of the voters' preferences for each party given an assumption for the rent production function. This procedure is similar to recovering a firm's unobserved productivity level by estimating its production function as in CPR.

Given the estimated party sizes and the assumed form of the rent production function, it is possible to undertake the duration analysis in the framework of a finite mixture model with known component densities. The club goods production function, the contribution of the observed characteristics to a politician's assets, and the distribution of unobserved heterogeneity in politicians' assets are estimated parametrically using this formulation.

The rent production function connects these two parts of estimation, i.e., recovering the party sizes from the vote shares and using the estimated party sizes for estimation of the other parameters. Scaling the rent production function in the first part results in scaled estimates of party sizes. Using these scaled estimates in the second part, in turn, it is possible to scale the other parameters to obtain the exact same likelihood value as the one that was obtained before scaling. Accordingly, the rent production function can be identified only up to a scale normalization.

The next subsection describes the estimation procedure in more detail.

6.1 Estimation procedure

The labor market transition parameters (λ, δ) are estimated by maximizing the unconditional likelihood of the observed party-membership durations (equation D.1). This section derives the likelihood functions for estimating the structural parameters and equilibrium objects. Note that, throughout estimation, the rent production function is normalized to be of the form $\theta(x) = log(x)$. To ensure that each party member has nonnegative assets, I assume that politician i's assets are given by

$$log(z_i) = \sum_m y_{im}\beta_m + \epsilon_i. \tag{6.1}$$

Moreover, I assume that $\psi(x) = x^{\eta_1}$, $\epsilon_i \sim iidN(0, \sigma_\epsilon^2)$ for $k = 1, ..., K$, and $\xi_{kct} \sim iidN(0, \sigma_{\xi_k}^2)$, and estimate η_1, σ_ϵ, $\{\sigma_{\xi_k}\}_{k=1}^K$, and $\{\beta_m\}_{m=1}^M$.

The section proceeds as follows. In section 6.1.1, I describe the procedure for estimating the party sizes using the vote shares. Section 6.1.2 describes the procedure for estimation of the other structural parameters using the duration data.

6.1.1 Estimation of party sizes

This section describes estimation of party types and the distribution of voters' preferences for each party using vote shares. Recall that, given voters' preferences for not voting for any party, ξ_{0ct}, and voting for each party, $\{\xi_{kct}\}_{k=1}^{K}$, there is a one-to-one relationship between the voters' choice specific value functions and the parties' vote shares (equation D.13),

$$\underbrace{log(\nu_{kct})}_{\substack{\text{log of the vote share of party } k \\ \text{in city } c \text{ at time } t}} - \underbrace{log(\nu_{0ct})}_{\substack{\text{log of the proportion of people who} \\ \text{did not vote in city } c \text{ at time } t}} = \theta(x_k) + \xi_{kct} - \xi_{0ct},$$

where $\xi_{kct} \sim iid(0, \sigma^2_{\xi_k})$. It is assumed that the utility a voter derives from not voting for any party is the same across different districts and constant over time, i.e., $\xi_{0ct} = \eta_0$, $\forall t$.[30] Then, the probability of observing $\{\nu_{kct}, \nu_{0ct}\}_{k=1,c=1,t=1}^{K\ \ C\ \ T}$ is

$$\begin{aligned}\prod_{k,c,t} p(\xi_{kct} &= log(\nu_{kct}) - log(\nu_{0ct}) - log(x_k) + \eta_0) \\ &= \prod_{k,c,t} \frac{1}{\sigma_{\xi_k}} \phi\Big(\frac{log(\nu_{kct}) - log(\nu_{0ct}) - log(x_k) + \eta_0}{\sigma_{\xi_k}}\Big), \end{aligned} \tag{6.2}$$

where $\phi(\cdot)$ denotes the standard normal density function. The likelihood function in equation 6.2 is maximized with respect to $\{x_k, \sigma_{\xi_k}\}_{k=1}^{K}$. The estimate of the sampling distribution, $F(x)$, is the cumulative distribution of the estimated party sizes.

6.1.2 Duration analysis

This section describes the implementation of the duration analysis for estimation of the structural parameters. Having estimated the labor market transition parameters, each party's size, and the sampling distribution, the contribution of the observable characteristics to a politician's assets, $\{\beta_m\}_{m=1}^{M}$, the standard deviation of the distribution of unobserved heterogeneity in party members' assets, σ_ϵ, and the parameter characterizing the club goods production function, η_1, are estimated by maximizing the likelihood of the observed membership spell durations conditional on party sizes and the observed politician characteristics.

The equilibrium equations in the model allow me to undertake the duration analysis in a known-component density finite mixture model framework. This is because the time-invariant hazard rate varies over the politician types in a systematic way. A politician may leave a party either through exogenous separation, which occurs at rate δ, or by receiving an offer from a party that he ranks more highly and accepting it. A low politician type in a type-x party ranks the bigger parties better, and, hence, the probability that he gets an acceptable offer is $\lambda\bar{F}(x)$. Similarly, a high politician type in a type-x party ranks the

[30] To ensure positive estimates for party sizes, throughout estimation, I set $\eta_0 = min_{k,c,t}\{log(\nu_{kct} - log(\nu_{0ct}))\}$.

smaller parties more highly than x, and, thus, the probability of getting an offer from a better party for him is $\lambda F(x)$. A medium-type politician may rank both a smaller and a bigger party better than a type-x party, as he has a U-shaped returns to party size. Indeed, if a type-x party is bigger than the lowest point of a medium type politician's U-shaped returns to party size, then he considers a type-x party as "big", and ranks all bigger parties more highly than it. A medium-type politician in a type-x party may also have a lower party switching threshold. For example, if a type-x_1 party such that $x_1 < x$ provides the same value to him as he obtains in a type-x party, then he is better-off in all parties smaller than x_1 and bigger than x. Accordingly, the probability of getting an offer from a better party for him is $\lambda[F(x_1) + \bar{F}(x)]$. As the number of parties is finite, there are only a finite number of possible hazard rates a party member can have. Accordingly, the continuous types of members of a party can be divided into a finite number of groups, each having a different hazard rate. When the party sizes and the labor market transition parameters are known, these hazard rates, that is, the component densities, are also known. Moreover, when both the party sizes and the functions for rent and club goods production are known, the threshold politician types that separate different hazard-rate groups can be solved. This is because, the threshold types have the same valuation for two parties. A party's value to a politician, in turn, is a function of the types of the politician and the party as well as the functions for rent and club goods production. Then, the probability of a politician being in a certain hazard-rate group is the probability of his assets, containing observable and unobservable elements, being within the thresholds which define that group. Conditional on having a certain hazard rate, the probability of observing a particular membership duration for a politician has the exponential form. The likelihood of observing that membership duration, in turn, integrates out the possible hazard rates a politician can have. As the possible hazard rates, that is, the component densities, are predetermined conditional on the party sizes, a politician's likelihood contribution is maximized by choosing the probabilities of the politician having each of the possible hazard rates. These probabilities, in turn, depend on the rent and club goods production functions, the politician's observed characteristics, and the distribution of the unobserved heterogeneity.

Formally, suppose that the party sizes are sorted in increasing order, i.e., $x_1 < x_2 < .... <$

x_K. Politician i in party k may have any of K+1 possible hazard rates,

$$a_{ik}(z_i, x_k) = \begin{cases} a_{ik0} = \delta + \lambda \bar{F}(x_k) & \text{if } exp(\sum_{m=1}^{M} y_{im}\beta_m + \epsilon_i) \leq \underline{z}, \\ \vdots & \\ a_{ikk+1} = \delta + \lambda[F(x_k) + \bar{F}(x_{k+1})] & \text{if } z_{kk} \leq exp(\sum_{m=1}^{M} y_{im}\beta_m + \epsilon_i) \leq z_{kk+1}, \\ \vdots & \\ a_{ikK} = \delta + \lambda F(x_k) & \text{if } exp(\sum_{m=1}^{M} y_{im}\beta_m + \epsilon_i) \geq z_{kK-1}, \end{cases}$$

where $\underline{z}$ is the threshold politician type that separates the medium types of politicians from the low types, z_{k1} is the threshold politician type that separates the politicians with hazard rate $\delta + \lambda[F(x_1) + \bar{F}(x_k)]$ from those with hazard rate $\delta + \lambda[F(x_2) + \bar{F}(x_k)]$, and so on.

As a type-z_{kj} politician obtains the same value in parties x_k and x_j, we have that $V(z_{kj}, \phi^{l^*}(z_{kj}, x_j), \phi^{l^*}(z_{kj}, x_j), x_j) = V(z_{kj}, \phi^{l^*}(z_{kj}, x_k), \phi^{l^*}(z_{kj}, x_k), x_k)$, which, after substituting the equilibrium value that $\phi^{l^*}(z_{kj}, x_j) = \phi^{l^*}(z_{kj}, x_k) = 1$ into equation 3.12, boils down to $z_{kj}\frac{\theta(x_j)}{x_j} + \psi(x_j) = z_{kj}\frac{\theta(x_k)}{x_k} + \psi(x_k)$. Moreover, a type-$z_{kk}$ politician's lowest point of the U-shaped returns to party size is a type-x_k party, so he ranks all parties better than x_k. This implies that the derivative with respect to party size of the value of membership in a type-x_k party for a type-z_{kk} politician is 0, i.e., $z_{kk}\frac{d}{dx}\left(\frac{\theta(x_k)}{x_k}\right) + \psi'(x_k) = 0$ (equation 3.13). Accordingly, the threshold politician types can be written as

$$z_{kj}(x_k, x_j) = \begin{cases} \frac{x_k^{\eta_1} - x_j^{\eta_1}}{log(x_j)/x_j - log(x_k)/x_k} & \text{if } j \neq k \\ \frac{-\eta_1 x_k^{\eta_1 - 1}}{(1 - log(x))/x_k} & \text{if } j = k. \end{cases}$$

Then, the probability that politician i in party k having the hazard rate a_{ik0} is

$$\begin{aligned} \pi_{ik0}(x_{min}, \{y_{im}\}_{m=1}^{M}) &= Pr[a_{ik}(z_i, x_k) = a_{ik0}] \\ &= Pr[exp(\sum_{m=1}^{M} y_{im}\beta_m + \epsilon_i) \leq \underline{z}] \\ &= \Phi\Big(\frac{log(\underline{z}) - \sum_{m=1}^{M} y_{im}\beta_m}{\sigma_\epsilon}\Big), \end{aligned}$$

where $\Phi(\cdot)$ denotes the standard normal distribution function. The probability of the politician having the hazard rate a_{ikj}, $\pi_{ikj}(x_k, x_j, \{y_{im}\}_{m=1}^{M})$, is written similarly. Given each of these probabilities, the likelihood contribution of politician i's lth spell is

$$L_{il}(t_{il}|x_k, \{y_{im}\}_{m=1}^{M}) = \sum_{j,k} s_{ilk}\pi_{ikj}(x_k, x_j, \{y_{im}\}_{m=1}^{M}) \times \Big[f(t_{il}|a_{ikj})^{d_{il}^{n}} S(t_{il}|a_{ikj})^{d_{il}^{r}} [S(t_{1il}|a_{ikj}) - S(t_{2il}|a_{ikj})]^{d_{il}^{i}}\Big]. \quad (6.3)$$

The likelihood function for observing the entire data is the product of the likelihood contribution of each spell of each politician,

$$L = \prod_{i,l} L_{il}(t_{il}|x_k, \{z_{im}\}_{m=1}^{M}), \quad (6.4)$$

which is maximized with respect to $\{\beta_m\}_{m=1}^{M}$, η_1, and σ_ϵ.

Table 3: Estimates of the labor market transition parameters

δ	0.0045 (0.0001)***
λ	0.0137 (0.0015)***
κ	3.0495 (0.3789)***
$1/\lambda$	72.9 (7.9)***
$1/\delta$	222.2 (4.9)***
$\frac{1}{N}\sum_i logL_i$	-0.6759

Notes: Estimates are per week. Standard errors are in parentheses. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

7 Results

This section presents the results. Table 3 shows the estimates of the labor market transition parameters. The estimates show that the rate at which a match occurs is bigger than the rate at which it exogenously breaks, since the estimated offer arrival rate, λ, is three times as high as the estimated exogenous separation rate, δ. While the average duration until having a need for a new party, $\frac{1}{\delta}$, is 222.2 weeks or 4.3 years, the average waiting time between two outside offers, $\frac{1}{\lambda}$, is 73 weeks or 1.4 years. On average, a politician is poached 3.05 times by the outside parties before his membership ends exogenously. The average length of a party affiliation spell implied by the model, $\int_0^\infty tp(t)dt = \frac{1}{2}[\frac{1}{\delta} + \frac{1}{\delta+\lambda}]$, is about 138 weeks or 2.66

Figure 4: Estimated party sizes

years.[31] The low average duration of party membership reflects the unstable party structures in a closed-list parliamentary democracy.

Figure 4 plots the maximum likelihood estimates of the party types.[32] The correlation between the vote shares and the estimated party sizes is 0.98. Thus, the ordering of the estimated party sizes is highly consistent with the ordering of the parties' vote shares. This result is not surprising because the party sizes are mapped directly into the vote shares through the maximum likelihood estimation of voters' choice probabilities. There are 23 small parties whose sizes are estimated to be within the 26-213 band and 9 parties whose sizes are distributed within the 458-2,736 band. There is one outstandingly large party, with an estimated size of 8,830, which formed a majority government in the last three electoral terms in the sample.

Table 4 presents the maximum likelihood estimates of the structural parameters. Throughout estimation, the parameter characterizing the club goods production function, η_1 is restricted to be between 0 and 0.99.[33] The estimate of η_1, 0.9878, is close to the upper-bound of the restriction. As the rent production function is normalized to be of the form $\theta(x) = log(x)$, the estimated value of η_1 indicates that, a party accumulates club goods faster than it produces rents.

Out of twenty observable politician characteristics, only female, retired, other, and pro-

[31]Using a matched employer-employee panel of French data covering the period 1993-2000, CPR finds that the mean employment duration in manufacturing, construction, trade, and services sectors are 9.1, 6.9, 6.4, and 7.1 years, respectively. The mean party membership duration is lower than the mean employment duration in the labor market, reflecting the instability in party structures.

[32]The estimates of the party types and the voters' preferences for each party are all significant at the 0.01 level.

[33]The upper bound of η_1 is chosen as 0.99 because the likelihood function behaves badly when η_1=1.

duction have positive coefficient estimates. Among these four characteristics, only production has a statistically significant coefficient estimate. These results may seem striking, especially because having a college degree, or an occupation in business, bureaucracy, healthcare, management, engineering, education, life sciences, bureaucracy, and the legal sector all have statistically significant negative coefficient estimates. On the other hand, the constant in a politician's assets is estimated to be a highly significant and large positive number. These results can be interpreted in the following way. Recall that the political rents are defined as the ability to influence government institutions in one's interest. The most important political assets that can be used to influence government institutions, like valence or other people skills, may have been captured by the constant term, and not by the observable politician characteristics such as occupation and education level. Having a college degree or specialization in a prominent occupation may prevent a politician from engaging in activities to influence government institutions in his interest, such as employing one's supporters in municipalities, which would explain the negative coefficient estimates of these variables. This reasoning still cannot explain the highly significant positive coefficient estimate of the production occupations. Note that, the majority of the politicians who are classified in this category declared their occupation as *laborer*. The social security system in Turkey divides the employees into two groups as public-servants and laborers. Some politicians may have declared their occupation based on this classification. Moreover, some politicians, especially members of the left-wing parties such as the Labor Party, may have declared their occupation as a laborer because of their ideologies. Accordingly, the unexpected coefficient estimate of the production occupation may be a result of data limitation.

Finally, although having a college degree is a negative political asset, half of the politicians and 90% of the MPs in the sample have one. Understanding this phenomenon requires studying the strategic ballot-list formation procedure. While forming the ballot lists, a party leader is likely to consider a politician's services to the party both before and after the elections. Before the elections, the politicians use their assets for pork-barrel spending, which would maximize the party's votes. After the elections, on the other hand, the politicians who gain a seat in the parliament work in legislative activities. Serving in committees such as healthcare, defense, or justice, require specialization in these fields. Accordingly, it is possible that, while forming the ballot lists, a party leader considers the politicians whose skills are more productive for legislative activities for the positions that have greater chances of winning.

Table 4: Maximum likelihood estimates of the structural parameters

Parameters			
η_2	0.99 (0.004)***		
σ_ϵ	2.80 (0.25)***		
Characteristics			
Constant	9,98 (0.29)***	College	-0.48 (0.23)**
Female	0.38 (0.27)	No occupation	-2.76 (0.60)***
Retired	0.34 (0.61)	Governance	-4.12 (1.36)***
Business	-0.56 (0.25)**	Life, physical, and social sciences	-1.50 (0.43)***
Other	0.17 (0.96)	Community and social service	-0.17 (0.53)
Production	1.81 (0.39)***	Legal occupations	-2.02 (0.45)***
Construction and extraction	-1.09 (0.58)*	Education, training, and library	-0.68 (0.32)**
Farming, fishing, and forestry	-0.16 (0.48)	Arts, sports, and media	-0.15 (0.41)
Office	-1.01 (1.02)	Architecture and engineering	-1.89 (0.41)***
Healthcare	-2.66 (0.54)***	Business and financial operations	-0.36 (0.35)
Management	-1.69 (0.47)***		
$\frac{1}{N}\sum_i logL_i$	-0.6628		

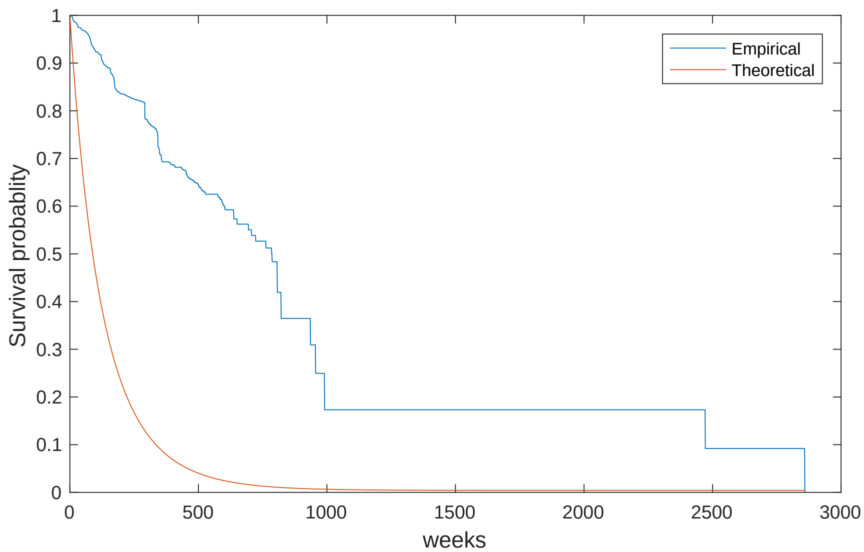
Notes: Standard errors are in parentheses. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

7.1 Model fit

In this section, I compare the frequency of party switching implied by the model to the observed frequency of party switching. Figure 5 plots the theoretical and empirical survival functions for the duration of a politician's party membership spell. In the figure, the empirical survival function is estimated by modifying the Turnbull estimator to account for the right- and interval-censored observations (Turnbull 1976, Klein and Melvin 2005). The theoretical survival function, on the other hand, is plotted after substituting the estimated labor market transition parameters into the likelihood of observing a given membership duration (equation D.1).

The figure shows that the model significantly overestimates the rate at which politicians switch their parties. The higher inertia in data can be due to at least three phenomena that are not accounted for in the model. First, politicians who approve their parties' ideological positions may be less interested in switching to parties that provide greater political benefits. Accordingly, taking into account the ideological match between politicians and party leaders would reduce the frequency of party-switching implied by the model. Second, elected

Figure 5: Survival function



politicians and politicians who never gain seats in parliament may have different offer arrival rates. This is because gaining public recognition may increase an MP's chances of meeting with other parties' leaders. Allowing the two groups of politicians to have different rates of offer arrival would allow the model to explain the electorally unsuccessful politicians' long party-membership durations. Third, parties that win seats in parliament or form the government coalition may produce additional club goods that have not been taken into account in the model. Allowing for additional sources of club goods production would explain the long membership spells of the electorally successful parties' members.

7.2 Specification tests

In this section, I present the results of two specification tests. First, I conduct a Wald test for the hypothesis that the contribution of all observed characteristics of a politician to his political assets is zero. This test yields a p-value less than 10^{-5}, and, hence, I conclude that the coefficient estimates of the politician characteristics are jointly statistically significantly different from zero.

Second, I conduct a Wald test for the hypothesis that a party that forms a majority government cannot produce additional club goods. To do this, I reestimate the model by allowing the party that formed a majority government during 2002-2014 to have additional club goods production.[34] Formally, if a type-z politician joins the governing party x_M with

[34]Note that, in the new estimation sample, I only included the last three electoral terms in my dataset, during which a party formed a majority government consecutively. Nevertheless, the restricted model's

share ϕ, then the politician's payoff from this membership is given by

$$u(z, \phi, x_M) = \frac{z\phi\theta(x_M)}{x_M} + \psi(x_M) + x_M^{\alpha}, \tag{7.1}$$

where x_M^{α} is the club goods that arise from party control over government functions. Testing $H_0 : \alpha = 0$ yields a p-value of 0.74. Thus, I fail to reject that party control over government functions do not generate additional club goods. Note that some of these additional club goods may be associated with winning seats in parliament. If this is true, then other parties that win seats in parliament should also be allowed to produce additional club goods. However, I cannot test this hypothesis because the model is identified only when at least 3 electoral terms are used in the estimation sample and there are no three consecutive terms in which the same subset of parties won seats in parliament.

8 Counterfactual analysis

In this section, I compare a candidate-centered system to a party-centered system. In section 8.1, I adjust the model in section 3 to a candidate-centered system and analyze its equilibrium properties. In section 8.2, I discuss estimation of the equilibrium of a candidate-centered system. In section 8.3, I derive the equilibriums of both a party-centered and a candidate-centered system when a politician has more bargaining power during membership negotiations. In section 8.4, I compare the expected rent share of a politician across two types of systems.

8.1 Equilibrium of a candidate-centered system

The model in section 3 studies the equilibrium of a party-centered democracy, where all members of a party combine their resources to collaboratively produce the party's rents. In a candidate-centered democracy, on the other hand, each politician produces rents in a party with more independence. In this section, I adjust the model in section 3 to a candidate-centered system and study its equilibrium properties. To do this, I preserve the structure of the model in section 3 but change its rent production process by assuming that each politician produces the same amount of rents independent of his membership status. If a type-z politician meets a type-x leader by the random matching process, they bargain over a share ϕ over the politician's rent production, $\theta(z)$. If the politician joins the party, he also utilizes the party's club goods. I assume that the club goods production function, $\psi(\cdot)$, is

estimates are fairly similar for the 1995-2014 and 2002-2014 terms.

the same across different systems.[35] Accordingly, if the politician joins the party with share ϕ, the politician's and the leader's payoffs are

$$u(z, \phi, x) = \phi\theta(z) + \psi(x)$$

and

$$w(z, \phi, x) = (1 - \phi)\theta(z),$$

respectively.

In a stationary equilibrium, a party leader seeks to fill her party with members who would join the party with the smallest rent shares. A capacity-constrained leader's rent maximization problem is the same as that of a leader in a party-centered democracy (equation 3.4). Accordingly, a leader maximizes her lifetime utility by following the stationary decision rules in equation 3.1.

I conjecture that, in the equilibrium of a candidate-centered system, the upper bound of the rent share a leader pays to the members is nonincreasing in party size, i.e., $\frac{d\phi^{l^*}(z,x)}{dx} \leq 0$. This implies that the value a politician receives from membership in a party has two components with different returns to party size. While the club goods are increasing in party size, the maximum amount of rents he can earn in a party is nonincreasing in party size. Accordingly, a politician has the same stationary decision rules for joining a party from the pool of independents and for switching a party in both systems (equations 3.5 and 3.6).

When a type-z politician joins a type-x party from the pool of independent politicians, the rent share he earns in the party is

$$\begin{aligned}
\phi^l(z,&x, x', \phi^{l^*}(z, x), \phi^{l^*}(z, x')) \\
&= [\rho V_0(z) - \psi(x)]\frac{1}{\theta(z)} \\
&- \frac{\lambda}{\theta(z)} \int_{q_b(\cdot)}^{x_b(\cdot)} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm} \bar{F}(m)dm \\
&+ \frac{\lambda}{\theta(z)} \int_{x_a(\cdot)}^{q_a(\cdot)} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm} F(m)dm, \qquad (8.1)
\end{aligned}$$

which is derived following CPR (see Appendix A for the derivation of the share equation for a party-centered system). Similarly, the rent share a type-(z, x') politician earns in a type-x

[35] The nonpecuniary benefits of affiliating with a party may change when a country adopts a different rent accumulation technology. For example, when politicians behave more independently in a party, the party leader may not be able to protect its members as strongly as she can when the party acts as a team. However, my results can be interpreted as the short-term effects of the institutional change

party is

$$
\begin{aligned}
\phi^{l}(z,&x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')) \\
&= \phi^{l^*}(z,x') + [\psi(x')-\psi(x)]\frac{1}{\theta(z)} \\
&- \frac{\lambda}{\theta(z)}\int_{q_b(\cdot)}^{x_b(\cdot)} \frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}\bar{F}(m)dm \\
&+ \frac{\lambda}{\theta(z)}\int_{x_a(\cdot)}^{q_a(\cdot)} \frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}F(m)dm.
\end{aligned}
\tag{8.2}
$$

Similar to the party-centered system, a type-z politician's profitability to a type-x leader, given in equation 3.23, is increasing in the upper bound of the rent share the leader pays to the politician's type, $\phi^{l^*}(z,x)$. Hence, just as in the party-centered system where a leader is willing to negotiate a politician's rents up to the match productivity, the Nash equilibrium of a candidate-centered system is characterized by the leader paying $\phi^{l^*}(z,x) = 1$ to all members of her party and not making acceptable offers to other types of politicians (see the proof of Theorem 1, which shows the existence of a similar equilibrium for the party-centered system).

Although a leader is willing to negotiate a politician's rents up to the politician's entire rent productivity in both types of systems, there are some different equilibrium features across these two systems as well. In a party-centered system, the politicians are heterogeneous in their value-ranking of the parties because a party's value to a politician has two components with different returns to party size. Bigger parties have more club goods, but a politician's rent productivity is decreasing in party size. Since richer politicians have to give up too much rents to join a big party, they prefer the smaller parties. In the Nash equilibrium of a candidate-centered system, on the other hand, the maximum amount of rents a politician can earn in a party is independent of the party size. As bigger parties provide more club goods, all politicians rank the parties vertically. Accordingly, the politicians switch only to the bigger parties as in CPR. Despite this similarity, the equilibrium of a candidate-centered system still differs from CPR because, while a party leader sorts members in my paper, a firm makes an acceptable offer to any worker it matches with in CPR.

8.2 Estimation

The online appendix shows that, under mild assumptions, a party leader in a party-centered democracy fills her party up to her capacity. In the Nash equilibrium of a candidate-centered democracy, a leader still benefits from filling her party up to her capacity. Since party leaders' capacities are exogenously determined, during the counterfactual analysis, one can use the estimated party sizes from the party-centered model.

Although the party leaders have the same stationary decision rules in both models, the range of politicians to whom a leader makes acceptable offers is likely to differ across the two types of systems. This is because the rent share that convinces a politician to join a party, and, in turn, a politician's profitability to a party leader is different across these systems. However, as all politicians rank the parties vertically, each politician has the same hazard rate of leaving a party. The estimation strategy in Section 6 uses the heterogeneity in politicians' conditional hazard rates, which identifies the structural parameters. Since politicians' conditional hazard rates are constant, it is not possible to undertake the duration analysis in a finite mixture model framework to estimate the politician types in a candidate-centered system. The heterogeneity in politicians' assets translates into their rent shares, which is not observable. So, it is not possible to estimate the politician types without having more information.

However, it is possible to compare the expected rents a given type of politician earns in a party across the two types of systems. As long as there is a subset of politician-types that exist in a party in both types of systems, this study would be informative about how politicians split a party's rents across different systems. One can implement this counterfactual analysis by computing the equilibrium rents a given type of politician earns in a party under both systems (equations 3.15 and 8.2) using the estimates of the party-centered system.

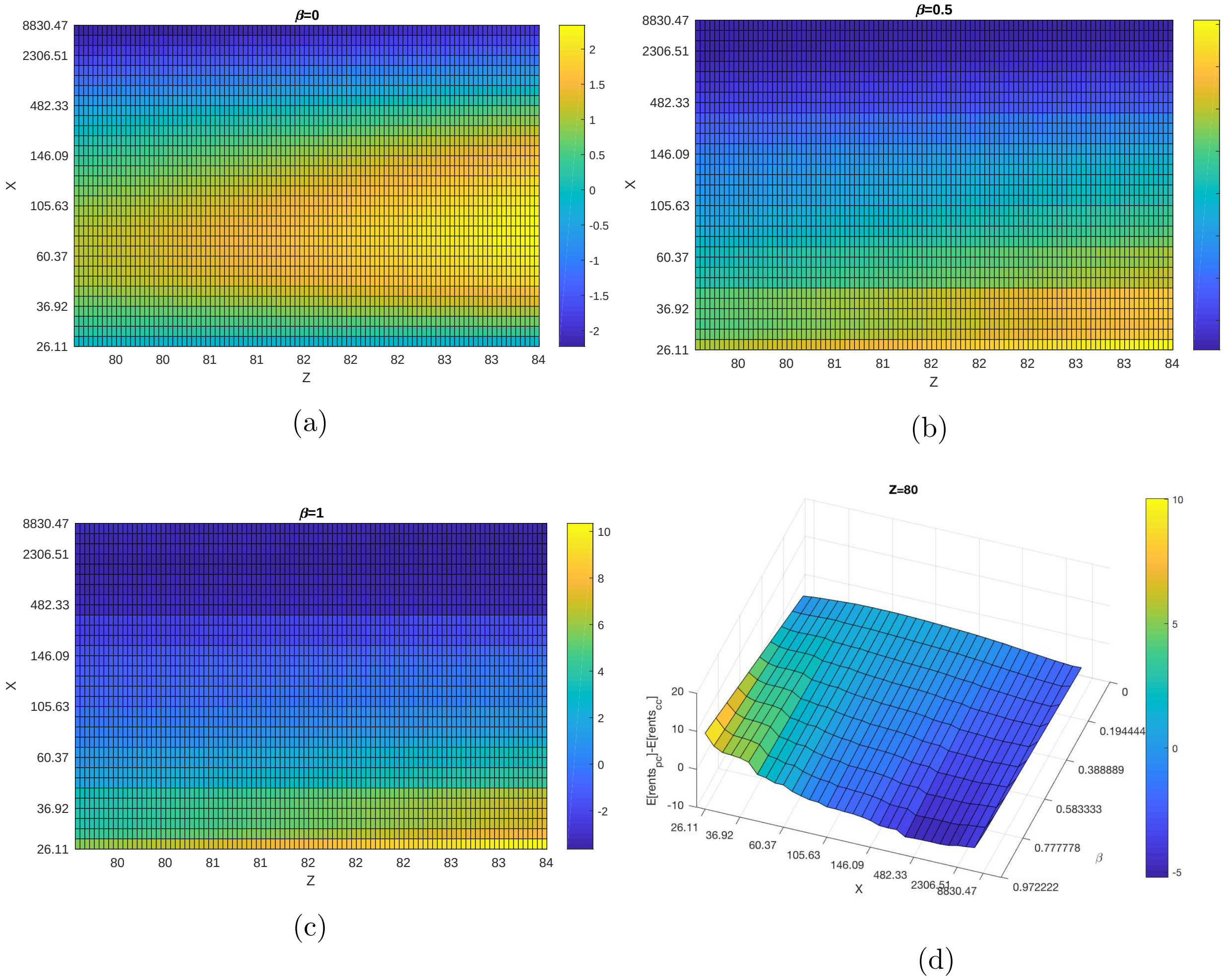


Figure 6: Comparison of the rents of a low-type politician across party-centered and candidate-centered systems

8.3 Mixed systems

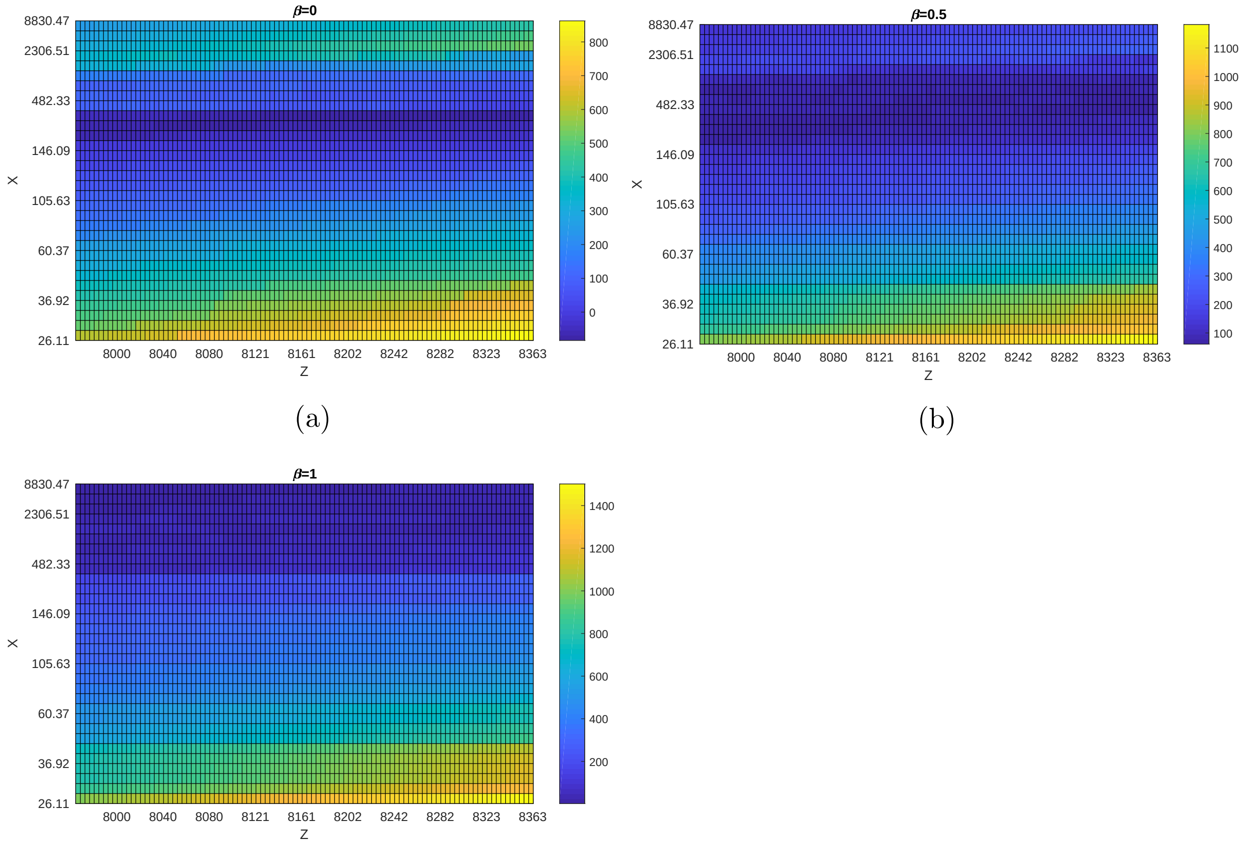
In both open and closed-list proportional representation systems, leaders select candidates and rank them in order of priority for winning seats. These two systems differ in the degree to which voters can influence the winning chances of the candidates. In the closed-list systems of Argentina, Israel, Italy, and Turkey, a voter can vote only for the party as a whole. In the open-list systems of Belgium, Finland, the Netherlands, and Sweden, among others, on the other hand, voters can also vote for their preferred candidates, and a candidate may take priority over the party's other candidates who are listed more highly if he gets sufficient preference votes. Accordingly, while a leader has monopsonistic power for recruiting members in a closed-list system, politicians have more bargaining power in open-list systems. So, an open-list system can be studied by assuming that a politician gets a share $\beta \in [0, 1]$ of his rent production in the party, where β is referred to as the politician's bargaining power as in CPR. The equilibrium of an open-list system can be derived by assuming that a type-x party leader's take-it-or-leave-it offer to a type-z politician is

$$V(z, \phi, 1, x) = min\{(1 - \beta)V(z, \phi^p, 1, x) + \beta(V(z, 1, 1, x), V(z, 1, 1, x)\} \tag{8.3}$$

where $\phi^p \leq 1$ is the minimum rent share that can convince the politician to join the party and $V(z, 1, 1, x)$ is the value that the politician receives in the party when he consumes his entire rent production. The equilibrium rent share of a politician in this system is derived by following the same steps as in a closed-list system.

8.4 Comparison of different systems

In this section, I compare the rents a politician earns in a party across candidate-centered and party-centered systems. The difference between the rents a politician earns in a party across different systems is determined by three elements of the model. First, a politician's rent productivity in a party is decreasing (constant) in party size in a party (candidate)-centered system. This implies that, a politician is more productive in small (large) enough parties in a party (candidate)-centered system. If that were all that mattered, politicians in small (large) parties would earn more (less) rents in a party (candidate)-centered system. However, the productivity difference across the two types of systems translates into a difference in the option value of party membership. A politician is willing to forgo today's rents in expectation of earning higher rents in the future that would arise from receiving lucrative outside offers. Depending on the difference in the values of the parties that would improve the politician's rents, the option value effect may be either smaller or bigger in a candidate-centered system. Hence, the effect of the option value on the difference between the rents of a politician across different systems can be either positive or negative. Third, the difference in party values across the two systems translates into a difference in the value of a politicians' outside option. If the parties that a politician ranks more lower than his party have a higher value in a party-centered system, then he would earn higher rents in a party-centered system.



(a) (b)

(c)

Figure 7: Comparison of the rents of a medium-type politician across party-centered and candidate-centered systems

Since low, medium, and high types of politicians differ in their ranking of the parties, in what follows, I analyze the difference between a politician's rents across the two political systems for each subgroup of politicians. Figure 6 shows the difference of the expected rents a low-type politician earns between party-centered and candidate-centered systems. Panel (a) shows that, when the politician has no bargaining power, he earns more rents in smaller (bigger) parties in a party (candidate)-centered system. This is because a low-type politician ranks the parties vertically in both systems and he is more productive in smaller (bigger) parties in a party (candidate)-centered system. Moreover, the expected rent difference is hump-shaped, which reflects the differences in the expected value of a politician's outside option as well as the option value of party membership across the two systems. In medium-sized parties, there are more politicians with better outside options, compared to, say, the first party in which all low-type members' outside option is to be an independent. However, the option value of membership in a candidate-centered system is bigger in larger parties where the politician is more productive. As a result, the expected rent difference increases in party size as long as the effect of the outside option dominates the effect of the option value. Panels (b) and (c) show the expected rent difference across the two systems when the politician has more bargaining power during the membership negotiations. As the politician's bargaining power increases, the effect of the outside option begins to disappear, since the politician is able to extract more out of the match surplus independent of his outside option. Accordingly, the rents the politician earns in a party-centered democracy in excess of the rents he earns in a candidate-centered democracy decrease monotonically in party size.

The expected rent difference of a medium-type politician across the two systems, as shown in Figure 7, is also determined by the differences in productivities and party values. However, since a medium-type politician's ranking of the party values differs across the two systems, the expected rent difference across the two systems is not monotone in party size. Panel (a) shows that, when the politician has no bargaining power, on average, the rents he earns in a party-centered system in excess of the rents he earns in a candidate-centered system are decreasing in party size from the smallest party to his lowest-ranked party and increasing afterwards. This is because, on average, the value of a politician's outside option decreases in party size over the first part of his U-shaped returns to party size and increases afterwards. When the politician is able to extract the entire match surplus, as in Panel (c), the effect of the outside option begins to disappear and the difference between the rents the politician earns across the two systems decreases monotonically.

For a high-type politician, the effects of outside options, option value, and productivity work in the same direction and make the smaller parties more attractive to the politician. As shown in Figure 8, a politician prefers the smaller parties.

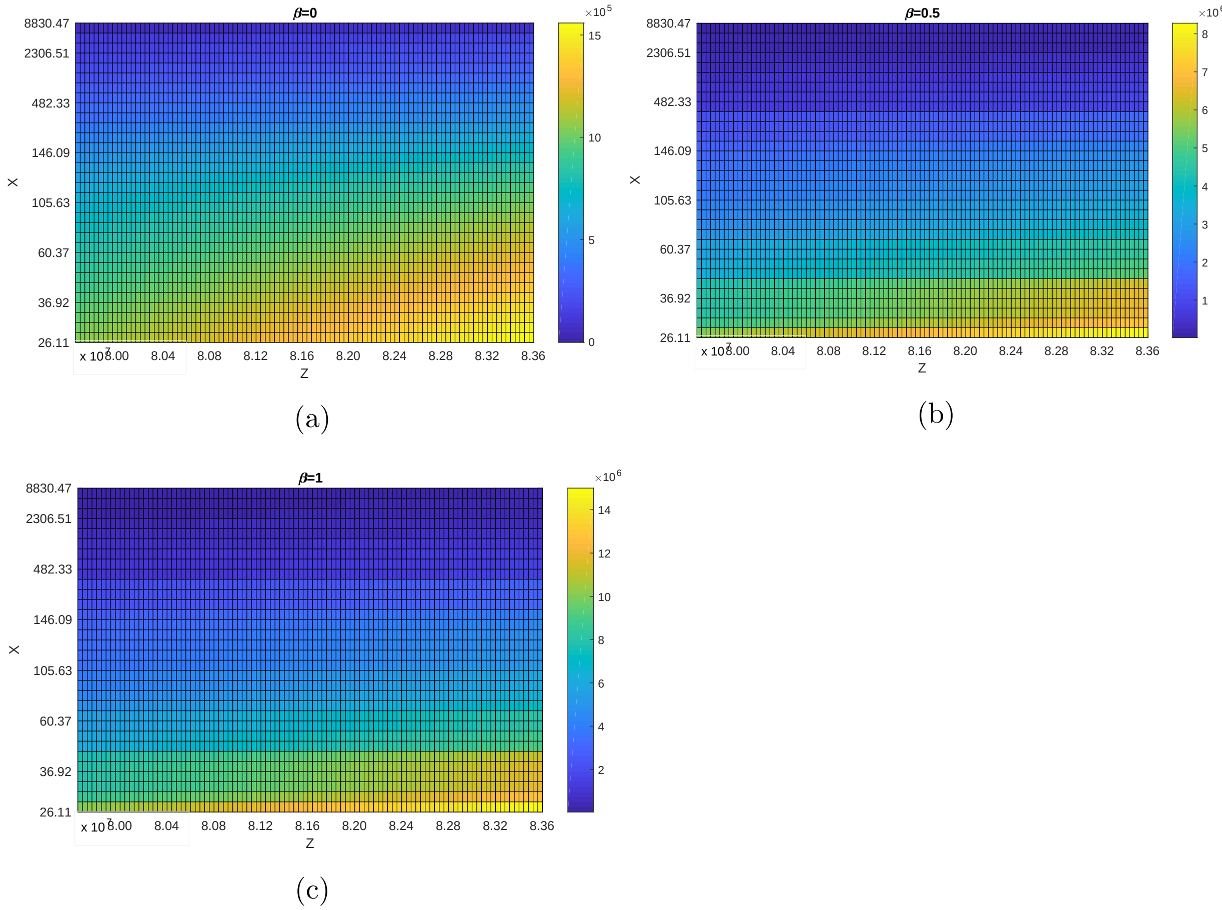


Figure 8: Comparison of the rents of a high-type politician across party-centered and candidate-centered systems

9 Other policies

In parliamentary democracies, the functioning of separation of powers is vulnerable to strong party leaders. This is because when a strong party produces large enough votes to form a majority government, both the legislative and the executive branches would be ruled by the same party. Moreover, the theory in section 3 shows that the existence of valuable club goods allows the party's monopsonistic leader to have a large control over her party. All of these suggest that having democratic elections does not necessarily prevent concentration of political power in a few hands. In this section, I draw several insights from the theoretical model to discuss the institutions that provide a broad distribution of political power.

First, limiting club goods production would require a party leader to pay greater rents to party members, which improves the within-party power distribution. Policies that target the limitation of club goods should better understand the nature of the club goods produced in a country. Club goods, which are nonexclusively provided to party members, may be justifiable, such as the prestige or the contentment of playing an active role in policy making. For example, if voters care about the identities of parties, belonging to a party that has a strong electoral support would provide nonpecuniary benefits to party members (see Ma and McLaren (2018) for the effects of local partisanship on electoral competition). If parties' club goods are of the justifiable type, then reducing political polarization would limit club goods production. However, club goods can also have an unjustifiable nature. For example, parties may give priority to their supporters for accessing the services provided by the government functions they control. If parties' club goods are mainly of the unjustifiable type, then strict enforcement of the rule of law would limit club goods production.

Second, depending on the distribution of the heterogeneous amounts of politicians' and leaders' political skills, imposing deterrent party-switching costs may either improve or worsen the power distribution in political parties. According to the model, imposing a party-switching ban has two opposing effects on a politician's rent share in a party. First, the option-value effect disappears, which increases the politician's rent share. Second, party competition over politicians' services vanishes, which decreases the average rent share of a given politician type in a party. The overall effect on the power distribution in parliament depends on the level of match frictions as well as the distribution of the politicians' political assets and the leaders' party-leading abilities. If the politicians in a country are rich in political assets, then the party leaders need to pay them large rents to convince them to join their parties. In a highly frictional political arena where politicians are rich in their political assets, imposing party-switching costs can improve the power distribution in parties. On the other hand, if the politicians are not rich in their political assets and they can easily meet with outside parties, then a ban on party switching would distort the power distribution in parliament in favor of the party leaders.

Third, reducing the match frictions would allow politicians to improve their outside options and get more out of their rent production in a party. One way to reduce match frictions can be to allow the voters to show their preferences for candidates in a primary election, which would increase politicians' recognizability by the party leaders.

Note that the policies that are discussed here do not characterize the complete set of policies that can affect the power distribution in parties. Designing party-subsidy mechanisms, imposing term limits, and developing many other types of institutions can be effective in improving the power distribution in parties.

10 Conclusion

In this paper, I develop an equilibrium model of team production in a labor search environment. Team production is characteristic of many industries, including the high-tech industry, academia, the healthcare provision industry, and the political arena. The workers' career choices and the distribution of production surplus in each of these industries depend on the characteristics of the industry as well as the features of team production. In each of these industries, a smaller team produces a lower amount of output but gives its members the opportunity to be more influential by using their skills more productively. In the high-tech industry, for example, we observe skilled engineers leaving giant tech companies to establish their start-ups, where the more productive use of their skills translates into higher earnings. In academia, we frequently observe established professors moving to smaller institutions where their skills play transformative roles. Similarly, in the political arena of a parliamentary democracy, we observefm politicians who switch to smaller parties and gain more say in party politics due to the smaller parties' greater needs for their political assets.

The effects of the distribution of production surplus in political arena extend beyond the agents in the political arena. This is because the political arena is a market for producing political power, and all social and economic policies in a country are influenced by the politicians who hold the political power. Parliamentary systems are especially vulnerable to strong party leaders. In closed-list systems of Argentina, Italy, Israel, Spain, and Turkey, for example, party leaders select candidates in the ballot lists and voters can vote only for a party. Hence, gaining influential positions in politics requires a party leader's approval. Moreover, party control over government functions in these countries generates valuable club goods, which increases the value of party membership. In this environment, politicians may relegate the use of their political power to party authorities when membership is more valuable than acting independently. This implies that the political power, and, in turn, the determination of all social and economic policies, may be left to a few strong party leaders instead of a broader set of representatives.

To understand how the political power is distributed in a political party, I adjust my model of team production to a party-centered parliamentary system. In my model, all politicians aggregate their resources in a party to produce party's rents and club goods. All members of a party non-exclusively benefit from their parties' club goods. In return, a politician is willing to share his private rent production with the party leader. The outcome of rent sharing depends on the politician's outside option. Unlike the standard search models, membership in smaller parties can be a better outside option for some politicians. This is because a smaller party's need for political assets is higher. As a result, smaller parties are willing to pay higher rents to politicians. This feature of the model allows me to estimate the model even without observing the outcome of rent sharing.

I structurally estimate my model for Turkey with a dataset I constructed of 33 parties, 2,000 politicians who gained seats in parliament, and 35,000 politicians who were on party ballot lists between 1995 and 2014. My model matches the high party-switching rate (28.5%) that is characteristic of many parliamentary democracies. I find that Turkish parties produce club goods more easily then rents, which leads to ever stronger party control by the leaders.

In a counterfactual analysis, I find that members of smaller (bigger) parties are more powerful in party-centered (candidate-centered) systems. This finding provides an explanation for the existence of small parties that continue to participate in elections despite having no chances of winning seats in parliament. My counterfactual results also provide an explanation for the existence of only two big parties that compete in the elections in the candidate-centered system of the United States.

Finally, my model provides several insights on the types of institutions that yield to a broader distribution of power in the political arena. Limiting club goods production, through either reduction of political polarization or strict enforcement of the rule of law, would yield a broader distribution of political power in a political party. Similarly, reducing the match frictions by allowing the voters to choose the candidates in the ballot lists would improve the distribution of political power.

11

References

[1] Acemoglu, Daron, and James Robinson. Why nations fail: The origins of power, prosperity, and poverty. Crown Business, 2012.

[2] Aizawa, Naoki, and Hanming Fang. "Equilibrium Labor Market Search and Health Insurance Reform, Second Version." (2015).

[3] Aldashev, Gani. "Voter turnout and political rents." Journal of Public Economic Theory 17.4 (2015): 528-552.

[4] Aldrich, John H., and William T. Bianco. "A game-theoretic model of party affiliation of candidates and office holders." Mathematical and Computer Modelling 16.8 (1992): 103-116.

[5] Anderson, Simon P., Andre De Palma, and Jacques François Thisse. Discrete choice theory of product differentiation. MIT press, 1992.

[6] Anderson, Simon P., and Kieron Meagher. "Choosing a champion: Party membership and policy platform." (2012).

[7] Ashworth, Scott, and Ethan Bueno De Mesquita. "Elections with platform and valence competition." Games and Economic Behavior 67.1 (2009): 191-216.

[8] Bajari, Patrick, et al. "Estimating static models of strategic interactions." Journal of Business Economic Statistics 28.4 (2010): 469-482.

[9] Banerjee, Abhijit V. "A simple model of herd behavior." The Quarterly Journal of Economics (1992): 797-817.

[10] Becker, Gary S. "A theory of marriage: Part I." The Journal of Political Economy (1973): 813-846.

[11] Berkovec, James, and Steven Stern. "Job exit behavior of older men." Econometrica: Journal of the Econometric Society (1991): 189-210.

[12] Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile prices in market equilibrium." Econometrica: Journal of the Econometric Society (1995): 841-890.

[13] Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. "A theory of fads, fashion, custom, and cultural change as informational cascades." Journal of political Economy (1992): 992-1026.

[14] Bonhomme, Stephane, and Gregory Jolivet. "The pervasive absence of compensating differentials." Journal of Applied Econometrics 24.5 (2009): 763-795.

[15] Buchanan, James M. "An economic theory of clubs." Economica 32.125 (1965): 1-14.

[16] Burdett, Ken, and Melvyn G. Coles. "Marriage and class." The Quarterly Journal of Economics (1997): 141-168.

[17] Burdett, Kenneth, and Dale T. Mortensen. "Wage differentials, employer size, and unemployment." International Economic Review (1998): 257-273.

[18] Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. "Wage bargaining with on-the-job search: Theory and evidence." Econometrica 74.2 (2006): 323-364.

[19] Caselli, Francesco, and Massimo Morelli. "Bad politicians." Journal of Public Economics 88.3 (2004): 759-782.

[20] Chen, Daniel L. "Club goods and group identity: Evidence from Islamic resurgence during the Indonesian financial crisis." Journal of Political Economy 118.2 (2010): 300.

[21] Çelen, Boğaçhan, and Shachar Kariv. "Distinguishing informational cascades from herd behavior in the laboratory." The American Economic Review 94.3 (2004): 484-498.

[22] Dagsvik, John K. "Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes." Econometrica: Journal of the Econometric Society (1994): 1179-1205.

[23] De Paula, Aureo. "Inference in a synchronization game with social interactions." Journal of Econometrics 148.1 (2009): 56-71.

[24] Desposato, Scott W. "Parties for rent? Ambition, ideology, and party switching in Brazil's chamber of deputies." American Journal of Political Science 50.1 (2006): 62-80.

[25] Dey, Matthew S., and Christopher J. Flinn. "An equilibrium model of health insurance provision and wage determination." Econometrica 73.2 (2005): 571-627.

[26] Dey, Matthew, and Christopher Flinn. "Household search and health insurance coverage." Journal of Econometrics 145.1 (2008): 43-63.

[27] Diermeier, Daniel, Hlya Eraslan, and Antonio Merlo. "A structural model of government formation." Econometrica 71.1 (2003): 27-70.

[28] Diermeier, Daniel, Michael Keane, and Antonio Merlo. "A political economy model of congressional careers." The American Economic Review 95.1 (2005): 347-373.

[29] Downs, Anthony. "An economic theory of political action in a democracy." The journal of political economy (1957): 135-150.

[30] Eckstein, Zvi, and Gerard J. Van den Berg. "Empirical labor search: A survey." Journal of Econometrics 136.2 (2007): 531-564.

[31] Eeckhout, Jan, and Philipp Kircher. "Identifying sorting-in theory." The Review of Economic Studies 78.3 (2011): 872-906.

[32] Enelow, James M., and Melvin J. Hinich. “Nonspatial candidate characteristics and electoral competition.” The Journal of Politics 44.01 (1982): 115-130.

[33] Evdokimov, Kirill. “Nonparametric identification of a nonlinear panel model with application to duration analysis with multiple spells.” Department of Economics, Princeton University (2011).

[34] Feddersen, Timothy J., Itai Sened, and Stephen G. Wright. “Rational voting and candidate entry under plurality rule.” American Journal of Political Science (1990): 1005-1016.

[35] Golder, Matt. “Democratic electoral systems around the world, 1946-2000.” Electoral Studies 24.1 (2005): 103-121.

[36] Grofman, Bernard. “Comparisons among electoral systems: Distinguishing between localism and candidate-centered politics.” electoral Studies 24.4 (2005): 735-740.

[37] Gronberg, Timothy J., and W. Robert Reed. “Estimating workers’ marginal willingness to pay for job attributes using duration data.” Journal of Human Resources (1994): 911-931.

[38] Hotelling, Harold. “Stability in Competition.” The Economic Journal 39.153 (1929): 41-57.

[39] Hotz, V. Joseph, and Robert A. Miller. “Conditional choice probabilities and the estimation of dynamic models.” The Review of Economic Studies 60.3 (1993): 497-529.

[40] Jolivet, Gregory, Fabien Postel-Vinay, and Jean-Marc Robin. “The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US.” European Economic Review 50.4 (2006): 877-907.

[41] Keane, Michael P., and Kenneth I. Wolpin. “The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence.” the Review of economics and statistics (1994): 648-672.

[42] Klein, John P., and Melvin L. Moeschberger. Survival analysis: techniques for censored and truncated data. Springer Science Business Media, 2005.

[43] Kunicova, Jana, and Susan Rose-Ackerman. “Electoral rules and constitutional structures as constraints on corruption.” British Journal of Political Science 35.04 (2005): 573-606.

[44] Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. "Matching, sorting and wages." Review of Economic Dynamics 19 (2016): 63-87.

[45] Lopes de Melo, Rafael. "Firm wage differentials and labor market sorting: Reconciling theory and evidence." Unpublished Manuscript (2013).

[46] Lopes de Melo, Rafael. "How Firms Affect Wages: a Structural Decomposition." Unpublished Manuscript (2015).

[47] Malmberg, Hannes. "Random Choice over a Continuous Set of Options." (2013).

[48] Ma, Xiangjun, and John McLaren. A swing-state theorem, with evidence. No. w24425. National Bureau of Economic Research, 2018.

[49] McKelvey, Richard D., and John W. Patty. "A theory of voting in large elections." Games and Economic Behavior 57.1 (2006): 155-180.

[50] Messner, Matthias, and Mattias K. Polborn. "Paying politicians." Journal of Public Economics 88.12 (2004): 2423-2445.

[51] Nickell, Stephen. "Estimating the probability of leaving unemployment." Econometrica: Journal of the Econometric Society (1979): 1249-1266.

[52] Peter, Coughlin. "Chapter Twenty-Six-Probabilistic and Spatial Models of Voting." Handbook of Social Choice and Welfare 2 (2011): 833-896.

[53] Polo, Michele. "Electoral Competition and Political Rents." SSRN Working Paper Series (1999).

[54] Ridder, Geert, and Gerard J. Berg. "Measuring Labor Market Frictions: A Cross-Country Comparison." Journal of the European Economic Association 1.1 (2003): 224-244.

[55] ROGERSON, RICHARD, ROBERT SHIMER, and RANDALL WRIGHT. "Search-Theoretic Models of the Labor Market: A Survey." Journal of Economic Literature 43.4 (2005): 959-988.

[56] Rosen, Sherwin. "The theory of equalizing differences." Handbook of labor economics 1 (1986): 641-692.

[57] Roth, Alvin E. Axiomatic models of bargaining. Vol. 170. Springer Science Business Media, 2012.

[58] Roth, Alvin E., and Marilda Sotomayor. "Two-sided matching." Handbook of game theory with economic applications 1 (1992): 485-541.

[59] Scotchmer, Suzanne. "Local public goods and clubs." Handbook of public economics 4 (2002): 1997-2042.

[60] Schofield, Norman, and Itai Sened. Multiparty democracy: elections and legislative politics. Cambridge University Press, 2006.

[61] Shimer, Robert, and Lones Smith. "Assortative matching and search." Econometrica 68.2 (2000): 343-369.

[62] Smith, Lones, and Peter Sørensen. "Pathological outcomes of observational learning." Econometrica 68.2 (2000): 371-398.

[63] Sorkin, Isaac. "Ranking Firms Using Revealed Preference." Manuscript, University of Michigan October (2015).

[64] Turnbull, Bruce W. "The empirical distribution function with arbitrarily grouped, censored and truncated data." Journal of the Royal Statistical Society. Series B (Methodological) (1976): 290-295.

Appendix A The share equation

In this section, I derive the closed-form solution of the rent share a politician earns in a party for both cases of the politician's initial party membership status. The steps taken for deriving these equations is exactly the same as in CPR; however, the resulting share equation reflects two different aspects of the model. First, while the authors allow a worker to earn a positive share of the match surplus when s/he has bargaining power, in this paper, a politician's rent share is independent of his bargaining power. This is because a leader chooses her stationary decision rules to extract the entire match surplus in any match. Second, in CPR, a worker ranks the firms vertically, and, therefore, he switches only to the more productive firms. However, in this paper, the politicians are heterogeneous in their preference ordering of the party sizes. Accordingly, depending on the degree to which he values the rents over the club goods, a politician may switch to either a bigger or a smaller party.

Equation 3.11 in section 3.6 showed that, substituting a leader's stationary decision rules into the value function of a medium type-z politician who earns a rent share ϕ in a type-x

party, one gets

$$
\begin{aligned}
&[\rho+\delta+\lambda\bar{F}(q_b(\cdot))+\lambda F(q_a(\cdot))]V(z,\phi,\phi^{l^*}(z,x),x)\\
&\quad=\phi\frac{z}{x}\theta(x)+\psi(x)+\delta V_0(z)+\lambda[F(x_a(\cdot))+\bar{F}(x_b(\cdot))]V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x)\\
&\quad+\lambda\int_{q_b(\cdot)}^{x_b(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m)+\lambda\int_{x_a(\cdot)}^{q_a(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m).
\end{aligned}
\tag{A.1}
$$

The share equation is obtained in three steps. First, integration by parts in equation A.1 yields an expression for $V(z,\phi,\phi^{l^*}(z,x),x)$. Next, another representation for $V(z,\phi,\phi^{l^*}(z,x),x)$ is obtained using the leader's stationary decision rules defined in equations 3.2 and 3.3. Finally, equating these two expressions yields the closed-form solution of the share equation.

Using integration by parts, the integral terms in equation A.1 can be written as

$$
\begin{aligned}
&\int_{q_b(\cdot)}^{x_b(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m)\\
&\qquad\qquad=-\bar{F}(x_b(\cdot))V(z,\phi^{l^*}(z,x_b(\cdot)),\phi^{l^*}(z,x_b(\cdot)),x_b(\cdot))\\
&\qquad\qquad+\bar{F}(q_b(\cdot)V(z,\phi^{l^*}(z,q_b(\cdot)),\phi^{l^*}(z,q_b(\cdot)),q_b(\cdot))\\
&\qquad\qquad+\int_{q_b(\cdot)}^{x_b(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}\bar{F}(m)dm,
\end{aligned}
$$

$$
\begin{aligned}
&\int_{x_a(\cdot)}^{q_a(\cdot)}V(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)dF(m)\\
&\qquad\qquad=F(q_a(\cdot))V(z,\phi^{l^*}(z,q_a(\cdot)),\phi^{l^*}(z,q_a(\cdot)),q_a(\cdot))\\
&\qquad\qquad-F(x_a(\cdot)V(z,\phi^{l^*}(z,x_a(\cdot)),\phi^{l^*}(z,x_a(\cdot)),x_a(\cdot))\\
&\qquad\qquad-\int_{x_a(\cdot)}^{q_a(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}F(m)dm.
\end{aligned}
$$

Accordingly, equation A.1 can be rewritten as

$$
\begin{aligned}
(\rho+\delta)V(z,\phi,\phi^{l^*}(z,x),x) &= \phi\frac{z}{x}\theta(x)+\psi(x)+\delta V_0(z) \\
&+ \lambda[F(x_a(\cdot))+\bar{F}(x_b(\cdot))]V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x) \\
&- \lambda[\bar{F}(q_b(\cdot))+F(q_a(\cdot))]V(z,\phi,\phi^{l^*}(z,x),x) \\
&- \lambda\bar{F}(x_b(\cdot))V(z,\phi^{l^*}(z,x_b(\cdot),\phi^{l^*}(z,x_b(\cdot),x_b(\cdot)) \\
&+ \lambda\bar{F}(q_b(\cdot)V(z,\phi^{l^*}(z,q_b(z\cdot)),\phi^{l^*}(z,q_b(z\cdot)),q_b(\cdot)) \\
&+ \lambda F(q_a(\cdot))V(z,\phi^{l^*}(z,q_a(\cdot)),\phi^{l^*}(z,q_a(\cdot)),q_a(\cdot)) \\
&- \lambda F(x_a(\cdot)V(z,\phi^{l^*}(z,x_a(\cdot)),\phi^{l^*}(z,x_a(\cdot)),x_a(\cdot)) \\
&+ \lambda\int_{q_b(\cdot)}^{x_b(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dx'}\bar{F}(m)dm \\
&- \lambda\int_{x_a(\cdot)}^{q_a(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}F(m)dm. \quad \text{(A.2)}
\end{aligned}
$$

Recall that the threshold party types for having a share improvement in the party, $q_b(\cdot)$ and $q_a(\cdot)$, solve

$$V(z,\phi^{l^*}(z,q_a(\cdot)),\phi^{l^*}(z,q_a(\cdot)),q_a(\cdot)) = V(z,\phi^{l^*}(z,q_b(\cdot)),\phi^{l^*}(z,q_b(\cdot)),q_b(\cdot)) = V(z,\phi,\phi^{l^*}(z,x),x),$$

and hence, the third line in equation A.2 cancels out with the fifth and the sixth lines. Similarly, since

$$
\begin{aligned}
V(z,\phi^{l^*}(z,x),\phi^{l^*}(z,x),x) &= V(z,\phi^{l^*}(z,x_a(\cdot)),\phi^{l^*}(z,x_a(\cdot)),x_a(\cdot)) \\
&= V(z,\phi^{l^*}(z,x_b(\cdot)),\phi^{l^*}(z,x_b(\cdot)),x_b(\cdot)),
\end{aligned}
$$

the second line cancels out with the fourth and the seventh lines. Simplifying these terms, one gets

$$
\begin{aligned}
(\rho+\delta)V(z,\phi,\phi^{l^*}(z,x),x) &= \phi\frac{z}{x}\theta(x)+\psi(x)+\delta V_0(z) \\
&+ \lambda\int_{q_b(\cdot)}^{x_b(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}\bar{F}(m)dm \\
&- \lambda\int_{x_a(\cdot)}^{q_a(\cdot)}\frac{dV(z,\phi^{l^*}(z,dm),\phi^{l^*}(z,m),dm)}{dm}F(m)dm. \quad \text{(A.3)}
\end{aligned}
$$

Recall that the maximum value a medium-type politician can obtain in a party is decreasing (increasing) in party size on $[x_{min},x_0(z))$ ($(x_0(z),x^{max}]$), where $x_0(z)$ denotes the lowest point of his U-shaped returns to party size. Accordingly, the integral terms on the right hand side of equation A.3 reflect the positive contribution of the possibility of gaining future share improvements in the party to membership value.

Now, suppose that the politician's outside option is a type-x' party. Following the leader's stationary decision rule in equation 3.2, one has

$$(\rho+\delta)V(z,\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')),x) = (\rho+\delta)V(z,\phi^{l^*}(z,x'),\phi^{l^*}(z,x'),x'),$$

which, using equation 3.12, can be rewritten as

$$(\rho+\delta)V(z,\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')),x) = \phi^{l^*}(z,x')\frac{z}{x'}\theta(x') + \psi(x') + \delta V_0(z). \quad \text{(A.4)}$$

Equation A.4 is a formal statement of the Bertrand competition of two parties. The higher-ranked party, x, wins the politician, and pays a rent share $\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x'))$ that provides the politician the same membership value as the the highest value he could have received in the lower-ranked party.

Equating the right-hand-side of equation A.4 with that of equation A.3 gives the equilibrium share $\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x'))$ that convinces the politician to join a type-x party when his outside option is membership in a type-x' party,

$$\begin{aligned}\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x')) \;&=\; \phi^{l^*}(z,x')\frac{x}{x'}\frac{\theta(x')}{\theta(x)} + [\psi(x')-\psi(x)]\frac{x}{z\theta(x)} \\ &-\; \frac{x}{z\theta(x)}\lambda\int_{q_b(\cdot)}^{x_b(\cdot)} \underbrace{\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}}_{>0}\bar{F}(m)dm \\ &+\; \frac{x}{z\theta(x)}\lambda\int_{x_a(\cdot)}^{q_a(\cdot)} \underbrace{\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}}_{<0}F(m)dm.\end{aligned} \quad \text{(A.5)}$$

Equation A.5 shows that, the utility flow to a type-(z,x') politician in a type-x party, $\frac{z\phi^l(z,x,x',\phi^{l^*}(z,x),\phi^{l^*}(z,x'))\theta(x)}{x}+\psi(x)$, is less than the utility flow the politician would have received if he had chosen his outside option, $\frac{z\phi^{l^*}(z,x')\theta(x')}{x'} + \psi(x')$, as the two integral terms on the right-hand-side are negative. This reflects an option value effect. The politician is willing to give up today's rents in expectation of future rent increases.

If the politician is a low type, the exact share equation still obtains with $x_a(\cdot) = x_{min}$. Similarly, if the politician is a high type, the share equation is found after substituting

$x_b(\cdot) = x^{max}$ in equation A.5.

When a type-z politician meets a type-x party from the pool of independents, the leader's stationary decision rule is to offer the politician the rent share $\phi^l(z, x, 0, \phi^{l^*}(z, x), 0)$. Notice that, when the politician's outside option is being an independent, the threshold party types that induce a share improvement in the party are given by $q_a(\cdot) = x_{a0}(\cdot)$ and $q_b(\cdot) = x_{b0}(\cdot)$. Accordingly, an expression for $V(z, \phi^l(z, x, 0, \phi^{l^*}(z, x), 0), \phi^{l^*}(z, x), x)$ can be obtained from equation A.3 as

$$\begin{aligned}(\rho+\delta)V(z, \phi^l(z, x, 0, \phi^{l^*}(z, x), 0), \phi^{l^*}(z, x), x) &= \phi^l(z, x, 0, \phi^{l^*}(z, x), 0)\frac{z}{x}\theta(x) + \psi(x) + \delta V_0(z) \\ &+ \lambda \int_{x_{b0}(\cdot)}^{x_b(\cdot)} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm} \bar{F}(m)dm \\ &- \lambda \int_{x_a(\cdot)}^{x_{a0}(\cdot)} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), dm)}{dm} F(m)dm.\end{aligned} \tag{A.6}$$

Moreover, following the leader's take-it-or-leave-it offer rule in equation 3.3, one has

$$(\rho+\delta)V(z, \phi^l(z, x, 0, \phi^{l^*}(z, x), 0), \phi^{l^*}(z, x), x) = (\rho+\delta)V_0(z). \tag{A.7}$$

Equating the right-hand-side of equation A.6 with that of A.7, one solves for the equilibrium share $\phi^l(z, x, 0, \phi^{l^*}(z, x), 0)$,

$$\begin{aligned}\phi^l(z, x, 0, \phi^{l^*}(z, x), 0) &= [\rho V_0(z) - \psi(x)]\frac{x}{z\theta(x)} \\ &- \frac{x}{z\theta(x)}\lambda \int_{x_{b0}(\cdot)}^{x} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm} \bar{F}(m)dm \\ &+ \frac{x}{z\theta(x)}\lambda \int_{x_a(\cdot)}^{x_{a0}(\cdot)} \frac{dV(z, \phi^{l^*}(z, m), \phi^{l^*}(z, m), m)}{dm} F(m)dm.\end{aligned} \tag{A.8}$$

Equation A.8 holds for all independent politicians. However, when a politician's value of being an independent is greater than the the minimum value he receives at the lowest point of his U-shaped returns to party size, i.e., when $V_0(z) > V(z, \phi^{l^*}(z, x_0(z)), \phi^{l^*}(z, x_0(z)), x_0(z))$, there is an equivalent expression. Suppose that $x_{a0}(\cdot)$ and $x_{b0}(\cdot)$ are the small and big party types that give the politician the same value as being an independent when the politician is paid the maximum share in those parties, i.e., $V_0(z) = V(z, \phi^{l^*}(z, x_{a0}(\cdot)), \phi^{l^*}(z, x_{a0}(\cdot)), x_{a0}(\cdot)) = V(z, \phi^{l^*}(z, x_{b0}(\cdot)), \phi^{l^*}(z, x_{b0}(\cdot)), x_{b0}(\cdot)) > V(z, \phi^{l^*}(z, x_0(z)), \phi^{l^*}(z, x_0(z)), x_0(z))$. The equivalent expression for $\phi^l(z,x,0,\phi^{l^*}(z,x),0)$ can be obtained by substituting $x' = x_{b0}(\cdot)$ in equation A.5,

$$\begin{aligned}\phi^l(z,x,0,\phi^{l^*}(z,x),0) &= \phi^{l^*}(z,x_{b0}(\cdot))\frac{x}{x_{b0}(\cdot)}\frac{\theta(x_{b0}(\cdot)}{\theta(x)} + [\psi(x_{b0}(\cdot)) - \psi(x)]\frac{x}{z\theta(x)} \\ &- \frac{x}{z\theta(x)}\lambda\int_{x_{b0}(\cdot)}^{x_b(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}\bar{F}(m)dm \\ &+ \frac{x}{z\theta(x)}\lambda\int_{x_a(\cdot)}^{x_{a0}(\cdot)}\frac{dV(z,\phi^{l^*}(z,m),\phi^{l^*}(z,m),m)}{dm}F(m)dm. \quad \text{(A.9)}\end{aligned}$$

Appendix B Steady-state flow equalities

This section derives the steady-state flow equalities by adjusting the steps taken in CPR for the possibility of a U-shaped returns to party size.

- The proportion of independent politicians φ is constant. However, as the parties differ in the range of the politician types to whom they make an acceptable offer, the proportion of independent politicians differ by the politician type. Let φ_z denote the proportion of type-z independent politicians. The flows into the stocks of independent type-z politicians are due to exogenous match break-ups, which occur at rate $M\ell(z)(1-\varphi_z)\delta$. The outflows from the stocks of independent type-z politicians occur as they get an acceptable offer, which occurs at rate $M\ell(z)\varphi_z\lambda[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]$. In a steady-state, the flows into and outflows from the stocks of independent politicians are equal, which gives the proportion of type-z independent politicians,

$$\varphi_z = \frac{\delta}{\delta + \lambda[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]}. \quad \text{(B.1)}$$

- The density of a politician type in a given type of party is constant. This density can be found from the steady-state flow equality of the politicians of a certain type entering into and leaving a type of party. Consider a medium type-z politician. Suppose that $x > x_0(z)$, i.e., the politician considers a type-x party as a big party. The outflows from the stocks $\Gamma_{\phi|z,x}(\phi|z,x)g(z,x|\Phi^{l^*}(z,x))M(1-\varphi_z)$ of politicians of type-z, member of parties of type-x, and paid less than $\phi \in [\phi(z,x,0,\phi^{l^*}(z,x),0), \phi^{l^*}(x)]$ leave this category in either of two ways. First, the match exogenously breaks up at rate δ. Second, they receive an offer from a party of type $x' \in [x_{min}, q_a(\cdot)] \cup [q_b(\cdot), x^{max}]$ that either causes a share improvement or induces them to leave their current party, which occurs at rate $\lambda[F(q_a(\cdot) + \bar{F}(q_b(\cdot)]$. The politicians enter this category either by switching from parties of type-$x' \in [q_a(\cdot), q_b(\cdot)]$ or from the pool of independents. The steady-state equality between flows into and outflows from the stocks $\Gamma_{\phi|z,x}(\phi|z,x)g(z,x|\Phi^{l^*}(z,x))M(1-\varphi_z)$ is

$$\begin{aligned}&[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot)]]M(1-\varphi_z)\Gamma_{\phi|z,x}(\phi|z,x)g(z,x|\Phi^{l^*}(z,x))\\ &\quad = \lambda M\varphi_z\ell(z)f(x) + \lambda f(x)M(1-\varphi_z)\int_{q_a(\cdot)}^{q_b(\cdot)} g(z,m|\Phi^{l^*}(z,x))dm\end{aligned}$$

This expression can be simplified by substituting $\lambda\varphi_z = \frac{\delta(1-\varphi_z)}{[F(x_{a0}(z))+\bar{F}(x_{b0}(\cdot))]}$ and simplifying the $(1-\varphi_z)M$ terms,

$$\begin{aligned}&[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]]\Gamma_{\phi|z,x}(\phi|z,x)g(z,x|\Phi^{l^*}(z,x))\\ &\quad = \frac{\delta}{[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]}\ell(z)f(x) + \lambda f(x)\int_{q_a(\cdot)}^{q_b(\cdot)} g(z,m|\Phi^{l^*}(z,x))dm, \qquad \text{(B.2)}\end{aligned}$$

Evaluating equation B.2 at $\phi = \phi^{l^*}(z,x)$, (which has the property that $\Gamma_{\phi|z,x}(\phi^{l^*}(z,x)|z,x) = 1$, $q_b(\cdot) = x$, and $q_a(\cdot) = x_a(\cdot)$), one gets

$$\begin{aligned}&[\delta + \lambda[F(x_a(\cdot)) + \bar{F}(x)]]\Gamma_{\phi|z,x}(\phi|z,x)g(z,x|\Phi^{l^*}(z,x))\\ &\quad = \frac{\delta}{[F(x_{a0}(\cdot)) + \bar{F}_{b0}(\cdot)]}\ell(z)f(x) + \lambda f(x)\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm,\end{aligned}$$

which can be rearranged as

$$\begin{aligned}\frac{\delta\ell(z)f(x)}{[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]} &= [\delta + \lambda[F(x_a(\cdot)) + \bar{F}(x)]]g(z,x|\Phi^{l^*}(z,x))\\ &- \lambda f(x)\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm. \qquad \text{(B.3)}\end{aligned}$$

Notice that the right-hand-side of this equality equals

$$\begin{aligned}&\frac{d}{dx}\Big\{[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))]]\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm\Big\}\\ &+ \frac{\partial x_a(\cdot)}{\partial x}\Big\{[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))]]g(z,x_a(\cdot)|\Phi^{l^*}(z,x))\\ &- \lambda f(x_a(\cdot))\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm\Big\},\end{aligned}$$

where the second term is equal to $\frac{\partial x_a(\cdot)}{\partial x}\frac{\delta\ell(z)f(x_a(\cdot))}{[F(x_{a0}(\cdot))+\bar{F}(x_{b0}(\cdot))]}$. Therefore, one can rewrite equation B.3 as

$$\frac{\delta\ell(z)}{[F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))]}\Big[f(x) - \frac{\partial x_a(\cdot)}{\partial x}f(x_a(\cdot))\Big]$$
$$= \frac{d}{dx}\Big\{[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))]]\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm\Big\},$$

which, after integrating both sides over x, becomes

$$\delta\ell(z)\Big[\frac{F(x) - F(x_a(\cdot))}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))}\Big]$$
$$= [\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))]\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm,$$

which can be rearranged to get

$$\int_{x_a(\cdot)}^{x} g(z,m|\Phi^{l^*}(z,x))dm = \frac{\delta[F(x) - F(x_a(\cdot))]}{(\delta + \lambda[\bar{F}(x) + F(x_a(\cdot)))}\frac{\ell(z)}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))}. \qquad \text{(B.4)}$$

Differentiating both sides of equation B.4 with respect to x, one gets

$$g(z,x|\Phi^{l^*}(z,x)) - \frac{\partial x_a(\cdot)}{\partial x}g(z,x_a(\cdot)|\Phi^{l^*}(z,x))$$
$$= \frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\Big]^2}\frac{\ell(z)f(x)}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))}$$
$$- \frac{\partial x_a(\cdot)}{x}\frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\Big]^2}\frac{\ell(z)f(x_a(\cdot))}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))}.$$

Then, the joint density of type-z politicians in parties of types x and $x_a(\cdot)$ are given by

$$\begin{aligned} g(z,x|\Phi^{l^*}(z,x)) &= \frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\Big]^2}\frac{\ell(z)f(x)}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))} \\ &= \frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\Big]^2}\tilde{\ell}(z)f(x), \end{aligned} \qquad \text{(B.5)}$$

and

$$g(z, x_a(\cdot)|\Phi^{l^*}(z,x)) = \frac{\delta(\delta+\lambda)}{\left[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\right]^2} \frac{\ell(z) f(x_a(\cdot))}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))}$$
$$= \frac{\delta(\delta+\lambda)}{\left[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\right]^2} \tilde{\ell}(z) f(x_a(\cdot)), \tag{B.6}$$

respectively, where

$$\tilde{\ell}(z) = \frac{\ell(z)}{F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))} \tag{B.7}$$

is defined as the *effective density* of type-z politicians, as it weights the politician's density by its demand by the parties. Note that the joint density of a politician in a party decrease in both the politician's probability of getting an acceptable offer conditional on getting an offer, $F(x_{a0}(\cdot)) + \bar{F}(x_{b0}(\cdot))$, and the probability of getting an offer from a party that the politician ranks better than types-x and $x_a(\cdot)$ parties, $\lambda[\bar{F}(x) + F(x_a(\cdot))]$, due to increased competition by the parties. Note also that, the conditional density of type-z politicians in type-x and type-$x_a(\cdot)$ parties are equal,

$$g(z|x, \Phi^{l^*}(z,x)) = g(z|x_a(\cdot), \Phi^{l^*}(z,x)) = \frac{\delta(\delta+\lambda)}{\left[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\right]^2} \tilde{\ell}(z),$$

as a medium type-z politician ranks the parties of types x and $x_a(\cdot)$ of equal value. Note that, the low (high) politician types switch only to the bigger (smaller) parties. Accordingly, the joint density of a low type-z politician in a type-x party is

$$g(z, x|\Phi^{l^*}(z,x)) = \frac{\delta(\delta+\lambda)}{[\delta + \lambda\bar{F}(x))]^2} \tilde{\ell}(z) f(x), \tag{B.8}$$

Similarly, the joint density of a high type-z politician in a type-x party is

$$g(z, x|\Phi^{l^*}(z,x)) = \frac{\delta(\delta+\lambda)}{[\delta + \lambda F(x)]^2} \tilde{\ell}(z) f(x). \tag{B.9}$$

Note that, equation B.8 is identical to its counterpart in CPR when $\bar{F}(x_{b0}(\cdot)) = 1$. This is because the low politician types rank the parties vertically, similar to the workers in CPR.

- The density of type-(z, x') politicians in type-x parties is constant. Consider a medium type-(z, x') politician in a type-x party. Suppose that both x and x' are "big" parties for the politician. Note that the politician's thresholds for switching to another party and having a share improvement in the party are $x_a(\cdot)$,$x_b(\cdot)$ and $q_a(\cdot)$,$q_b(\cdot)$, respectively. Moreover, when member of a big party, he ranks all bigger parties better, thus $x_b(\cdot) = x$. Similarly, when his outside option is a big party, an offer from a party that is bigger than his outside option and smaller than his bigger-party switching threshold cause a share improvement in the party, and, hence, $q_b(\cdot) = x'$. The outflows from the stocks of type-$(z, q_b(\cdot))$ politicians, member of parties of type-x, and paid $\phi^l(z, x, q_b(\cdot), \phi^{l^*}(z, x), \phi^{l^*}(z, q_b(\cdot)))$ leave this category in either of two ways. First, the match exogenously breaks up at rate δ. Second, they get an offer from a party of type $x'' \in \{[x_{min}, q_a(\cdot)] \cup [q_b(\cdot), x^{max}]\}$ that either causes a share improvement or induces them to leave their party, which occurs at rate $\lambda[F(q_a(\cdot)) + \bar{F}(q_b(\cdot)]$. The politicians enter this category in either of two ways. First, they switch from parties of type-$x'' \in \{q_a(\cdot), q_b(\cdot)\}$. Second, if they were already a member of a type-x party and had a worse outside option than $q_b(\cdot)$, they get an offer from an outside party of type-$x' \in \{q_a(\cdot), q_b(\cdot)\}$. Then, the steady-state equality between flows into and outflows from the stocks $M(1-\varphi_z)\mu_{z,q_b(\cdot),x}(z, q_b(\cdot), x|\Phi^{l^*}(z, x))$ is

$$\begin{aligned}
&[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot)]]M(1-\varphi_z)\mu_{z,q_b(\cdot),x}(z, q_b(\cdot), x|\Phi^{l^*}(z, x)) \\
&\quad = \lambda M(1-\varphi_z) f(x) g(z, q_b(\cdot)|\Phi^{l^*}(z, x)) \\
&\quad + \lambda M(1-\varphi_z) f(q_b(\cdot)) \int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x)) dm
\end{aligned} \tag{B.10}$$

which, after simplifying the $(1-\varphi_z)M$ terms can be rearranged as

$$\begin{aligned}
&\lambda f(x) g(z, q_b(\cdot)|\Phi^{l^*}(z, x)) = [\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot)]]\mu_{z,q_b(\cdot),x}(z, q_b(\cdot), x|\Phi^{l^*}(z, x)) \\
&\quad - \lambda f(q_b(\cdot)) \int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x, \Phi^{l^*}(z, x)) dm.
\end{aligned} \tag{B.11}$$

The right hand side of this equality can be rewritten as

$$\begin{aligned}
&\frac{d}{dq_b(\cdot)}\{[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot)]] \int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x)) dm\} \\
&+ \frac{dq_a(\cdot)}{dq_b(\cdot)}\{-\lambda f(q_a(\cdot)) \int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x)) dm \\
&+ [\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot)]]\mu_{z,q_a(\cdot),x}(z, q_a(\cdot), x|\Phi^{l^*}(z, x))\},
\end{aligned} \tag{B.12}$$

where the second term equals $\frac{dq_a(\cdot)}{dq_b(\cdot)}\lambda f(x)g(z, q_b(\cdot)|\Phi^{l^*}(z, x))$. Then,

$$\begin{aligned}
&\lambda f(x)[g(z, q_b(\cdot)|\Phi^{l^*}(z, x)) - \frac{dq_a(\cdot)}{dq_b(\cdot)}g(z, q_a(\cdot)|\Phi^{l^*}(z, x))] \\
&= \frac{d}{dx'}\{[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]][\int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x))dm\}
\end{aligned} \tag{B.13}$$

Substituting equations B.5 and B.6 into equation B.15,

$$\begin{aligned}
&\lambda f(x)\tilde{\ell}(z)\delta(\delta+\lambda)\frac{[f(q_b(\cdot)) - \frac{dq_a(\cdot)}{dq_b(\cdot)}f(q_a(\cdot))]}{\Big[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))\Big]^2} \\
&= \frac{d}{dq_b(\cdot)}\{[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a())]][\int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x))dm\},
\end{aligned}$$

and integrating both sides with respect to x', we have

$$\begin{aligned}
&f(x)\tilde{\ell}(z)\frac{\delta(\delta+\lambda)}{\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]} \\
&= [\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]][\int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x))dm,
\end{aligned}$$

which can be rearranged to get

$$\int_{q_a(\cdot)}^{q_b(\cdot)} \mu_{z,m,x}(z, m, x|\Phi^{l^*}(z, x))dm = \frac{\delta(\delta+\lambda)f(x)\tilde{\ell}(z)}{\Big[[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]]\Big]^2} \tag{B.14}$$

Differentiating both sides of this equation with respect to $q_b(\cdot)$, one gets

$$\begin{aligned}
&\mu_{z,q_b(\cdot),x}(z, q_b(\cdot), x|\Phi^{l^*}(z, x)) - \frac{dq_a(\cdot)}{dq_b(\cdot)}\mu_{z,q_a(\cdot),x}(z, q_a(\cdot), x|\Phi^{l^*}(z, x)) \\
&\qquad = -2\frac{\delta(\delta+\lambda)\lambda f(x)\tilde{\ell}(z)\Big(- f(q_b(\cdot)) + \frac{dq_a(\cdot)}{dq_b(\cdot)}f(q_a(\cdot))\Big)}{\Big[[\delta + \lambda[\bar{F}(q_b(\cdot)) + F(q_a(\cdot))]]\Big]^3}
\end{aligned}$$

and, hence

$$\mu_{z,q_b(\cdot),x}(z, q_b(\cdot), x|\Phi^{l^*}(z, x)) = 2\frac{\delta(\delta+\lambda)\lambda f(x)\tilde{\ell}(z)f(q_b(\cdot))}{[\delta+\lambda[\bar{F}(q_b(\cdot))+F(q_a(\cdot))]]^3}, \tag{B.15}$$

and

$$\mu_{z,q_a(\cdot),x}(z, q_a(\cdot), x|\Phi^{l^*}(z, x)) = -2\frac{\delta(\delta+\lambda)\lambda f(x)\tilde{\ell}(z)f(q_a(\cdot))}{\Big[[\delta+\lambda[\bar{F}(q_b(\cdot))+F(q_a(\cdot))]]\Big]^3}. \tag{B.16}$$

- The density of type-$(z, 0)$ politicians in type-x parties is constant. The flows into this category occurs as a type-x leader meets a type-z independent politician at rate λ. The outflows occur either through an exogenous match break up, occurring at rate δ, or when a politician gets an offer from a party of type $x' \in \{[x_{min}, x_{a0}(\cdot) \cup [x_{b0}(\cdot), x^{max}]\}$ that either induces the politician to switch the party or improves his outside option in the party. The steady-state equality of the flows into and the outflows from the stocks $M(1-\varphi)\mu_{z,0,x}(z, 0, x|\Phi^{l^*}(z, x))$ is

$$\varphi M\lambda\ell(z)f(x) = M(1-\varphi)\mu_{z,0,x}(z, 0, x|\Phi^{l^*}(z, x))[\delta+\lambda[F(x_{a0}(z)+\bar{F}(x_{b0}(z))]]$$

which, after simplifying the M term and imposing $\lambda\varphi_z = \frac{\delta(1-\varphi_z)}{[F(x_{a0}(z))+\bar{F}(x_{b0}(z))]}$ becomes

$$\mu_{z,0,x}(z, 0, x|\Phi^{l^*}(z, x)) = \frac{\delta}{[\delta+\lambda[F(x_{a0}(z)+\bar{F}(x_{b0}(z))]]}\tilde{\ell}(z)f(x) \tag{B.17}$$

- The within-party share distribution of the politicians, $\Gamma_{\phi|z,x}(\phi|z, x)$, is constant. This distribution can be found by substituting equations B.4 and B.5 into equation B.2,

$$\Gamma_{\phi|z,x}(\phi|z, x) = \Big(\frac{\delta+\lambda[\bar{F}(x_b(\cdot))+F(x_a(\cdot))]}{\delta+\lambda[\bar{F}(q_b(\cdot))+F(q_a(\cdot))]}\Big)^2, \tag{B.18}$$

and the joint density of type-z politicians in type-x parties with a share less than $\phi \in [\phi(z, x, 0, \phi^{l^*}(z, x), 0), \phi^{l^*}(x)]$ is

$$\begin{aligned}\Gamma_{\phi,z,x}(\phi, z, x) &= \Gamma_{\phi|z,x}(\phi|z, x)g(z, x)\\ &= \frac{\delta(\delta+\lambda)}{(\delta+\lambda[\bar{F}(q_b(\cdot)+F(q_a(\cdot))])^2}\tilde{\ell}(z)f(x).\end{aligned} \tag{B.19}$$

- Let $z_0(x)$ denote the medium-type politician for whom the lowest point of the U-shaped returns to party size is x. Accordingly, all $z < z_0(x)$ consider a type-x party as a big party, while all $z > z_0(x)$ consider it as small. In a steady state, a party's size is equal to the sum of its members' resources,

$$\begin{aligned} x &= \int_0^{z^{max}} z g(z, x|\Phi^{l^*}(x)) dz \\ &= \int_0^{z_0(x)} z \frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[\bar{F}(x) + F(x_a(\cdot))\Big]^2} \tilde{\ell}(z) dz \\ &+ \int_{z_0(x)}^{z^{max}} z \frac{\delta(\delta+\lambda)}{\Big[\delta + \lambda[F(x) + \bar{F}(x_b(\cdot))\Big]^2} \tilde{\ell}(z) dz. \end{aligned} \tag{B.20}$$

Appendix C The unconditional likelihood of a party affiliation duration

This section derives the unconditional likelihood of observing a party affiliation duration following Ridder and van den Berg (2003) and CPR. Let d_n, d_r, and d_i denote the indicator functions for the uncensored, right-censored, and interval-censored observations, respectively. I begin by deriving the likelihood contribution of the uncensored observations, and then derive the contributions of the censored observations. Since the low, medium, and high politician types follow different decision rules for switching a party, the likelihood function takes the probability of the politician belonging to a particular type into account. Formally, the unconditional likelihood of a membership duration of t for an uncensored observation is

$$\begin{aligned} p(t|d_n = 1) &= L(\underline{z})p(t|z \leq \underline{z}, d_n = 1) + (L(\bar{z}) - L(\underline{z}))p(t|z \in \{\underline{z}, \bar{z}\}, d_n = 1) \\ &+ (1 - L(\bar{z}))p(t|z \geq \bar{z}, d_n = 1), \end{aligned} \tag{C.1}$$

where $\underline{z}$ and $\bar{z}$ are the threshold politician types that separate the low and the high types from the medium types of politicians, respectively.

Since all party transition processes are Poisson, all corresponding durations are exponentially distributed. The rate at which a low-type politician leaves a type-x party is $\delta[1+\kappa\bar{F}(x)]$. Accordingly, the density of a membership duration of t in a type-x party for a low type-z politician is

$$p(t|z \leq \underline{z}, x, d_n = 1) = \delta[1 + \kappa\bar{F}(x)]e^{-\delta[1+\kappa\bar{F}(x)]t}. \tag{C.2}$$

I treat the party type as unobserved heterogeneity and integrate equation C.2 over the density of the party types, $g(x|z,\Phi^{l^*}(z,x)) = \frac{1+\kappa}{[1+\kappa\bar{F}(x)]^2}f(x)$, which was derived in equation 3.17. Accordingly, the likelihood of observing a party affiliation duration of t for a low-type politician is

$$\begin{aligned} p(t|z \leq \underline{z}, d_n = 1) &= p(t|z \leq \underline{z}, x, d_n = 1)g(x|z,\Phi^{l^*}(z,x)) \\ &= \int_{x_{min}}^{x^{max}} \delta[1+\kappa\bar{F}(x)]e^{-\delta[1+\kappa\bar{F}(x)]t}\frac{1+\kappa}{[1+\kappa[\bar{F}(x)]^2}f(x)dx \\ &= \frac{\delta(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta a t}}{a}da, \end{aligned} \tag{C.3}$$

where $a = 1+\kappa\bar{F}(x)$ is the probability of leaving a type-x party as a fraction of the probability of having a need for new party membership, δ.

Similar to low-type politicians, the hazard of leaving a type-x party for a high type-z politician is $\delta[1+\kappa F(x)]$, and the joint density of high type-z politicians in type-x parties is $g(z,x|\Phi^{l^*}(z,x)) = \frac{1+\kappa}{[1+\kappa\bar{F}(x)]^2}\tilde{\ell}(z)f(x)$ (equation 3.17). Accordingly, the likelihood of observing a party affiliation duration of t for a high-type politician is

$$\begin{aligned} p(t|z \geq \bar{z}, d_n = 1) &= p(t|z \geq \bar{z}, x, d_n = 1)g(x|z,\Phi^{l^*}(z,x)) \\ &= \int_{x_{min}}^{x^{max}} \delta[1+\kappa F(x)]e^{-\delta[1+\kappa F(x)]t}\frac{1+\kappa}{[1+\kappa F(x)]^2}f(x)dx \\ &= \frac{\delta(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta a t}}{a}da, \end{aligned} \tag{C.4}$$

where $a = 1+\kappa F(x)$.

Recall that a medium type-z politician has a threshold party type $x_0(z)$ such that he considers all smaller parties than $x_0(z)$ as small, and the others as big. Due to the U-shaped returns to party size, he may consider two parties with different sizes of equal value. Accordingly, when a type-z politician is member of a small type-x party, he is better-off in all smaller parties than the current party and all parties that are larger than his bigger party-switching threshold, $x_b(z,x)$. Then, the hazard of leaving a small type-x party is $\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]$, and the joint density of medium type-z politicians in type-x parties is $g(z,x|\Phi^{l^*}(z,x)) = \frac{1+\kappa}{[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]^2}\tilde{\ell}(z)f(x)$ (equation 3.17). Similarly, when $x > x_0(z)$, the hazard of leaving a type-x party is $\delta[1+\kappa\bar{F}(x)+\kappa F(x_a(z,x))]$, and the joint density of medium type-z politicians in type-x parties is $g(z,x|\Phi^{l^*}(z,x)) = \frac{1+\kappa}{[1+\kappa\bar{F}(x_b(z,x))+\kappa F(x)]^2}\tilde{\ell}(z)f(x)$. Accordingly, the likelihood of observing a party affiliation duration of t for a medium-type politician is

$$
\begin{aligned}
p(t|z \in \{\underline{z}, \bar{z}\}, d_n = 1) &= p(t|z \in \{\underline{z}, \bar{z}\}, x, d_n = 1) g(x|z, \Phi^{l^*}(z, x)) \\
&= \int_{x_{min}}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]] e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t} \\
&\times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]]^2} f(x) dx \\
&- \int_{x^{max}}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x) + F(x_a(z, x))]] e^{-\delta[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]t} \\
&\times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x) + F(x_a(z, x))]]^2} f(x) dx. \qquad \text{(C.5)}
\end{aligned}
$$

Now, suppose that $x_b(z, x_{min}) < x^{max}$, i.e., no smaller party provides a greater value to the politician when he is a member of a type-$x_b(z, x_{min})$ party. Accordingly, the politician behaves like a low-type over the range $[x_b(z, x_{min}), x^{max}]$. Note that

$$
\begin{aligned}
&\int_{x_{min}}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]] e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t} \times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]]^2} f(x) dx \\
&- \int_{x_b(z,x_{min})}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x) + F(x_a(z, x))]] e^{-\delta[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]t} \times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x) + F(x_a(z, x))]]^2} f(x) dx \\
&= \int_{x_{min}}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]] e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t} \\
&\times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]]^2} [f(x) - f(x_b(z, x))] dx. \qquad \text{(C.6)}
\end{aligned}
$$

Substituting equation C.6 into equation C.5, one obtains

$$
\begin{aligned}
p(t|z \in \{\underline{z}, \bar{z}\}, d_n = 1) = &\int_{x_{min}}^{x_0(z)} \delta[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]] e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t} \\
&\times \frac{1 + \kappa}{[1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]]^2} [f(x) - f(x_b(z, x))] dx \\
&- \int_{x^{max}}^{x_b(z,x_{min})} \delta[1 + \kappa\bar{F}(x)] e^{-\delta[1+\kappa\bar{F}(x)]t} \times \frac{1 + \kappa}{[1 + \kappa\bar{F}(x)]^2} f(x) dx. \qquad \text{(C.7)}
\end{aligned}
$$

Applying change of variable in the first term with $a = 1 + \kappa[\bar{F}(x_b(z, x)) + F(x)]$, $da = \kappa[f(x) - f(x_b(z, x))\frac{dx_b(z,x)}{dx}]dx = \kappa[f(x)dx - f(x_b(z, x))dx_b(z, x)]$, and in the second term with

$a = 1 + \kappa\bar{F}(x), da = -\kappa f(x)dx$, one gets

$$\begin{aligned} p(t|z \in \{\underline{z}, \bar{z}\}, d_n = 1) &= \frac{\delta(1+\kappa)}{\kappa}\int_{1+\kappa\bar{F}(x_b(z,x_{min}))}^{1+\kappa}\frac{e^{-\delta at}}{a}da + \frac{\delta(1+\kappa)}{\kappa}\int_{1}^{1+\kappa\bar{F}(x_b(z,x_{min}))}\frac{e^{-\delta at}}{a}da \\ &= \frac{\delta(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta at}}{a}da. \end{aligned} \tag{C.8}$$

Finally, substituting equations C.3, C.4, and C.8 into equation C.1, the unconditional likelihood of a membership duration of t for an uncensored observation is

$$\begin{aligned} p(t|d_n = 1) &= L(\underline{z})p(t|z \leq \underline{z}, d_n = 1) + (L(\bar{z}) - L(\underline{z}))p(t|z \in \{\underline{z}, \bar{z}, d_n = 1\}) \\ &+ (1 - L(\bar{z})p(t|z \geq \bar{z}, d_n = 1) \\ &= \frac{\delta(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta at}}{a}da. \end{aligned} \tag{C.9}$$

There are three sources of right-censorship in data: death, the Constitutional Court banning the politician from affiliating with a political party (which is the case for only a few observations), and the politician being a member of a party in the last period of data. The likelihood contribution of a right-censored observation is the probability that the membership did not end until the censoring time. Suppose that the politician's true duration of membership in the party is T, but the data sample ends at $t < T$. Then, the probability that a low type-z politician's membership did not end by t is

$$\begin{aligned} p(T > t|z \leq \underline{z}, d_r = 1) &= p(T > t|z \leq \underline{z}, d_r = 1, x)g(x|z, \Phi^{l^*}(z,x)) \\ &= \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa\bar{F}(x)]t}\frac{1+\kappa}{[1+\kappa\bar{F}(x)]^2}f(x)dx \\ &= \frac{(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta at}}{a^2}da, \end{aligned} \tag{C.10}$$

where $a = 1 + \kappa\bar{F}(x)$.

Similar to the low types, the probability that a high type-z politician's membership did not end by t is

$$\begin{aligned} p(T > t|z > \bar{z}, d_r = 1) &= p(T > t|z > \bar{z}, d_r = 1, x)g(x|z, \Phi^{l^*}(z,x)) \\ &= \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa F(x)]t}\frac{1+\kappa}{[1+\kappa F(x)]^2}f(x)dx \\ &= \frac{(1+\kappa)}{\kappa}\int_{1}^{1+\kappa}\frac{e^{-\delta at}}{a^2}da, \end{aligned} \tag{C.11}$$

where $a = 1 + \kappa F(x)$.

When a medium type-z politician's membership did not end by t, the likelihood contribution of this observation is

$$
\begin{aligned}
p(T > t | z \in \{\underline{z}, \bar{z}\}, d_n = 1) &= p(T > t | z \in \{\underline{z}, \bar{z}\}, x, d_n = 1) g(x | z, \Phi^{l^*}(z, x)) \\
&= \int_{x_{min}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x_b(z,x)) + F(x)]]^2} f(x) dx \\
&- \int_{x^{max}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]t} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x) + F(x_a(z,x))]]^2} f(x) dx \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t}}{a^2} da, \qquad \text{(C.12)}
\end{aligned}
$$

where $a = 1 + \bar{F}(x_b(z,x)) + F(x)]]$ and the last equality follows from applying the operations defined by C.5-C.8. Then, the unconditional likelihood of membership duration of t for a right-censored observation is

$$
\begin{aligned}
p(t | d_r = 1) &= L(\underline{z}) p(t | z \leq \underline{z}, d_r = 1) + (L(\bar{z}) - L(\underline{z})) p(t | z \in \{\underline{z}, \bar{z}, d_r = 1\}) \\
&+ (1 - L(\bar{z}) p(t | z \geq \bar{z}, d_r = 1) \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t}}{a^2} da, \qquad \text{(C.13)}
\end{aligned}
$$

where the second equality substitutes equations C.10-C.12.

Interval censoring occurs when a member of a parliament loses an election, but reappears on the ballot lists of a different party in a consecutive election. The likelihood contribution of an interval-censored observation is the probability that the membership ended over the interval $T \in (t_1, t_2)$. Formally, the probability that a small type-z politician switches from a type-x party at $T \in (t_1, t_2)$, is

$$
\begin{aligned}
p(t_2 > T > t_1 | z \leq \underline{z}, d_i = 1) &= p(t_2 > T > t_1 | z \leq \underline{z}, x, d_i = 1) g(x | z, \Phi^{l^*}(z, x)) \\
&= \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa\bar{F}(x)]t_2} \frac{1+\kappa}{[1+\kappa\bar{F}(x)]^2} f(x) dx \\
&- \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa\bar{F}(x)]t_1} \frac{1+\kappa}{[1+\kappa\bar{F}(x)]^2} f(x) dx \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2}}{a^2} a \\
&- \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_1}}{a^2} da \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2} - e^{-\delta a t_1}}{a^2} da, \qquad \text{(C.14)}
\end{aligned}
$$

where $a = 1 + \kappa\bar{F}(x)$.

Similarly, the probability that a high type-z politician switches from a type-x party at $T \in (t_1, t_2)$, is

$$
\begin{aligned}
p(t_2 > T > t_1 | z \leq \underline{z}, d_i = 1) &= p(t_2 > T > t_1 | z \leq \underline{z}, x, d_i = 1) g(x | z, \Phi^{l^*}(z, x)) \\
&= \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa F(x)]t_2} \frac{1+\kappa}{[1+\kappa F(x)]^2} f(x) dx \\
&- \int_{x_{min}}^{x^{max}} e^{-\delta[1+\kappa F(x)]t_1} \frac{1+\kappa}{[1+\kappa F(x)]^2} f(x) dx \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2}}{a^2} da \\
&- \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_1}}{a^2} da \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2} - e^{-\delta a t_1}}{a^2} da, \qquad \text{(C.15)}
\end{aligned}
$$

where $a = 1 + \kappa F(x)$.

Lastly, the probability that a medium type-z politician switches from a type-x party at $T \in (t_1, t_2)$ is

$$
\begin{aligned}
p(t_2 > T > t_1 | z \in \{\underline{z}, \bar{z}\}, d_n = 1) &= p(t_2 > T > t_1 | z \in \{\underline{z}, \bar{z}\}, x, d_n = 1) g(x|z, \Phi^{l^*}(z,x)) \\
&= \int_{x_{min}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t_2} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]^2} f(x)dx \\
&- \int_{x^{max}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]t_2} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]^2} f(x)dx \\
&- \int_{x_{min}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]t_1} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x_b(z,x))+F(x)]]^2} f(x)dx \\
&+ \int_{x^{max}}^{x_0(z)} e^{-\delta[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]t_1} \\
&\times \frac{1+\kappa}{[1+\kappa[\bar{F}(x)+F(x_a(z,x))]]^2} f(x)dx \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2} - ^{-\delta a t_1}}{a^2} da, \qquad \text{(C.16)}
\end{aligned}
$$

where $a = 1 + \bar{F}(x_b(z,x)) + F(x)]]$ and the last equality follows from applying the operations defined by C.5-C.8. Then, the unconditional likelihood of membership duration of t for an interval-censored observation is

$$
\begin{aligned}
p(t|d_i = 1) &= L(\underline{z}) p(t|z \leq \underline{z}, d_i = 1) + (L(\bar{z}) - L(\underline{z})) p(t|z \in \{\underline{z}, \bar{z}, d_i = 1\}) \\
&+ (1 - L(\bar{z}) p(t|z \geq \bar{z}, d_i = 1) \\
&= \frac{(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2} - ^{-\delta a t_1}}{a^2} da, \qquad \text{(C.17)}
\end{aligned}
$$

where the second equality substitutes equations C.14-C.16. Accordingly, the unconditional likelihood of observing a membership duration of t is

$$
\begin{aligned}
p(t) &= p(t|d_n = 1)^{d_n} \times p(t|d_r = 1)^{d_r} \times p(t|d_i = 1)^{d_1} \\
&= \Big(\frac{\delta(1+\kappa)}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t}}{a} da\Big)^{d_n} \Big(\frac{1+\kappa}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t}}{a^2} da\Big)^{d_r} \Big(\frac{1+\kappa}{\kappa} \int_1^{1+\kappa} \frac{e^{-\delta a t_2} - e^{-\delta a t_1}}{a^2} da\Big)^{d_i}
\end{aligned}
$$
(C.18)

where the second equality subsitutes equations C.9, C.13, and C.17.

Appendix D Identification

In this section, I show that all structural parameters and functions except for the distribution of leaders' exogenous leading abilities are nonparametrically identified. Note that, I discuss identification of the model in the equilibrium, where the maximum rent share a leader pays to a politician, $\phi^{l^*}(z, x)$, is 1. Accordingly, the arguments of the party-switching threshold functions, $x_a(\cdot)$ and $x_b(\cdot)$, defined in equation 3.6, reduce to the types of a politician and his party, z and x. Similarly, the arguments of the functions defining the thresholds for joining a party from the pool of independent politicians, $x_{a0}(\cdot)$ and x_{b0}, defined in equation 3.5, reduce to the politician's type, z.

D.1 Transition parameters

In this section, I discuss identification of the parameters that characterize a politician's transition between parties and two party membership statuses using only the observed duration of party membership. To do so, I follow Ridder and van den Berg (2003), henceforth RB, and define $\kappa = \frac{\lambda}{\delta}$ as the average number of offers a politician receives between two spells of being an independent, where λ is the rate at which a politician gets an offer and δ is the rate at which a match exogenously breaks up. As κ determines the duration of job spells in conventional search-theoretic models of transitions in the labor market, RB show that it can be identified using only the duration data without using the wages. To do so, they treat the firm types as unobserved heterogeneity and estimate the unconditional distribution of the employment durations. They further show that this method of unconditional inference is robust to any specification of wage determination.[36] In Appendix C, I show that, although different politician types have different decision rules for party switching, the method of unconditional inference can be used to identify the transition parameters.

Intuitively, suppose that all politicians ranked the party values in the same order and the parties did not sort their members. Then, the random matching process implies that the initial party membership of the politicians is randomly distributed according to the distribution that the politicians sample offers from. RB note that, if they received an offer

[36] Many recent models of labor market equilibrium, including CPR, use unconditional inference to estimate job market frictions. Note that this method requires a time-invariant hazard function to achieve identification. For example, in a model of on-the-job search with two different sources of exogenous separation, Jolivet, Postel-Vinay, and Robin (2006) show that duration data do not contain the needed information to identify the job market transition parameters separately when there is duration dependence.

just after joining a party, the probability that a politician will accept a new offer is a half given both the original and new offers are from the same distribution. Accordingly, the average switching rate just after time zero is equal to $\delta + \frac{\kappa}{2}$, as a politician leaves a party either because of an exogenous shock or because he receives a new offer and accepts it. The longer membership durations largely belong to the politicians who started in better parties. As they have lower switching rates, the observable average exit rate decreases with membership duration. Specifically, for long enough durations, the exit rate reduces to δ. In other words, κ is identified from membership spells that just started and δ is identified from spells that have been around for a long time. Although the politicians differ in their ranking of the parties, this analysis is valid for each subgroup, as also noted by Barlevy and Nagaraja (2010). Moreover, Appendix C shows that, although the party leaders do not make acceptable offers to some politician types, using the effective density of politicians in parties (equation 3.18), one obtains the same result as RB after integrating over the politician types together with the party types.

The dataset includes observations with interrupted, uncensored, right-censored, or interval-censored spells. Heckman and Singer (1984) show that when the duration in a state of interest has the exponential density, conditional on the agents' characteristics, an interrupted spell's duration after the origin date of the sample is also exponentially distributed. Accordingly, left-censoring is not an issue for applying the unconditional inference method of RB. However, the empirical analysis must account for the right- or the interval-censored observations. The latter occurs when a politician loses an election, but reappears in the ballot list of a different party in a consecutive election. Oller, Gomez, and Calle (2004) find that, when the censoring times of an interval-censored observation are not influenced by the specific value of the failure time, the likelihood-based inferences of the transition parameters are consistent. Since the election dates are not influenced by the time a politician leaves a party, the noninformativeness condition holds, and, hence, I adjust the likelihood function for the interval-censored observations. Let d_n, d_r, and d_i denote the indicator functions for the uncensored, right-censored, and interval-censored spells, respectively, that are equal to 1 if a spell has the related type of censoring and 0 otherwise. Appendix C shows that the likelihood of observing a membership duration of t is

$$p(t) = \Big(\frac{\delta(1+\kappa)}{\kappa}\int_1^{1+\kappa}\frac{e^{-\delta at}}{a}da\Big)^{d_n}\Big(\frac{1+\kappa}{\kappa}\int_1^{1+\kappa}\frac{e^{-\delta at}}{a^2}da\Big)^{d_r}\Big(\frac{1+\kappa}{\kappa}\int_1^{1+\kappa}\frac{e^{-\delta at_1}-e^{-\delta at_2}}{a^2}da\Big)^{d_i}, \tag{D.1}$$

where a is the hazard of switching a party. To see the intuition behind identification, notice that the probability of observing a membership duration is decreasing with the length of the duration. This is because the hazard rate is a decreasing function of membership spell duration due to unobserved heterogeneity. Therefore, while the slope of the membership spell duration identifies κ, δ is identified as $t \to \infty$. Note that, since in many samples, there are only a small number of observations with long spells, RB suggests a two-step procedure, in which they first estimate δ from flows into unemployment, and then substitute the estimated value of δ into equation D.1 to estimate κ.

D.2 The hazard function and the agents' types

Having identified the transition parameters (κ, δ), the distribution of the politicians' types can be recovered from the hazard of leaving a party, which, in turn, can be nonparametrically identified. Suppose that we observe M characteristics of politician i's private assets, $\{y_{im}\}_{m=1}^{M}$. There is also an unobserved component of politician i's assets, ϵ_i, distributed $H(\cdot)$ with $h(\epsilon) > 0$ on $(-\infty, \infty)$. Given the unobserved heterogeneity, politician i's assets are equal to

$$log(z_i) = \sum_{m=1}^{M} y_{im}\beta_m + \epsilon_i. \tag{D.2}$$

Accordingly, identification of the distribution of politicians' types is equivalent to identification of the contribution of the observed characteristics to a politician's assets and the distribution of the unobserved heterogeneity. Let

$$y_i = \begin{bmatrix} y_{i1} & y_{i2} & \dots & y_{iM} \end{bmatrix}$$

be the vector of politician i's observable characteristics. For notational simplicity, from now on, I represent a politician's type as

$$log(z_i) = \bar{z}(y_i, \beta) + \epsilon_i. \tag{D.3}$$

Before discussing identification of the distribution of the unobserved heterogeneity, note that only the switches across parties by the medium-type politicians provide identifying information for the unobserved heterogeneity. As the low (high) politician types switch only to the bigger (smaller) parties, they have identical hazard rates conditional on being in the same party. As the unobserved heterogeneity in these politicians' private assets do not affect their hazard rates as long as they fall into these categories, their switches across parties do not provide identifying information about the distribution of the unobserved heterogeneity.

However, for the medium-type politicians with the same observable characteristics, a slight difference in unobserved heterogeneity changes the hazard of leaving a party through the change in politicians' party-switching thresholds. Accordingly, the identification strategy in this section focuses on the medium types of politicians.

Section D.2.1 discusses the identification of the hazard function and the distribution of the unobserved heterogeneity following Evdokimov (2011). Having identified these objects, section D.2.2 discusses identification of the agents' types using the equilibrium density of politician types in parties.

D.2.1 Application of Evdokimov (2011)

In this section, I discuss the applicability of the results of Evdokimov (2011) for the identification of the hazard function and the distribution of the unobserved heterogeneity. Since all party transition processes are Poisson, all corresponding durations are exponentially distributed conditional on a politician's and his party's types, and, hence, the stationary decision rules in equation 3.6 imply that the hazard of a type-z politician leaving a type-x party is

$$\begin{aligned} a(y, x, \epsilon) &= \delta[1 + \kappa[F(x_a(\cdot)) + \bar{F}(x_b(\cdot))]], \\ &= \delta[1 + \kappa[F(x_a(\bar{z}(y_i, \beta), \epsilon_i, x)) + \bar{F}(x_b(\bar{z}(y_i, \beta), \epsilon_i, x)]], \end{aligned} \tag{D.4}$$

with corresponding survivor function

$$S(t|\tilde{z}, x) = \int_0^\infty e^{-\int_0^t a(y,x,\epsilon) ds} h(\epsilon) d\epsilon. \tag{D.5}$$

The left-hand side of equation D.5 is identified from data given x. The hazard function $a(y(\beta), x, \epsilon)$ and the conditional distribution of the unobserved heterogeneity, $h(\epsilon|y(\beta), x)$, on the right-hand side are also identified nonparametrically by theorem 5 in Evdokimov (2011). This theorem shows the sufficient conditions for the identification of the transformation models of the form

$$\Lambda_j(t_j, y, x_j) = m(y_j, x_j, \epsilon) + u_j. \tag{D.6}$$

The hazard function in equation D.4 is a special case of the models defined by equation D.6, where $F(u_j|y, x_j) = 1 - exp(-e^u)$, $exp(\Lambda_j(t_j))$ is the integrated baseline hazard of spell j of length t_j, and $a(y, x_j, \epsilon) = exp(-m(y, x_j, \epsilon))$.

The types of models that Evdokimov considers are more general than the model in equation D.4, and, in general, the required conditions for identification include a random sample of two spells for each individual. In particular, for time-variant hazard models, two spells for a subsample with time-invariant covariates are required for identification of the time-varying component of the hazard function in a fashion similar to Honore (1993). Identification of the time-invariant component of the hazard function, on the other hand, requires a subsample with time-variant covariates, the necessity of which stems from the need for a scale normalization for the unrestricted distribution of the unobserved heterogeneity. With time-variant covariates and two spells for each individual, it is possible to normalize the value of the time-invariant component of the hazard function for a specific covariate in a given spell. Then, the value of the hazard function for all possible covariates in the other spell can be identified up to this normalization. Since he requires two spells for each individual, he chooses the covariate to impose a normalization on as one that each individual can obtain in their second spell independent of what their covariates were in their first spell.

In this paper, the politicians' types are constant over time, but their party affiliations may change. Note that, two spells for each individual is not a requirement for applying Evdokimov's theorem to a time-invariant hazard function, as one only needs to normalize the value of the hazard function for a specific covariate. However, to follow the notation in Evdokimov, I temporarily assume that the data includes two spells for each individual and choose the largest party type, x^{max}, as the covariate to impose a normalization for the unrestricted distribution of the unobserved heterogeneity. Accordingly, define ϵ^* such that

$$\epsilon^* = a(y, x^{max}, \epsilon). \tag{D.7}$$

Recall that the analysis in this section focuses on the medium-type politicians, who consider the largest party as a big party. Since their bigger-party switching threshold when member of a type-x^{max} party is x^{max}, we have that $\bar{F}(x_b(\bar{z}(y,\beta), x^{max})) = \bar{F}(x^{max}) = 0$ for all politicians. Then, the hazard rate in equation D.4 takes the form

$$a(y, x^{max}, \epsilon) = \delta[1 + \kappa F(x_a(\bar{z}(y,\beta) + \epsilon, x^{max}))].$$

Next, define $a^*(\bar{z}(y,\beta), x^{max}, \epsilon^*)$ such that

$$a^*(y, x^{max}, \epsilon^*) \quad = \quad \epsilon^*, \tag{D.8}$$

and note that $\frac{\partial a^*(y(\beta), x^{max}, \epsilon)}{\partial \epsilon} > 0$ as $\frac{\partial x_a(y(\beta)+\epsilon, x^{max})}{d\epsilon} > 0$. Considering the model in equation D.6, suppose that we have a random sample of two spells for each individual,

$$\{T_{i1}, T_{i2}, Y_i, X_i = (X_{i1}, X_{i2}), U_{i1}, U_{i2}, E_i\}.$$

Next, define $\tau_j = \Lambda(T_j, Y, X_j)$ and note that for any $y_i(\beta)$, $x_1 \in [x_{min}, x^{max}]$, and $x_2 = x^{max}$, the conditional characteristic function of τ_j can be written as

$$c_{\tau_1|Y,X_1,X_2}(s|y,x,x^{max}) = c_{m(y,x,E)|Y,X_1,X_2}(s|y,x,x^{max})c_{U_1}(s),$$

and

$$c_{\tau_2|Y,X_1,X_2}(s|(y,x,x^{max}) = c_{m(y,x^{max},E)|Y,X_1,X_2}(s|y,x,x^{max})c_{U_2}(s),$$

where $c_{U_j}(s)$ does not depend on X and are known. Moreover, $c_{U_j}(s) \neq 0$ for all s. Then, the conditional characteristic functions $c_{m(y,x,E)|Y,X_1,X_2}(s|y,x,x^{max})$ and $\phi_{m(y(\beta),x^{max},E)|Y,X_1,X_2}(s|y(\beta),x,x^{max})$ are identified. Since identification of these characteristic functions is equivalent to identification of the distributions, Evdokimov shows identification of the distributions of $-m(y,x,E)$ and $-m(y,x^{max},E)$, conditional on the event $(Y,X_1,X_2) = (y,x,x^{max})$.

Next, for all $\epsilon \in (0,\infty)$ we obtain

$$\begin{aligned}
&exp\{Q_{-m(y,x,E)|X_1,X_2}(F_{-m(y,x^{max},E)|X_1,X_2}(ln\epsilon|x,x^{max})|x,x^{max})\}\\
&\quad = Q_{exp(-m(y,x,E))|X_1,X_2}(F_{exp(-m(y,x^{max},E)|X_1,X_2}(\epsilon|x,x^{max})|x,x^{max})\\
&\quad = Q_{a(y,x,E)|X_1,X_2}(F_{a(y,x^{max},E)|X_1,X_2}(\epsilon|x,x^{max})|x,x^{max})\\
&\quad = Q_{a^*(y,x,E)|X_1,X_2}(F_{a^*(y,x^{max},E)|X_1,X_2}(\epsilon^*|x,x^{max})|x,x^{max})\\
&\quad = a^*(y,x,Q_{E|X_1,X_2}(H_{E|X_1,X_2}(\epsilon^*|x,x^{max})|x,x^{max}))\\
&\quad = a(y,x,\epsilon),
\end{aligned}$$

where the first equality follows by the properties of quantiles and CDFs, the second equality follows from $a(y,x,\epsilon) = exp(-m(y,x,\epsilon))$, the third equality follows by equation D.6 and since ϵ^* is a monotonic transformation of ϵ, the fourth equality follows by the properties of the quantiles and equation D.8, and the last equality follows because E is assumed to be distributed continuously. Accordingly, the function $a(y,x,\epsilon)$ is identified for all x and all ϵ.

D.2.2 The types of the agents

Having identified the structural hazard function conditional on the agents' observed characteristis, the contribution of the observables to a politician's private assets can be identified from the derivatives of equation D.4 with respect to β_m for $m = 1,2,...,M$. Then, the types of the parties are also identified using the equilibrium density of each politician type in a party which was derived in equations 3.17. Formally,

$$\begin{aligned} x &= \int_0^{z^{max}} zg(z|x, \Phi^{l^*}(z,x))dz \\ &= \int_0^{z^{max}} z\frac{\delta(\delta+\lambda)}{\Big[\delta+\lambda[F(x_a(z,x))+\bar{F}(x_b(z,x))\Big]^2}\tilde{\ell}(z)dz \end{aligned} \tag{D.9}$$

where the denominator is the square of the hazard of leaving a type-x party for a type-z politician, which is nonparametrically identified. Moreover, the effective distribution of politician types, $\tilde{L}(z)$, is the convolution of the distributions of $\sum_m y_m\beta_m$ and ϵ. Since the distribution of $\sum_m y_m\beta_m$ is identified from the data and $h(\epsilon|y,x)$ is derived nonparametrically, $\tilde{\ell}(z)$ can be derived by straightforward algebra.

D.3 Rent production function

In this section, I discuss identification of the rent production function and the voters' time-varying preferences for political parties using the vote shares. To do this, I deviate from the model in section 3 in two ways. First, although the assumption of constant party rents is preserved, I deviate from stationarity by assuming that the voters have unobserved preferences for each party that is common among the voters but varies over time.[37] Second, although the model assumes the existence of a continuum of parties, in the empirical section, I assume that there are a finite number of parties, and I use discrete choice theory to identify the rent production function.[38] In the rest of this subsection, I assume that there are $k \in \{1, 2, ..., K\}$ parties participating in an election.

In section 3.3, the value of voting for party k for voter i in district c, v_{ikct}, is given by

$$v_{ikct} = \theta(x_k) + \xi_{kct} + \epsilon_{ikct}, \tag{D.10}$$

[37] Recall that the political rents were defined as the ability to influence the government institutions. I assume that the change in the vote share of a party does not influence how a team with certain assets can affect the decision makers in the government institutions. Time-varying preferences are included into the model solely for explaining the variation in the vote shares over time. Accordingly, I assume that the non-stationarity in voter preferences does not affect the equilibrium conditions derived in section 3.

[38] The discrete choice theory is extended to the case with a continuum of options by Dagsvik (1994) when the distribution of the choices satisfies the independence of irrelevant alternatives (IIA) property, which is generalized further by Malmberg (2013) to the case where this property does not hold. Although these authors' theories can be used to study the voters' preferences when they face a continuum of parties, I model the voters' preferences in a discrete choice framework as there are about 10-15 parties participating in an election during the data's time period. On the other hand, leaving the continuum assumption requires studying the problem in a two-sided matching theoretical framework, which complicates the model tremendously as the number of players are very large (Roth and Sotomayor 1992).

where ξ_{kct} is the electorate's unobserved preference for party k in district c at time t and ϵ_{ikct} is an idiosyncratic taste shock.[39] When ϵ_{ikct} is generated from an extreme value distribution as in the logit model (Anderson, de Palma, and Thisse 1992, Berry, Levinsohn, and Pakes 1995), the vote share of party k in district c at time t is

$$\nu_{kct} = \frac{exp(\theta(x_k) + \xi_{kct})}{\sum_{k=0}^{K} exp(\theta(x_k) + \xi_{kct})}, \tag{D.11}$$

where $k = 0$ is the outside option, i.e., not voting for any party. Let voter i's value of not voting for any party be $v_{i0ct} = \xi_{0ct} + \epsilon_{i0ct}$. Then the probability of not voting for any party is given by

$$\nu_{0ct} = \frac{exp(\xi_{0ct})}{\sum_{k=0}^{K} exp(\theta(x_k) + \xi_{kct})}. \tag{D.12}$$

The inversion theorem in Hotz and Miller (1993) implies that the vote shares, ν_{kct}, have a one-to-one relationship to the choice specific value functions, $\theta(x_k) + \xi_{kct}$. To see this, one can take the logs of each sides in equations D.11 and D.12, and subtract $log(\nu_{0ct})$ from $log(\nu_{kct})$ to get

$$log(\nu_{kct}) - log(\nu_{0ct}) = \theta(x_k) + \xi_{kct} - \xi_{0ct}. \tag{D.13}$$

Normalizing the voters' preferences for not voting for any party to zero, i.e., $\xi_{0ct} = 0, \forall t$, one can identify the rent production function and the distribution of the voters' preferences for a party from equation D.13 conditional on having identified $x_k, \forall k$.[40]

D.4 Club goods production function

Having identified the rent production function, and given the types of politicians and parties, the conditional likelihood of the observed membership durations contains the necessary information to identify both the rent and club goods production functions. Hence the rent production function is overidentified. To see this, consider the probability of observing a membership duration of t for a type-z politician in a type-x party,

$$p(t|z,x) = \delta[1 + \kappa[\bar{F}(x_b(z,x)) + F(x_a(z,x))]]e^{-\delta[1+\kappa[\bar{F}(x_b(z,x))+F(x_a(z,x))]]t}. \tag{D.14}$$

Since the low (high) types of politicians switch only to the bigger (smaller) parties, for these politicians, the likelihood of observing a particular membership duration is indepen-

[39] This voting model's plausibility and divergence from the literature is discussed in section 3.3.

[40] This normalization is common in estimation of static games of strategic interactions, where one inverts the equilibrium choice probabilities for nonparametric identification of the choice specific value functions. See Bajari, Hong, Krainer, Nekipelov (2010).

dent of the politician's private assets, and determined solely by the labor market transition parameters and the sampling distribution. Accordingly, given the sizes of the parties (and, hence, the distribution of party sizes), these politicians' switches across parties do not contain any information about the parties' rents or club goods. This is because, the low (high) types' private assets are too low (high) that, the loss in terms of the private rents (club goods) never dominates the gain in terms of the club goods (rents) as the party size increases.

The hazard out of a party, and, in turn, a politician's switch across parties contains information about the parties' rents and club goods only if the politician is a medium type, meaning that a politician's valuation of the private rents over the club goods may cause him to rank two parties with different sizes of equal value. This is because, the hazard rate of leaving a type-x party increases in politician type on $(\underline{z}, z_0(x))$ and decreases on $(z_0(x), \bar{z})$, as $z_0(x)$ is the politician type who considers a type-x party as the worst. Accordingly, the derivative of the likelihood of a membership duration with respect to politician type should be decreasing on $(\underline{z}, z_0(x))$ and increasing on $(z_0(x), \bar{z})$. Then, given the types of agents and having identified the rent production function, the variation in the hazard of leaving a party across politician types identifies the club goods production function when the maximum rent share a party offers to its members is known. Intuitively, one can think of the change in politician type as providing a variation in the demand for club goods, which identifies the supply of club goods. Moreover, given a politician type, the variation in the hazard of leaving a party across different parties identifies the overall ranking of party values, and, hence, enables one to identify the rent production function given club goods production function.

D.5 Discount rate

This section discusses the identification of the discount rate from a politician's spell of being an independent after having identified the primitive functions and the other parameters. Recall that, the lifetime utility of being an independent for a type-z politician is $V_0(z) = \frac{1+\rho}{\rho}\theta(z)$ (equation 3.8), and he accepts the membership offers of the parties that provide a lifetime utility that is at least as much as that of being an independent. When a politician's private assets are too low that he can accept the membership offer of any party, his spell of being an independent is not informative about the discount rate, as it only depends on the offer arrival rate. However, when the politician's assets are high enough that he rejects the membership offers of some parties, then the hazard of joining a party depends on the discount rate. Therefore, to discuss identification, from now on I focus on the politician types for whom the lifetime utility of being an independent is strictly greater than the value of joining some parties. Suppose that the politician accepts the offer of a type-x party when $x \in [x_{min}, x_a(z)] \cup [x_b(z), x^{max}]$, where the thresholds $x_a(z)$ and $x_b(z)$ are the types of two parties that provide the same lifetime utility to the politician as that of being an independent. Accordingly, these threshold party types solve

$$V_0(z) = V(z, 1, 1, x_a(z)) = V(z, 1, 1, x_b(z)), \tag{D.15}$$

where the equilibrium property that $\phi^{l^*}(z, x) = 1, \forall z, x$ is substituted into the value functions.

An independent type-z politician joins a party at rate $\lambda[F(x_a(z)) + \bar{F}(x_b(z))]$. So, the conditional probability of observing a spell of being an independent of length t_0 is

$$p(t_0|z) = \lambda[F(x_a(z)) + \bar{F}(x_b(z))]e^{-\lambda[F(x_a(z))+\bar{F}(x_b(z))]t_0}. \tag{D.16}$$

The derivatives of equation D.16 with respect to t_0 and z provide two equations that can be used to identify the hazard of joining a party as well as the variation across the independent politicians of the hazard of joining a party. Identifying the hazard of joining a party, in turn, identifies the discount rate, as equation D.15 implies

$$(1 + \rho)\theta(z) = \frac{z\theta(x_a(z))}{x_a(z)} + \psi(x_a(z)) = \frac{z\theta(x_b(z))}{x_b(z)} + \psi(x_b(z)). \tag{D.17}$$

Having identified z, $\theta(x)$, and $\psi(x)$, equation D.17 depends only on ρ. Hence, the discount rate is identified after recovering the politicians' party-switching thresholds.