

The Politics of Fiscal Federalism: Building a Stronger Decentralization Theorem

Raúl A. Ponce-Rodríguez

Department of Economics
Universidad Autónoma de Ciudad Juárez

Charles R. Hankla

Department of Political Science
Georgia State University

Jorge Martínez-Vazquez

Department of Economics
International Center for Public Policy
Georgia State University

Eunice Heredia-Ortiz¹

Development Alternatives Inc., DAI

We explore how party structures can condition the benefits of decentralization in modern democracies. In particular, we study the interaction of two political institutions: democratic (de)centralization (whether a country has fiscally autonomous and elected local governments) and party (non)integration (whether power over local party leaders flows upwards through party institutions, which we model using control over candidate selection). We incorporate these institutions into our *strong decentralization theorem*, which expands on Oates (1972) to examine when the decentralized provision of public services will dominate centralized provision even in the presence of inter-jurisdictional spillovers. Our findings suggest that, when externalities are present, democratic decentralization will be beneficial *only* when parties are integrated. In countries with non-integrated parties, we find that the participation rules of primaries have implications for the expected gains from democratic decentralization. Under blanket primaries, Oates' conventional decentralization theorem holds but our strong decentralization theorem does not. By contrast, when primaries are closed, not even Oates' conventional decentralization theorem holds.

Keywords: Decentralization, Fiscal Federalism, Formal Theory, Political Parties, Public Goods.

¹The order of authors has been drawn randomly. We would like to thank Roger Myerson, Irfan Nooruddin, Craig Volden, Katrina Kosec, Emmanuel Haruna, Kyle Hanniman, and participants in the World Bank's 2018 Conference in Development Economics for their helpful feedback. Our thanks also go to the International Center for Public Policy at Georgia State's Andrew Young School of Policy Studies for financial support. Finally, we would like to thank the members of the Deil Wright Awards committee for naming a much earlier version of this work the best paper on federalism presented at the 2011 American Political Science Association Convention in Seattle.

1. Introduction

In the last few decades, numerous countries – China, Indonesia, South Africa, India, the United Kingdom, and many others – have engaged in decentralization reforms. These reforms, at least in the developing world, have been supported both by the aid dollars of multilateral agencies such as the World Bank and USAID, and by the research findings of many scholars. Central to these positive scholarly judgments is the “decentralization theorem,” which was developed by Oates (1972) and states that “. . . in the absence of cost-savings from the centralized provision of a (local public) good and of inter-jurisdictional externalities, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained across all jurisdictions” (p.54).

However, as the process of decentralization has continued apace, some scholars have begun to question whether devolving authority to regional and local governments is a universal good. Among other things, they have pointed out that Oates, in developing his theorem, assumes not only the absence of economies of scale and externalities, but also that policies are implemented by benevolent welfare-maximizing governments. While this latter assumption may be useful for creating a simple and elegant theory of decentralization, it hardly accords with empirical realities. More to the point, it begs the question of how different political processes and institutions might shape the fiscal choices made by policy makers, and, with them, the outcomes of fiscal federalism.

In this paper we seek to answer that question by analyzing the interaction of two particular political institutions: democratic (de)centralization (whether a country has fiscally autonomous and elected local governments) and party (non)integration (whether power over local party leaders

flows upwards through party institutions to leaders at the national level, which we model using candidate selection procedures).² We incorporate these institutions into a rigorous and formal extension of Oates (1972), which we term our *strong decentralization theorem*.

Our findings suggest that, when public goods have externalities, a move towards democratic decentralization will only produce the benefits predicted by Oates if parties are integrated. When local governments are elected and autonomous, their leaders possess the accountability necessary to incentivize public goods provision. But when national party leaders also control access to the ballot, local leaders have strong reasons to provide the efficient level of goods even when their benefits spillover across jurisdictions. Put differently, democratic (de)centralization affects public spending because local elections and nationwide elections create different incentives for the officials who design public policy. Upward accountability through integrated party mechanisms can influence these incentives.

Narrowing our focus to non-integrated parties, we also study how the institutions of primary elections shape decisions over public spending. We find that the participation rules of primaries, whether *closed* or *blanket*, have important implications for the expected gains from democratic decentralization. Blanket primaries are an extreme form of open primary in which the top two candidates, regardless of party, compete in the general election.³ Under such system, democratic decentralization only produces the expected benefits when inter-jurisdictional spillovers are absent. By contrast, when primaries are closed, not even Oates' conventional

² Note that, in Ponce et al. 2018, we use the term “party (de)centralization” for a similar concept. As we explain later in this paper, we adopt the term “party (non)integration” here because it corresponds more closely to our intended meaning and is consistent with the use of the term in our recent book, Hankla et al. 2019.

³ See the National Council of State Legislatures for more on types of primary elections: <http://www.ncsl.org/research/elections-and-campaigns/primary-types.aspx>.

decentralization theorem holds, meaning that democratic decentralization will produce no consistent benefits vis-à-vis centralized provision. This is because, in a closed primary system, parties will hew too closely to the specific preferences of their local co-partisans rather than to the aggregate interests of all constituents.

This article contributes to the extensive literature on decentralization in two important ways. First, it significantly expands the main theoretical construct in fiscal decentralization—Oates’s decentralization theorem—by showing that under the right political institutions, the fiscally decentralized provision of local public goods can be more efficient even in the presence of inter-jurisdictional externalities. It also shows that under certain political institutions, Oates’s decentralization theorem does not continue to hold. Second, and more generally, the paper clearly demonstrates that different forms of decentralization—fiscal versus democratic versus party – are closely intertwined and cannot be analyzed separately, as has been previously done in most of the literature.

Given the extensive variation in party organization that exists in the real world, these findings are of great empirical significance. Of the 157 countries that we have coded for our empirical work (see Hankla et al. 2019; Ponce et al. 2018), 106 were characterized by some degree of democratic decentralization. Of these 106, sixty-seven countries have had integrated parties for at least some years and forty-nine have consistently had non-integrated parties.

Moreover, the results from our empirical work point – as expected – to the benefits of combining democratic decentralization with party integration. Our statistical analyses show that countries with such institutions possess an advantage in a number of primary health, primary education, and local infrastructure outcomes. The benefits of integrated parties also show up in our comparison of party integrated Senegal with party non-integrated Nigeria (Hankla et al. 2019).

This article provides a rigorous theoretical foundation for the empirical work presented in Ponce et al. (2018). Moreover, while the models developed here build on those in our book (Hankla et al. 2019), we introduce several significant extensions. These include relaxing the assumption that local public goods are homogeneous and analyzing majoritarian electoral systems with both blanket and closed primary elections.

We structure the rest of the article as follows: Section 2 reviews the relevant previous literature. Section 3 contains a summary and detailed explanation of the main findings of our theory. Section 4 includes a benchmark model in which local public goods are efficient and match the heterogeneous preferences over public spending from voters. Section 5 incorporates the analysis of party integration and democratic centralization. Section 6 studies the case of party integration and democratic decentralization. Section 7 contains the introduction of our analysis of non-integrated parties. Section 8 considers the case of nonintegrated parties and democratic centralization. Section 9 studies non-integrated parties and democratic decentralization. Section 10 develops a comparative analysis between democratic centralization versus decentralization in economies with non-integrated parties. Section 11 includes a comparison of our findings and a brief discussion of the main assumptions driving our results. Section 12 concludes.

2. *Literature background*

There has been considerable previous research on the role of political institutions in public goods delivery, but most of it has concerned the national level. In this vein, scholars have investigated the impact of a number of specific institutions (e.g., electoral systems, legislative-executive relations, legislative and coalition party fragmentation) on policy outcomes such as

government spending, free trade, balanced budgets, and economic growth (see Persson and Tabellini 2003, O'Halloran 1994, Nielson 2003, Haggard and Kaufman 1995, Hallerberg and Marier 2004, Roubini and Sachs 1989 and Volkerink and de Haan 2001). Work on the influence of political institutions on subnational public goods delivery has been much sparser, as we discuss below.

Researchers focused on decentralization, for their part, have examined the relationship between the organization of parties, on the one hand, and the devolution of state power, on the other. For example, Chhibber and Kollman (2004) make the case that countries decentralizing authority to the subnational level are likely to have more localized party systems, while Fabre et al. (2005) find that such countries will also be characterized by more decentralized parties.

Perhaps William Riker is the most prominent scholar to have taken up this question. He argues in his 1987 book that the American “decentralized party system is the main protector of the integrity of states in our federalism” (p. 221).⁴ By contrast, Filippov, Ordeshook, and Shvetsova (2004) emphasize the benefits of more integrated parties, making the case that party systems which successfully link the national and sub-national levels of government are the best guarantors of a stable federal system. Myerson (2006) concurs with the latter authors, arguing that regional and local elections provide opportunities for potential national candidates to prove themselves at the sub-national level.

Oates' (1972) decentralization theorem has also inspired considerable work in the last few decades. While its assumption of benevolent public officials has been confronted in the “second generation” of research on decentralization, few scholars have examined which political

⁴ See also Volden 2004 for an excellent summary of Riker's thought on Federalism.

institutions support the social welfare gains expected from the theorem.⁵ In particular, second generation scholars have tended to focus on the problems of assignment and soft budget constraints, and much less on the relationship between specific political institutions at the subnational level and the efficient provision of local public goods in fiscally decentralized systems.

Despite these limitations, recent work on the political economy of fiscal federalism highlights the need to examine more closely how political institutions can influence the expected gains from decentralization (see Lockwood 2015). For instance, Lockwood (2002) uses a model of legislative bargaining to show that welfare is not increasing with higher spillovers under centralization, which is one of the main advantages of centralization suggested by Oates (1972). Lockwood (2008) studies whether the decentralization theorem holds when collective choices are made by majority rule and lobbying. And Bordignon et al. (2008) characterize conditions in which lobbying through campaign contributions induces a decentralized provision that is not Pareto efficient when politicians become too greedy.

Even more to the point, Besley and Coate (2003), in a legislative model, show that local public goods are not automatically Pareto efficient since public spending maximizes the utility of a median voter instead of a social welfare function. This outcome means that the decentralization theorem identified by Oates (1972) is not necessarily compatible with the electoral incentives of politicians. However, in spite of their focus on the role of legislatures in fiscal decentralization, Besley and Coate (2003) do not consider how the structures of political parties influence public spending behavior. In this paper we develop such an analysis.

⁵ For more on second generation research on fiscal federalism, see Weingast (2014), Weingast (2009), and Oates (2005).

Other contributions in political economy have focused on providing rationales for the superiority of fiscal decentralization that are distinct from the one provided by Oates (1972). For instance, scholars have argued that fiscal decentralization is preferable because of its impact on corruption (Bardhan and Mookherjee 2000), accountability (Seabright 1996), and, under certain conditions, government capture (Bardhan and Mookherjee 2002).

Enikolopov and Zhuravskaya (2007) is one rare study that focuses on the political and institutional conditions under which decentralization can deliver on its promises. These scholars argue convincingly that fiscal decentralization produces better outcomes in the presence of party integration. That said, they are not interested in democratic decentralization (i.e. elections) as we are here, and their paper is entirely empirical, with no formal component.

Building on all of this prior work, we demonstrate in this paper that the decentralization theorem, which lies at the heart of the fiscal federalism literature, is dependent on the structure of political institutions. Local elections and certain forms of party institutions, we argue, must be in place before decentralization can deliver on its promises.

3. *Brief introductory tour to the theory*

As noted above, the main goal of this paper is to create a rigorous and formal extension of Oates' influential "decentralization theorem." Of interest for us is the interaction between democratic (de)centralization and party (non)integration in the provision of local public goods. Under democratic decentralization local public goods are provided by democratically elected local governments, while under democratic centralization they are provided by the national government.

Our notion of what constitutes an integrated party is the same regardless of the electoral system used in a country. For us, a party is integrated when power over local party officials flows upwards through party institutions to leaders at the national level. This concept, more fully defined in Hankla et al. (2019), encompasses a number of characteristics. For example, integrated parties are governed by strong institutions rather than personalist leaders, and they are present at all tiers of the polity. For modeling simplicity, however, we focus on what is perhaps the most salient feature of an integrated party – that its national leaders control access to the party name in local elections. When local politicians can run for office using the name of a party without the permission and oversight of party leaders in the capital, the party cannot be considered integrated.

For space reasons, we focus our formal analysis in this paper on majoritarian, single-member-jurisdiction systems, defining non-integrated parties as those that hold blanket or closed primaries (modeled separately) to choose candidates, as opposed to those having national party leaders nominate them.⁶ While we understand that many non-integrated parties practice free candidate nomination procedures (i.e. by collecting signatures or paying a fee) rather than primaries, we believe these decentralized structures will have many of the same effects as primaries (see Carey and Shugart 1995).

In the majoritarian system models that are our focus here, we find that both the conventional decentralization theorem (which assumes away spillovers) and our new *strong decentralization theorem* (which allows for spillovers) hold—under democratic decentralization—when parties are integrated. In other words, we find that, when integrated national parties are active

⁶ See Hankla et al. (2019) for a previous analysis of systems using proportional representation, and Ponce et al. (2018) for an empirical test of our arguments. Note also that the model assumes that parties possess meaningful organizational structures and are not purely dominated by personalist “bosses.” It also assumes competitive elections at the local level.

in local politics, locally elected governments will outperform central governments in providing public goods. This pattern will hold even in the presence of inter-jurisdictional spillovers.

In such situations, democratic decentralization creates local governments that are accountable to their voters' preferences, while the vertical responsibility within integrated parties incentivizes these local governments to finance the efficient provision of local goods and services, even those with benefits that spill across jurisdictions. This is the key implication of our *strong decentralization theorem*.

The situation is different when parties are non-integrated. In this case, it is necessary to consider the structure of primary elections in order to predict varying incentives for the provision of public goods and to determine whether democratic decentralization dominates centralization. For instance, we find that, under blanket primaries, our *strong* theorem *is not* satisfied but the conventional decentralization theorem is. Primary elections do not create the necessary incentives for local leaders to internalize spillovers, which means that the *strong decentralization theorem* does not generally hold for democracies with non-integrated parties. However, systems with blanket primaries are an exception. Here, local public goods with no spillovers are Pareto efficient and the benefits of policy differentiation are captured. By contrast, while central governments also provide Pareto efficient local public goods, they are not able to maximize the gains associated with the heterogeneity of preferences.

By contrast, when parties are non-integrated and primaries are closed, neither the *strong* nor the conventional decentralization theorem holds. For countries with such nominating institutions, parties operating under both democratically centralized and decentralized structures will lack incentives to provide the efficient level of local public goods. This is because the participation rules of closed primaries allocate voting rights only to members of a specific party

rather than to the electorate in general. For this reason, parties operating under both structures of government have political incentives to target goods only to those voters who count. This intuition suggests why local public goods provision might not be efficient and might not successfully maximize the gains expected from matching policy with the heterogeneous preferences of voters.

These findings have significant implications for the scholarly understanding of decentralization among both political scientists and economists. Most obviously, they show that political institutions mediate the effects of the decentralized provision of public goods to an extent not previously realized. In addition, our findings contribute to the theory of fiscal federalism and help make sense of the mixed results that characterize the empirical scholarship on the actual impact of decentralization on service delivery, economic growth and other dimensions (see, for example, Martinez-Vazquez et al. 2016). And, for development practitioners, they have the potential to encourage a deeper examination of the types of political institutions that may be necessary for decentralization reforms to produce fuller positive results.

4. *The Benchmark and Definitions*

We begin by characterizing the set of local public goods that maximize society's surplus from public goods. This approach to benchmarking our results allows us to compare them to those in the normative analysis of Oates (1972) and the more recent political economy analyses of Besley and Coate (2003) and Lockwood (2015).

Consider an economy composed of jurisdictions $i = \{1,2\}$ with a corresponding population of n_i in each jurisdiction. Moreover, we assume $n_1 \neq n_2$. Following most of the literature, and for simplicity of the analysis, in our economy individuals do not have mobility across jurisdictions.

The utility of an individual h in jurisdiction 1 is $v_{h1}(\alpha_{h1}, g_1, g_2) = \alpha_{h1}[\ln(g_1) + k_2 \ln(g_2)]$ where g_1 and g_2 are public goods provided by jurisdictions 1 and 2 and $\alpha_{h1} > 0$ is a parameter of intensity of preferences for public goods consumed by a resident h of jurisdiction 1. The parameter $k_2 \in [0,1]$, measures the extent of inter-regional spillovers of g_2 over residents of jurisdiction 1. For local public goods without spillovers $k_2 = 0$, and when jurisdiction 2 provides a nationwide pure public good, $k_2 = 1$. A similar expression is given by an individual h living in jurisdiction 2 such that $v_{h2}(\alpha_{h2}, g_2, g_1) = \alpha_{h2}[\ln(g_2) + k_1 \ln(g_1)]$. In this economy, preferences are heterogeneous within and across districts. Hence, for individuals h and h' : $h \neq h'$, $\alpha_{h1} \neq \alpha_{h'1} \neq \alpha_{h2} \neq \alpha_{h'2}$. In our model, the size of spillovers between jurisdictions is asymmetric therefore $k_1 \neq k_2$. The marginal cost of producing a public good in all jurisdictions is q . The nationwide net social welfare for this economy is given by NS :

$$NSW = \sum_{h=1}^{n_1} v_{h1}(\alpha_{h1}, g_1, g_2) + \sum_{h=1}^{n_2} v_{h2}(\alpha_{h2}, g_2, g_1) - q(g_1 + g_2) \quad (1)$$

Proposition 1 characterizes a set of local public goods that are Pareto efficient and maximize the gains attributed to matching the size of local public spending to the heterogeneous preferences of individuals across jurisdictions.

Proposition 1. *The Pareto efficient local public goods $\hat{g}^* = [\hat{g}_1^*, \hat{g}_2^*]$ that maximize the gains from inter-regional policy differentiation are given by:*

$$[\hat{g}_1^*, \hat{g}_2^*] = \left[\frac{\alpha_1 + k_1 \alpha_2}{q}, \frac{\alpha_2 + k_2 \alpha_1}{q} \right] \quad (2)$$

Where $\alpha_1 = \sum_{h=1}^{n_1} \alpha_{h1}$ is the aggregate intensity of preferences for local public goods of all residents in district 1 and $\alpha_2 = \sum_{h=1}^{n_2} \alpha_{h2}$ is the corresponding aggregate intensity of preferences of all residents of district 2.

Proof.

Find $\frac{\partial NSW}{g_1} = \frac{\sum_{h=1}^{n_1} \alpha_{h1}}{\hat{g}_1^*} + k_1 \frac{\sum_{h=1}^{n_2} \alpha_{h2}}{\hat{g}_1^*} - q = 0$ and $\frac{\partial NSW}{g_2} = \frac{\sum_{h=1}^{n_2} \alpha_{h2}}{\hat{g}_2^*} + k_2 \frac{\sum_{h=1}^{n_1} \alpha_{h1}}{\hat{g}_2^*} - q = 0$ re-arrange terms to obtain the result in (2).

In (2), local public goods with and without spillovers in each jurisdiction, \hat{g}_i^* for $i = 1, 2$, are provided at the point in which the nationwide marginal social benefits of \hat{g}_i^* in both jurisdictions are equivalent to the social marginal costs, q , of producing the local public good in jurisdiction i , therefore, \hat{g}_i^* is Pareto efficient. For instance, the nationwide marginal social benefits of \hat{g}_1^* include the marginal utility gains of all residents of jurisdiction 1, $\frac{\sum_{h=1}^{n_1} \alpha_{h1}}{\hat{g}_1^*}$, and due to inter-regional spillovers, the marginal utility gains of all residents of jurisdiction 2, $k_1 \frac{\sum_{h=1}^{n_2} \alpha_{h2}}{\hat{g}_1^*}$. Therefore, \hat{g}_1^* , depends positively on the intensity of preferences of all residents of jurisdiction 1, which are given by $\alpha_1 = \sum_{h=1}^{n_1} \alpha_{h1}$, the intensity of preferences of all residents of jurisdiction 2, which are given by $\alpha_2 = \sum_{h=1}^{n_2} \alpha_{h2}$, the extent of inter-regional spillovers (if $k_1 > 0$), and negatively on the marginal cost of the public good q . The heterogeneity of preferences of individuals across jurisdictions implies that in general $\hat{g}_1^* \neq \hat{g}_2^*$.

5. *Party Integration in a Democratically Centralized Government*

In this section we introduce a model of electoral competition in a democracy characterized by party integration (modeled as party leaders controlling the nomination of candidates participating in elections) and democratic centralization (modeled as voters electing an official to run the central government and with no local elections). We assume that if systems are democratically centralized then they are also fiscally centralized. Similarly, democratically decentralized systems, i.e., those with elected subnational governments, are also fiscally decentralized, with subnational officials having autonomy over spending and taxing decisions in their respective jurisdictions.

In our economy, there are two parties $P = \{A, B\}$. The parties' problem is to aggregate the heterogeneous and conflicting views of voters over public spending into a policy platform that maximizes the parties' probabilities of winning the election. We characterize the political equilibrium of this economy throughout a sequential game of complete and perfect information in which all actions of all players in each stage of the game are common knowledge. In the first stage of the game, candidates announce policies seeking their party's nomination and party leaders nominate the candidate that will run in the general election. Party leaders nominate a candidate by selecting a candidate $j = \{1, 2\}$ to the nomination set NS_P with $P = \{A, B\}$. In the second stage voters observe the parties' policies and vote. In the third stage, the winning candidate takes all, forms the government, and implements policy.

In the first stage, two candidates $j = \{1, 2\}$ in each party seek the nomination of their parties. For an economy with party integration, party leaders have full command over policy making by nominating only those candidates who adopt the ideal policy of party leaders. In our

economy, the two parties compete in the election to form the government in a majoritarian electoral system with single member jurisdictions. Under a central government, local public goods are provided by a single government that represents voters of all jurisdictions. The government finances its expenditures through a uniform tax on residents of all jurisdictions τ satisfying the budget constraint $\tau(n_1 + n_2) = q(g_1 + g_2)$, where $\tau(n_1 + n_2)$ is the tax revenue and $q(g_1 + g_2)$ is the nationwide spending in local public goods. We assume the central government provides a uniform local public good.⁷

In the second stage of the game, voters observe the parties' policies and vote. Voting is sincere and all individuals vote.⁸ Denote $\Psi_{Ah1}(\alpha_{h1})$ as the net payoff received by a voter with preference α_{h1} in jurisdiction 1 from party A relative to that of party B where $\Psi_{Ah1}(\alpha_{h1}) = \{v_{Ah1}(\alpha_{h1}, g_{A1}, g_{A2}) - \tau_A + \sigma_{Ah1}\} - \{v_{Bh1}(\alpha_{h1}, g_{B1}, g_{B2}) - \tau_B + \sigma_{Bh1}\}$ and where $v_{Ah1}(\alpha_{hi}, g_{A1}, g_{A2})$ is the welfare of voter when party A selects spending policies g_{A1}, g_{A2} , $v_{Bh1}(\alpha_{hi}, g_{B1}, g_{B2})$ is the welfare when party B selects spending policies g_{B1}, g_{B2} , and τ_A and τ_B are taxes on residents of all jurisdictions under parties A and B. We follow McKelvey and Patty (2006), and Coughlin (1992) in assuming that the net payoff $\Psi_{Ah1}(\alpha_{h1})$ also depends on stochastic factors σ_{Ah1} and σ_{Bh1} . Voter type α_{h1} in jurisdiction 1 votes for party A if $\Psi_{Ah1}(\alpha_{h1}) > 0$, for party B if $\Psi_{Ah1}(\alpha_{h1}) < 0$, and flips a fair coin if $\Psi_{Ah1}(\alpha_{h1}) = 0$. An equivalent interpretation is

⁷ The assumption of uniform central provision is quite standard in the literature. It is now agreed that central governments may provide heterogeneous services but that the central government may also have less information (Cremer and Palfrey 1996) and fewer incentives. Moreover, although federations may lead to horizontal fiscal externalities (Wilson 1999), and citizens could benefit from a coordinated central policy, this could also lead to less accountability and efficiency (Seabright 1996). Related to this same issue, Tomassi and Weinschelbaum (2003) characterize a tradeoff between externalities and accountability in which decentralization might be preferred even if preferences are identical. In summary, assuming that the central government provides uniform public goods could be interpreted as just convenient shorthand for assuming that centralized and decentralized governments have different mechanisms to match local preferences and needs with policy.

⁸ The assumption of sincere voting seeks to simplify the analysis and it ignores strategic voting behavior such as credible threats of some coalition of voters who might abstain from voting for the nominated candidate in the general election if the candidate changes the policy position he previously announced.

given to $\Psi_{Ah2}(\alpha_{h2})$ which is the net payoff received by a voter with preference α_{h2} in jurisdiction 2 from party A relative to that of party B.

Define $\varepsilon_{h1} = \sigma_{Bh1} - \sigma_{Ah1} = \{v_{Ah1}(\alpha_{h1}, g_{A1}, g_{A2}) - \tau_A\} - \{v_{Bh1}(\alpha_{h1}, g_{B1}, g_{B2}) - \tau_B\}$ and assume $\varepsilon_{h1} \in [\underline{\varepsilon}_{h1}, \bar{\varepsilon}_{h1}]$, then probability that a voter type α_{h1} votes for party A in jurisdiction i is $F_{Ah1} = \int_{\underline{\varepsilon}_{h1}}^{\bar{\varepsilon}_{h1}} f_{Ah1}(\varepsilon_{h1}) d\varepsilon_{h1}$, where $f_{Ah1}(\varepsilon_{h1})$ is a continuous probability distribution over ε_{h1} . The expected vote of party A in jurisdiction 1 is $\phi_{A1} = \sum_{h=1}^{n_1} F_{Ah1}(\varepsilon_{h1})$ and the nationwide expected vote of party A is $\phi_A = \sum_{i=1}^2 \phi_{Ai}$ and that of party B is $\phi_B = \sum_{i=1}^2 \phi_{Bi}$. Define $\pi_A : \rho_A \rightarrow [0,1]$ as a continuous cumulative distribution over the plurality of the party, $\rho_A = \phi_A - \phi_B$, where $\rho_A \in [\underline{\rho}_A, \bar{\rho}_A]$ and $\pi_A = \int_{\underline{\rho}_A}^{\bar{\rho}_A} w_A(\rho_A) d\rho_A$, where $w_A(\rho_A)$ is the probability distribution over the party's plurality. Following our previous discussion, we define π_B as the probability that party B wins the election. We follow the previous literature (see, for example, Coughlin 1992) and assume π_A is strictly concave on g_{A1}, g_{A2} and π_B is strictly concave on g_{B1}, g_{B2} .

In the third stage, we define Ω as a non-decreasing cumulative distribution of the sequence $\{\Psi_{Ahi}(\alpha_{hi})\}_{\forall h, \forall i}$, hence $\Omega \in [0, 1]$. Therefore, if there exists a majority of individuals voting for party A then $\Omega(\forall \alpha_{hi}: \Psi_{Ahi}(\alpha_{hi}) > 0) > 1/2$ and party A wins the election and implements its policy platform.⁹ Otherwise, party B wins the election and implements its policy platform.

The Subgame Perfect Nash Equilibrium of our political process is constituted by the policy platforms of candidates seeking nomination g_{Pi}^* for $P = \{A, B\}$ and $i = \{1, 2\}$, the nomination choices of party leaders NS_P for both parties, and the individuals' choices of the vote in all

⁹ Note that $\Omega \in [0, 1]$ is a non-decreasing cumulative distribution of the sequence $\{\Psi_{Ahi}(\alpha_{hi})\}_{\forall h, \forall i}$, hence, given some policies g_{A1}, g_{A2} of party A and policies g_{B1}, g_{B2} of party B, if $\Omega(\forall \alpha_{hi}: \Psi_{Ahi}(\alpha_{hi}) > 0) = 0.3$ then 30% of voters vote for party A and 70% for party B.

jurisdictions $\Psi_{Ahi}(\alpha_{hi}) \stackrel{>}{<} 0$ for all voters in all districts. For a formal definition of the equilibrium see definition 1 in the appendix. On what follows, the equilibrium provision of local public goods is characterized in Lemma 1. In our economy, parties converge in their policy platforms since they maximize a continuous and strictly concave probability function of winning the election based on a common system of beliefs and strategy policy set.¹⁰ In this section, and with the purpose of differentiating the outcome from this section with upcoming sections of the paper, we define the equilibrium policy as g_c^* . Formally, our main result of this section is the following.

Lemma 1 *Local public goods with and without spillovers are Pareto efficient for an economy with a majoritarian electoral system, single member jurisdictions, a democratically centralized government and integrated parties. All parties converge in providing a uniform local public good across jurisdictions, $g_1^* = g_2^* = g_c^*$ where g_c^* is defined as follows:*¹¹

$$g_c^* = \frac{\hat{g}_1^* + \hat{g}_2^*}{2} \quad (3)$$

Where

$$[\hat{g}_1^*, \hat{g}_2^*] = \left[\frac{\alpha_1 + k_1 \alpha_2}{q}, \frac{\alpha_2 + k_2 \alpha_1}{q} \right] \quad (4)$$

Proof

See the appendix

¹⁰ For a formal proof of convergence in probabilistic voting models with homogeneous parties see Coughlin (1992).

¹¹ Because all candidates of all parties converge in selecting their policy platforms, in Lemma 1 we have drop the subscripts associating policy with candidates and parties.

Lemma 1 says g_c^* is a Pareto efficient provision of local public goods since the institution of democratic centralization provides voting rights to residents of all jurisdictions and parties have electoral incentives to take into account the marginal benefits and costs of public goods of all residents of all jurisdictions. As a result, parties select a policy in which the nationwide social marginal benefits of public goods are equal to the society's marginal costs. In addition, since the central government provides a uniform public good across jurisdictions then democratic centralization and party integration create incentives for politicians to provide a local public good, g_c^* , in all jurisdictions that is given by $g_c^* = \frac{\hat{g}_1^* + \hat{g}_2^*}{2}$, where g_c^* corresponds to the average of the socially optimal provision of public goods in jurisdiction 1 and 2, that is, \hat{g}_1^* and \hat{g}_2^* .

6. *Party Integration and Democratic Decentralization*

In this section we consider the case of party integration and democratic decentralization (i.e. local government elections with fiscal decentralization). To anticipate the results that follow below, this section has two main findings. First, we show that party integration in a system of local governments leads to Pareto efficient local public goods with and without inter-regional spillovers. This finding that local elections might lead to Pareto efficient local public goods with inter-regional spillovers is different to the findings of most political economy models in the literature (see for instance Besley and Coate 2003 and many others). Second, local public spending is differentiated to match the heterogeneous spending policies demanded by voters across jurisdictions. Third, we show that, under party integration, democratic decentralization dominates democratic centralization.

To prove these claims, we develop again a sequential game of complete and perfect information with local elections. In the first stage of the game, candidates announce policies seeking their party's nomination and party leaders nominate the candidate that will run in the local election of jurisdictions 1 and 2. In the second stage, voters observe the parties' policies and vote in the local election of their jurisdiction. In the third stage, the winning candidate in the local election takes all, forms the government and implements policy.

In a federation with integrated parties, party leaders want to win multiple local elections and nominate candidates who propose policies that maximize the party's *joint* probability of winning elections in jurisdictions 1 and 2.¹² In this case, party leaders have full command on policy making by nominating only those candidates who adopt the ideal policy of party leaders. Party Leaders nominate a candidate by selecting a candidate $j = 1, 2$ in jurisdictions 1 and 2 to the nomination set NS_P for parties $P = \{A, B\}$.

Two parties compete in the local election of each jurisdiction to form the local government. For the analysis that follows, we define the joint probability of party A winning local elections in jurisdictions 1 and 2 by $\pi_{LA} = \pi_{LA}(\rho_{LA1}, \rho_{LA2})$ where π_{LA} is a function of the pluralities of the party in both jurisdictions, ρ_{LA1} and ρ_{LA2} , where $\rho_{LAi} = \phi_{LAi} - \phi_{LBi}$, and $\phi_{LAi} = \sum_{h=1}^{n_1} F_{LAhi}(\varepsilon_{hi})$ is the number of votes that party A expects to receive in the local election of jurisdiction i and $F_{LAhi}(\varepsilon_{hi})$ is the marginal probability that a voter type α_h^i votes for party A in the local election of jurisdiction i (a similar interpretation is given to ϕ_{LBi}). A similar expression is given for the

¹² In the previous section, under democratic centralization, parties design spending policies to maximize the parties' probability of winning a single national election in which voters from all jurisdictions vote. In this section, under democratic decentralization, parties with nationwide presence, design policies to maximize the joint probability of winning all local elections.

probability of party B of winning the local election π_{LB} . We assume that π_{LA} and π_{LB} are continuous and strictly concave function of public goods in jurisdictions 1 and 2.

In the second stage of the game, voters observe the parties' policies and vote. All individuals vote and voting is sincere. For the case $i = 1$, we define $\Psi_{Ah1}(\alpha_{h1})$ as the net payoff received by a voter with preference α_{h1} in jurisdiction 1 from party A relative to that of party B where $\Psi_{Ah1}(\alpha_{h1}) = \{v_{Ah1}(\alpha_{h1}, g_{A1}, g_{A2}) - \tau_{A1} + \sigma_{Ah1}\} - \{v_{Bh1}(\alpha_{h1}, g_{B1}, g_{B2}) - \tau_{B1} + \sigma_{Bh1}\}$.¹³ In a system of local governments, politicians pay for local public goods by setting a head tax on residents of their respective jurisdiction $i = \{1, 2\}$ with $\tau_{Ai} = \frac{qg_{Ai}}{n_i}$. In addition, the net payoff Ψ_{Ah1} also depends on stochastic factors determined by σ_{Ah1} and σ_{Bh1} . Voter type α_{h1} votes for party A in the local election of jurisdiction i if $\Psi_{Ah1}(\alpha_{h1}) > 0$; if $\Psi_{Ah1}(\alpha_{h1}) < 0$ he or she votes for party B, and the voter flips a fair coin if $\Psi_{Ah1}(\alpha_{h1}) = 0$.

In the third stage, in the local election of jurisdiction i , if there exists a majority of individuals voting for party A then $\Omega_i(\forall \alpha_{hi}: \Psi_{Ahi}(\alpha_{hi}) > 0) > 1/2$ then party A wins the election and g_{LAi}^* is implemented, otherwise party B wins and g_{LBi}^* is implemented.

The subgame perfect Nash equilibrium of the political process with local elections is constituted by policy platforms of candidates $J = \{1, 2\}$ seeking nomination from their parties in their respective jurisdictions, the nomination choices of party leaders $NS_{LPi} \forall P, \forall i$, that is to say, party leaders nominate a candidate that will run with the party label in each local election, and the individuals' choices of the vote in their respective local election. For a formal characterization of this equilibrium see definition 2 in the appendix. In this economy, all candidates of all parties

¹³ Notice that from the perspective of politicians in jurisdiction 1, the welfare of the voter associated with local spending is $v_{Ah1}(\alpha_{h1}, g_{A1}, g_{A2})$ and g_{A1} is a choice variable but g_{A2} is exogenous and determined by the policy maker of jurisdiction 2.

converge in their policy platforms since they maximize a continuous and strictly concave probability function of winning multiple elections based on a common system of beliefs and strategy policy set. Lemma 2 characterizes the equilibrium spending policies for this economy and Theorem 1 shows that if parties are integrated then democratic decentralization dominates democratic centralization.

Lemma 2 *Party integration and democratic decentralization lead to a set of Pareto efficient local public goods with and without spillovers $\mathbf{g}_L^* = [g_{L1}^*, g_{L2}^*]$. At the equilibrium, g_{L1}^*, g_{L2}^* , satisfy the following:*

$$[g_{L1}^*, g_{L2}^*] = \left[\frac{\alpha_1 + k_1 \alpha_2}{q}, \frac{\alpha_2 + k_2 \alpha_1}{q} \right] \quad (5)$$

Proof

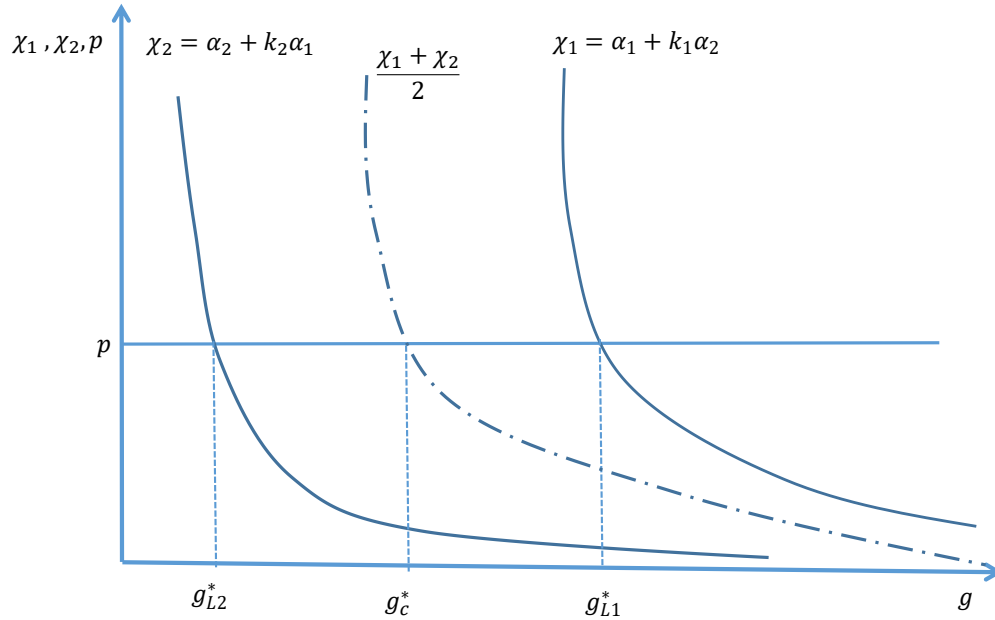
See the appendix.

Lemma 2 says that parties have electoral incentives to choose a policy that maximizes the joint probability of winning all local elections. Lemma 2 shows that this policy is equivalent to choosing a policy that maximizes an anonymous utilitarian nationwide social welfare function subject to the constraint that the local public good of the jurisdiction is financed by the residents of the jurisdiction. To see this, note the equivalence between the results in expressions (2) and (5) implying $g_{Li}^* = \hat{g}_i^* \forall i$. In (5), parties have electoral incentives to provide local public goods with and without spillovers at the point in which the nationwide marginal social gains, constituted by the sum of the marginal utility gains of all voters in jurisdiction i plus the marginal utility gains of all voters in the other jurisdiction (that are associated with the effect of spillovers from g_{Li}^*) are

equal to the nationwide social marginal costs of the public good. At $g_{Li}^* > 0 \forall i$, the nationwide aggregate surplus from local public goods is maximized. The heterogeneity of preferences of individuals across jurisdictions implies that $g_{L1}^* \neq g_{L2}^*$.

The decentralized provision of local public goods with and without spillovers is Pareto efficient because the political process is centralized and national parties seek to win elections in all jurisdictions and recognize that the inter-regional externalities of local public goods create an interdependence between the parties' expected votes in the elections of jurisdictions 1 and 2. Thus, parties have electoral incentives to coordinate local policies and propose spending policies that internalize the inter-regional spillovers in order to maximize the party's joint probability of winning the elections in all jurisdictions. In addition, if parties are integrated, local governments maximize the gains from differentiating local public goods according to the heterogeneous preferences of voters across jurisdictions.

Figure 1. Public Goods With Party Integration and Democratic (De)Centralization



To contrast the outcomes between democratic centralization and decentralization with party integration, in Figure 1, we show the allocation under democratic decentralization given by g_{L1}^*, g_{L2}^* and g_c^* which is the equilibrium policy under democratic centralization. In figure 1, p is the party's electoral costs of public goods, and χ_1 and χ_2 represent the party's electoral gains from the provision of public goods in jurisdictions 1 and 2 in economies with party integration and democratic decentralization. The expected votes in jurisdiction 1, χ_1 , is equivalent to the marginal utility gains of all voters in jurisdiction 1 plus the marginal utility gains of all voters of jurisdiction 2. A similar interpretation is given to χ_2 , and without loss of generality, we assume $\chi_1 > \chi_2$. In contrast, for an economy with party integration and democratic centralization, $\frac{\chi_1 + \chi_2}{2}$, represents the parties' nationwide expected proportion of votes in a nationwide election from policy platform g_c^* , which, is also equivalent to the average nationwide marginal utility gains of providing g_c^* .

Figure 1 shows the loss of welfare associated with a move from democratic decentralization to centralization under party integration. A move from g_{L2}^* to g_c^* implies that, from the perspective of residents of jurisdiction 2, public spending is too high which leads to a loss of welfare for residents of that jurisdiction. Similarly, a move from g_{L1}^* to g_c^* implies that, from the perspective of residents of jurisdiction 1, public spending is too low which also leads to a loss of welfare for residents of that jurisdiction. More formally, in Theorem 1, our *strong Decentralization Theorem* shows that democratic decentralization welfare dominates centralization even in the presence of spillovers.

Theorem 1 “Strong Decentralization Theorem”. *If parties are integrated, then the provision of local public goods with and without inter-regional spillovers by a system of democratically elected local governments welfare-dominates the democratically centralized provision.*

Proof

See the appendix.

Theorem 1 says that the nationwide welfare of voters is maximized when local public goods with and without inter-regional spillovers are provided by a system of local governments. This outcome is a stronger version of the decentralization theorem proposed by Oates (1972), since democratic decentralization dominates democratic centralization even if local public goods show inter-regional spillovers.

Note, first, that local public goods with and without inter-regional spillovers are Pareto efficient under both a central government and a system of local governments. Second, by matching the individuals' demand for heterogeneous public spending across jurisdictions, the decentralized provision maximizes the gains associated with inter-regional policy differentiation. Since the central government does not maximize the gains from differentiating local public goods to match local preferences, democratic decentralization is welfare superior to democratic centralization.

For reasons of space and mathematical simplicity, we have assumed that the central government is constrained to provide uniform local public goods. However, in Hankla et al (2019), we relax this assumption and identify conditions to show that our analysis is robust. Decentralization dominates centralization even when the central government can differentiate across jurisdiction in the provision of public goods.

7. *Non-Integrated Parties*

We turn our analysis to the provision of local public goods in a democracy with a majoritarian electoral system, single member jurisdictions, and non-integrated parties. In contrast to the case of integrated parties in which the decision to nominate candidates can rely on a small group of politicians (or even a single politician) inside a party, in the case of non-integrated parties the nomination process is determined by a primary election in which residents first vote to nominate a candidate while later in the general election voters elect a public official. In this setting, we study how the political institutions of primaries create incentives for parties to represent into policy platforms either the interests of a broad set of voters in the electorate or else the preferences of a minority coalition of voters.

Proponents of non-integrated parties argue that primary elections promote the political participation of voters and the representation of their interests in the policies eventually implemented by the government. However, the participation rules of primaries might actually limit both the voters' participation in elections and their effective political influence on policy design. Primary elections can be blanket, open and closed (see Kaufman, Gimpel and Hoffman 2003). In blanket primaries voters of any affiliation may vote in the primary and voters are given a ballot listing all candidates of all parties (see Keefe 1998). In contrast, in closed primaries only those voters affiliated with a party (probably partisan voters) can vote in the party's primary.

Blanket primaries provide the whole electorate with the opportunity to nominate candidates and parties have incentives to consider the whole distribution of voters' views while designing policy. However, in closed primaries candidates have electoral incentives to weigh (discount) heavily the preferences over policy of those voters who can (not) participate in the primary election. Hence, parties might have electoral incentives to implement the ideal policy of primary voters. This might be considered socially undesirable because, in this case, public spending does

not maximize the society's net fiscal incidence associated with public goods but the net surplus from public goods for a minority coalition of voters (the primary voters).¹⁴

The main results of this section are, first, that the *strong decentralization theorem* does not hold for economies with non-integrated parties. We also find that the specific institutions of primaries might (not) lead to the expected benefits of democratic decentralization. In particular, Oates' conventional decentralization theorem (which assumes no spillovers) holds only for economies with *blanket* primaries. If primaries are *closed* then the decentralization theorem, in general, does not hold. Local leaders will be too beholden to their co-partisans rather than to the overall interests of their constituents. These are important and novel results. In summary, in this section we clearly show that the political institutions of non-integrated parties may also matter considerably in determining the gains from decentralization.

8. *Non-Integrated Parties and Democratic Centralization*

In this section we analyze a model with non-integrated parties (an economy in which nominations are determined through a primary election) and democratic centralization (voters elect public officials only for the central government). Corresponding to the latter, the economy is also fiscally centralized. That is, local public goods are provided, and taxes levied by a single government that represents voters of all jurisdictions. The government finances its expenditures through a tax on residents of all jurisdictions.

¹⁴ The net fiscal incidence or fiscal surplus of local public goods reflects the following tradeoff: on the one hand, an increase of public spending leads to higher utility for voters (this is the marginal social benefit). On the other hand, higher spending requires higher taxes and lower consumption of private goods (this is the marginal social cost), see Martinez-Vazquez (1982).

We consider a sequential game of complete and perfect information with a primary and a general election to form the central government. In the first stage, two individuals, denoted by $j = \{1,2\}$ in each party seek the nomination of their party by declaring their binding policy platforms over public spending. In the second stage, all voters observe the candidates' policies, but only qualified voters vote in the primary election.¹⁵ In our economy, under a closed primary the right to vote is limited only to voters affiliated with the party. In contrast, in blanket primaries voters are given a ballot listing all candidates of all parties (see Keefe 1998) and voters vote in the primary of each party.

To model the right to vote in primary elections, and for simplicity of the analysis, we assume *Nature* moves and selects a set of voters who can participate in the primary. Hence, in each jurisdiction, *Nature* selects the distribution of qualified voters to participate in the primary of party P .¹⁶ The distribution of qualified voters selected by *Nature* is given by $\{\alpha_{P1i}, \alpha_{P2i} \dots \dots \alpha_{Pm_i i}\}$ where α_{phi} identifies a voter with some preference for public goods, the subscript P shows the label of the party of the primary in which the voter will be participating and qualified voters are indexed by $h = 1,2 \dots m_i$ where m_i is the number of voters participating in the primary election of jurisdiction i .¹⁷ In a blanket primary, all voters are given a ballot listing all candidates of all parties (see Keefe 1998) and voters vote in the primary of each party, therefore $m_i = n_i$ where n_i

¹⁵ The assumption that, in the first stage candidates announce a binding policy platform is for simplicity of the analysis and it ignores dynamic inconsistency issues such as the possibility that candidates might announce different policies in the primary and general elections to please, respectively, primary and general election voters.

¹⁶ In our model, the move by *Nature* could be interpreted as a move by an exogenous player with no strategic interest in the outcome of the game.

¹⁷ In the real world, the number of primary voters participating in the closed primary of each party will be different. Here, however, we assume that m_i voters participate in the primary of each party. This assumption helps to simplify our models, and, at the same time, it does not change the main results of our analysis.

is the number of residents of jurisdiction i . In a closed primary, only $m_i < n_i$ qualified voters participate in the primary of one party.

In the third stage of the game, the candidate who receives the majority of votes across all jurisdictions wins the primary and obtains the nomination of his or her party. In the fourth stage, the general election takes place and all voters in the electorate vote from the set of *nominated* candidates to elect a public official. Voting is sincere at the different stages of the electoral contest. In the fifth stage, votes are counted and the party with a majority of votes wins the general election, forms the government and implements its proposal on public spending.

In the first stage of the game, candidates $j = \{1,2\}$ of party $P = \{A, B\}$ announce spending platforms to maximize, Φ_{jP} , which is the candidates' *joint* probability of winning the nationwide primary and general elections. We assume Φ_{jP} is continuous and strictly concave on spending on public goods. Hence, candidates of, let's say party A, propose a policy platform that is sequentially rational and, therefore, their policy platform must consider two different states that might be played next: candidates might compete in the general election against candidate 1 or 2 of party B.

Define $\Phi_{jP} = \int_{\underline{\rho}_0}^{\bar{\rho}_0} \int_{\underline{\rho}_1}^{\bar{\rho}_1} \int_{\underline{\rho}_2}^{\bar{\rho}_2} w_{jP}(\rho_{jP0}, \rho_{jP1}, \rho_{jP2}) d\rho_{jP0} d\rho_{jP1} d\rho_{jP2}$ as the joint cumulative probability that candidate j of party P wins the primary and general elections where $w_{jP} = \partial\Phi_{jP}/\partial\rho_{jP0}\partial\rho_{jP1}\partial\rho_{jP2} > 0$ is the joint probability distribution function and ρ_{jP0} is the nationwide plurality in the primary election of candidate j of party P versus that of the other candidate running in the primary and ρ_{jP1} is the nationwide plurality in the general election for the state in which candidate j of party P runs against candidate 1 of the competing party. A similar

interpretation is given to ρ_{jP2} . In each case, the pluralities ρ_{jPs} are defined in the interval $\rho_{jPs} \in [\underline{\rho}_s, \bar{\rho}_s]$ for $s = \{0,1,2\}$.^{18,19}

For the case of party $P = A$, each of the nationwide pluralities are given by the difference of the expected votes between party A, ϕ_{jAs} , and party B, ϕ_{jBs} , that is, $\rho_{jAs} = \phi_{jAs} - \phi_{jBs}$ for the states $s = \{0,1,2\}$ and the nationwide sum of expected votes in the primary (for the case $s = 0$) is $\phi_{jA0} + \phi_{jB0} = m_1 + m_2$ and in the general election is $\phi_{jAs} + \phi_{jBs} = n_1 + n_2$ for $s = 1,2$. The expected votes of candidate j of party A in the nationwide primary election is $\phi_{jA0} = \sum_{i=1}^2 \phi_{jA0i}$, and $\phi_{jA1} = \sum_{i=1}^2 \phi_{jA1i}$ is the expected votes of candidate j of party A in the general election in the event this candidate faces candidate 1 of party B and $\phi_{jA2} = \sum_{i=1}^2 \phi_{jA2i}$ is the corresponding nationwide expected vote of candidate j of party A when this candidate faces candidate 2 of party B.

The expected vote of candidate j of party A in the primary of jurisdiction i is $\phi_{jA0i} = \sum_{h=1}^{m_i} F_{jA0i}(\varepsilon_{jA0i}) \forall i$, where $F_{jA0i}(\varepsilon_{jA0i}) = \partial^2 F_{jAi}(\Psi_{jA0i}, \Psi_{jA1i}, \Psi_{jA2i}) / \partial \Psi_{jA1i} \partial \Psi_{jA2i}$ is the continuous marginal probability that a voter type α_{Ahi} in jurisdiction i votes for candidate j of party A in the primary, and $F_{jAi} = F_{jAi}(\Psi_{jA0i}, \Psi_{jA1i}, \Psi_{jA2i})$ is the joint probability that the voter type α_{Ahi} votes for candidate j of party A in the primary and the general election. Similarly, the sum of expected votes for candidate j of party A in the general election from jurisdiction i is $\phi_{jAsi} = \sum_{h=1}^{n_i} F_{jAsi}(\varepsilon_{jAsi}) \forall i$ for $s = \{1,2\}$ where F_{jAsi} is the corresponding marginal probability that a

¹⁸ Note that ρ_{jPs} is the plurality of candidate j of party P in state $s = \{0,1,2\}$. When $s = 0$ then candidate j is competing in the primary election. In the event that candidate j obtains the nomination then this candidate must consider two different states that might be played in the general election: candidates might compete against candidate $s = 1$ or $s = 2$ of the opposite party.

¹⁹ It is not restrictive to assume $\rho_{jPs} \in [\underline{\rho}_s, \bar{\rho}_s]$ since, in general, this interval can take a wide range of values if $\underline{\rho}_s \rightarrow -\infty$ and $\bar{\rho}_s \rightarrow +\infty$.

voter type α_{hi} in jurisdiction i votes for j of party A if he or she faces candidates $s = 1$ or $s = 2$ of party B in the general election.²⁰

The joint probability that a voter with preference α_{Ahi} in jurisdiction i votes for candidate j of party A in the primary and the general election is $F_{jAi}(\Psi_{jA0i}, \Psi_{jA1i}, \Psi_{jA2i})$, which is assumed to be continuous and non-decreasing with Ψ_{jA0i} , Ψ_{jA1i} , and Ψ_{jA2i} . Recall Ψ_{jA0i} is the individual's net payoff for voter with preference α_{hi} in jurisdiction i from voting for candidate j of party A instead of the competing candidate of the same party in the primary, and Ψ_{jA1i} and Ψ_{jA2i} are the net payoffs of voting for candidate j of party A instead of candidates 1 or 2 of party B in the general election. If $\Psi_{jA0i} > 0$, $\Psi_{jA1i} > 0$, $\Psi_{jA2i} > 0$, a voter type α_{Ahi} in jurisdiction i votes for candidate j of party A in the nationwide primary and general elections.

We characterize the electoral equilibrium for this economy in definition 3 (see the appendix) and the equilibrium level for public goods (the main result of this section) in Lemma 3. Candidates facing sequential primary and general election face several tradeoffs. The first tradeoff is between the ideal policies of primary voters versus the socially optimal policies in each jurisdiction.²¹ Since candidates have incentives to choose a policy that maximizes their joint probability of winning the primary and general election, candidates need to choose a policy that reflects a compromise between the ideal policies of primary voters and a linear combination of the policies that maximize the nationwide welfare. If the ideal policies of primary voters are not middle

²⁰ Our notation means that a voter type α_{Ahi} has the right to vote in the primary of party A in the second stage of the game while the same voter, now characterized by α_{hi} has the right to participate in the general election (which is not conditioned to the election organized by a certain party) that takes place in the fourth stage of the game.

²¹ A nationwide general election gives voting rights to residents of all jurisdictions. For this reason, in our economy, the ideal policies of general election voters are equivalent to the socially optimal policies.

of the road policies, and primary voters have a significant electoral weight, then candidates might have incentives to design polarized policies with too much or too little government spending.

The second tradeoff that candidates also need to take into account is the distribution of votes in each jurisdiction. Hence, there is a tradeoff between the ideal policy of voters of jurisdiction 1 versus the policy of voters of jurisdiction 2. This particular tradeoff might also create incentives for politicians to focalize local public goods and concentrate a significant amount of resources in jurisdictions that have significant electoral influence. If local public goods are focalized in some jurisdiction then the scope of net benefits of public goods might be greatly diminished (that is, local public goods might not be designed to benefit a broad portion of voters but to benefit a minoritarian coalition of voters).

For the analysis that follows, we define the following concepts: Let G_{jPc}^* be the equilibrium policy of candidate j of party P under non-integrated parties and democratic centralization. We also denote G_{P01}^* and G_{P02}^* as the ideal policies of voters participating in the nationwide primary election of party P in jurisdictions 1 and 2 and W_{jP01} and W_{jP02} are the weights, or relative electoral importance, that candidate j of party P assigns to the corresponding ideal policies of voters participating in the nationwide primary election in jurisdictions 1 and 2. For the analysis that follows $W_{jP0i} = \beta_{jP0i}$ $i = 1, 2$ where β_{jP0i} is the change in the nationwide plurality of votes that candidate j of party P expects to receive in the primary election from jurisdiction i .

In addition, \hat{g}_1^* and \hat{g}_2^* are the socially optimal policies in jurisdictions 1 and 2 and W_{jP1} and W_{jP2} are the weights that candidate j of party P assigns to the socially optimal policies in jurisdictions 1 and 2 (or equivalently the ideal policies of general election voters). For the analysis that follows W_{jP1} and W_{jP2} are the change in the nationwide plurality of votes that candidate j of

party P expects to receive in the general election from jurisdictions 1 and 2, respectively. Recall that in the nationwide general election, candidate j of party P runs against candidates 1 or 2 of the competing party. Hence the weight $W_{jP1} = \beta_{jP11} + \beta_{jP21}$ reflects the sum of β_{jP11} which is the expected plurality of votes from jurisdiction 1 when candidate j of party P runs against candidate 1 of the competing party and β_{jP21} which is the plurality of expected votes from jurisdiction 1 when this candidate runs against candidate 2 of the competing party. A similar interpretation is given to $W_{jP2} = \beta_{jP12} + \beta_{jP22}$ for the expected plurality of votes in the general election from jurisdiction 2. In what follows, Lemma 3 shows the equilibrium policy G_{jPc}^* :

Lemma 3 *For economies with non-integrated parties and democratic centralization, a candidate j of party P proposes a uniform and Pareto efficient local public good G_{jPc}^* in jurisdictions 1 and 2 satisfying the following:*

$$G_{jPc}^* = Y_{jPc} (W_{jP01} G_{P01}^* + W_{jP02} G_{P02}^* + W_{jP1} \hat{g}_1^* + W_{jP2} \hat{g}_2^*) \quad (6)$$

Where Y_{jPc} is a proportionality parameter given by:

$$Y_{jPc} = \frac{1}{2(\Theta_{jPc} + \Gamma_{jPc}) - \frac{Cov_0 + Cov_{GE}}{q}} \quad (7)$$

And

$$\Theta_{jPc} = \frac{m_1 + m_2}{(\beta_{jP01} m_1 + \beta_{jP02} m_2)} \quad \text{and} \quad \Gamma_{jPc} = \frac{n_1 + n_2}{(\beta_{jP11} + \beta_{jP21}) n_1 + (\beta_{jP12} + \beta_{jP22}) n_2} \quad (8)$$

Moreover, Cov_0 is a weighted covariance of the marginal probability of the vote of candidate j of party P in the primary election, $\frac{\partial F_{jP0hi}}{\partial \varepsilon_{jP0hi}}$, and $\frac{\partial \varepsilon_{jP0hi}}{\partial G_{jPc}}$ which is the change in welfare of each

individual voter due in jurisdiction i to a marginal change in G_{jPc}^* . In addition, $\frac{\partial \Phi_{jP}}{\partial \rho_{jP0}}$ is the candidate's marginal change in the probability of winning the primary election:

$$Cov_0 = \frac{\partial \Phi_{jP}}{\partial \rho_{jP0}} \sum_{i=1}^2 m_i Cov_{jP0i} \left(\frac{\partial F_{jP0hi}}{\partial \varepsilon_{jP0hi}}, \frac{\partial \varepsilon_{jP0hi}}{\partial G_{jPc}} \right) \quad (9)$$

And Cov_{jP0i} is the covariance of the marginal probability of voters in jurisdiction i for candidate j of party P in the primary and the change in welfare of each resident of jurisdiction i due to a marginal change in G_{jPc}^* .

Moreover, Cov_{GE} is a weighted covariance of the marginal probability of the vote of candidate j of party P in the general election, $\frac{\partial F_{jPshi}}{\partial \varepsilon_{jPshi}}$, and $\frac{\partial \varepsilon_{jPshi}}{\partial G_{jPc}}$ which is the change in the welfare of each resident in jurisdiction i due to a marginal change in G_{jPc}^* . Moreover, $\frac{\partial \Phi_{jP}}{\partial \rho_{jPs}}$ is the candidate's marginal change in the probability of winning the general election, hence:

$$Cov_{GE} = \sum_{s=1}^2 \frac{\partial \Phi_{jP}}{\partial \rho_{jPs}} \left(\sum_{i=1}^2 n_i Cov_{jPsi} \left(\frac{\partial F_{jPshi}}{\partial \varepsilon_{jPshi}}, \frac{\partial \varepsilon_{jPshi}}{\partial G_{jPc}} \right) \right) \quad (10)$$

And Cov_{jPsi} is the covariance of the marginal probability of voters in jurisdiction i for candidate j of party P in the general election and the change in welfare of each resident of jurisdiction i due to a marginal change in G_{jPc}^* .

Proof

See the appendix

Lemma 3 says that G_{jPc}^* is proportional to a linear combination of the ideal policies of all voters participating in the primary election of party P in jurisdictions 1 and 2 determined by $W_{jP01}G_{P01}^* + W_{jP02}G_{P02}^*$ and a linear combination of the socially optimal policies in jurisdictions 1 and 2 determined by $W_{jP1}\hat{g}_1^* + W_{jP2}\hat{g}_2^*$.²² The parameters $W_{jP01} > 0, W_{jP02} > 0$ represent the relative electoral influence of primary voters in jurisdictions 1 and 2. The higher W_{jP01} , the higher

²² Recall that the institution of democratic centralization induces parties to take into account the marginal benefits and costs of public goods of all residents of all jurisdictions, which explains why parties have electoral incentives to consider a linear combination of the socially optimal policies \hat{g}_1^*, \hat{g}_2^* given by $W_{jP1}\hat{g}_1^* + W_{jP2}\hat{g}_2^*$.

is the change in the expected number of votes that candidate j of party P expects to receive from residents of jurisdiction 1 in the primary election and candidates have electoral incentives to choose a policy closer to the ideal policy of residents of jurisdiction 1 who participate in the primary election. A similar interpretation is given to the weights of electoral influence of voters in the general election in jurisdictions 1 and 2 given by $W_{jP1} > 0, W_{jP2} > 0$.

Moreover, it is relevant to point out that the relative plurality gain between the primary and general election is determined by the relative values of W_{jP01}, W_{jP02} vis-à-vis W_{jP1}, W_{jP2} . For instance if $W_{jP01} + W_{jP02} > W_{jP1} + W_{jP2}$ then primary voters in jurisdictions 1 and 2 have a higher electoral weight than general election voters in jurisdictions 1 and 2. The higher the difference the higher is the electoral incentive for candidate j of party P to propose the ideal policy of the average nationwide primary voter. Another example is that in the case in which $W_{jP01} > W_{jP1}$ then primary voters in jurisdiction 1 have a higher electoral weight than general election voters in jurisdictions 1. In this case, parties weigh more heavily the preferences of primary voters in jurisdiction 1 versus general election voters of jurisdiction 1, and the electoral incentive is higher for candidates to focalize local public spending and propose the ideal policy of the average primary voter of jurisdiction 1.

The equilibrium level of G_{jPC}^* also depends on a proportionality parameter given by $Y_{jPC} = \frac{1}{2(\Theta_{jPC} + \Gamma_{jPC}) - \frac{Cov_0 + Cov_{GE}}{q}}$ which depends on the cost of producing public goods q and the parameters Θ_{jPC} and Γ_{jPC} that explain how the distribution of electoral costs associated with the tax burden are aggregated through the political process. That is to say, the parameters Θ_{jPC} and Γ_{jPC} show the distribution of burden of costs of producing public goods among primary and general election voters in jurisdictions 1 and 2. In particular, the higher is Θ_{jPC} the higher is the electoral cost

associated with the burden of financing public goods for primary voters in jurisdictions 1 and 2. Similarly, the higher is Γ_{jPC} the higher is the burden of costs of local public goods for general election voters in jurisdictions 1 and 2. In both cases, increases in Θ_{jPC} and Γ_{jPC} lead to a lower provision of G_{jPC}^* .

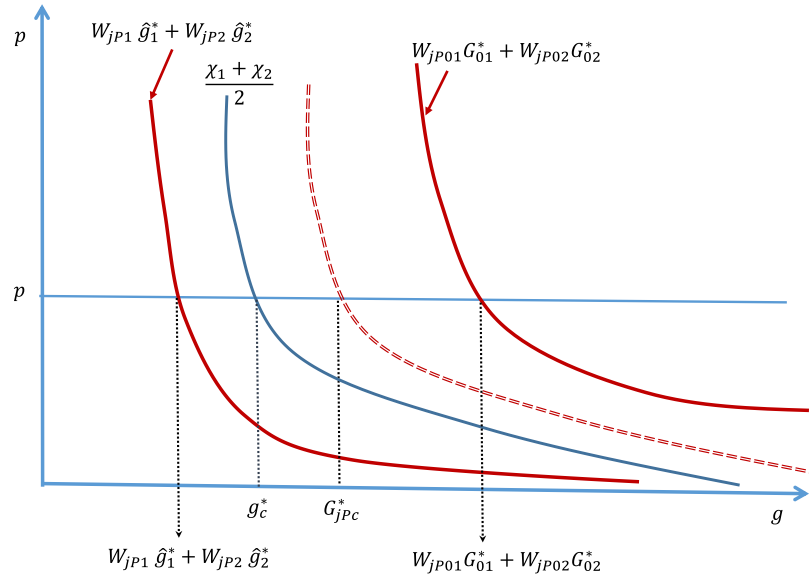
In addition, Cov_{GE} is a weighted covariance between the marginal probability of voting for candidate j of party P from individual type α_{hi} in jurisdiction i in the general election and the change in the well-being of each voter from an increase in the provision of the local public good. Candidates will have electoral incentives to increase the size of G_{jPC}^* when Cov_{GE} is positive; that is, when the electorate is constituted by voters who simultaneously have higher than average marginal probabilities of voting for candidate j of party P and higher than average demands for local public goods. A similar interpretation is given to Cov_0 which is a weighted covariance between the marginal probability of voting for candidate j of party P from individual type α_{phi} in jurisdiction i in the primary election of party P and the change in wellbeing of the voter from an increase in the provision of the local public good.

It is relevant to point out that positive and negative values of Cov_{GE} and Cov_0 tend to move the design of public spending from the center of the distribution of ideal policies of general and primary voters towards a policy that is polarized, and leading towards too much government spending if Cov_{GE} and Cov_0 are positive, and too little government spending if Cov_{GE} and Cov_0 are negative.

It should be clear from our previous discussion that different political institutions such as party integration and democratic decentralization lead to different processes of aggregating the conflicting preferences of voters into policy platforms. For this reason, the equilibrium level of

government spending in party integrated versus non-integrated regimes will be, in general, different. To provide a contrast for the results of this section, figure 2 shows the marginal electoral benefits for parties from public goods in economies with party integration and democratic centralization which leads to a policy given by g_c^* (see Lemma 2 and the curve given by $\frac{\chi_1 + \chi_2}{2}$) and the marginal electoral benefits when parties are non-integrated and there is democratic centralization (see the dashed red line) which leads to policy G_{jPc}^* .

Figure 2. Integrated versus Non-Integrated Party Systems and Democratic Centralization



In figure 2, the marginal electoral benefits for parties in the primary and general election are defined, respectively, by the marginal utility of the public good of the nationwide weighted average primary voter, $W_{jP01}G_{01}^* + W_{jP02}G_{02}^*$, and the marginal utility of the nationwide weighted average general voter $W_{jP1}\hat{g}_1^* + W_{jP2}\hat{g}_2^*$. For purposes of exposition, figure 2 shows the case in which the ideal policy of the weighted average voter of the general election in systems with non-integrated parties, $W_{jP1}\hat{g}_1^* + W_{jP2}\hat{g}_2^*$, is smaller than the ideal policy of the average voter in systems with integrated parties, $\frac{\chi_1 + \chi_2}{2}$, which, in turn, is also assumed to be smaller than the ideal

policy of the weighted average primary voter. That is to say, $W_{jP1} \hat{g}_1^* + W_{jP2} \hat{g}_2^* < \frac{\chi_1 + \chi_2}{2} < W_{jP01} G_{P01}^* + W_{jP02} G_{P02}^*$. Therefore, figure 2 shows the particular case in which parties have incentives to select more government spending in systems with non-integrated parties relative to the equilibrium policy of systems with integrated parties, that is to say, $G_{jPc}^* > g_c^*$.

9. *Non-Integrated Parties and Democratic Decentralization*

In this section, we characterize the provision of local public goods for an economy with sequential primary and general elections and democratic decentralization (voters elect public officials at the local level). The structure of the game is easily extended from our previous discussion: Local public goods are chosen by the local government and expenditure is financed by a uniform tax on residents in each jurisdiction. To form local governments, primary and general elections take place in each jurisdiction.

In the first stage, candidates $j = \{1,2\}$ of each party in jurisdiction i , announce policy platforms that maximize the joint probability of candidate j of party $P = \{A, B\}$ of winning the local primary and general elections $\Phi_{jPi} = \Phi_{jPi}(\rho_{jP0i}, \rho_{jP1i}, \rho_{jP2i})$ that depends on the plurality of the candidate in states $s = 0,1,2$ given by ρ_{jP0i} , ρ_{jP1i} and ρ_{jP2i} . In the second stage, residents of jurisdiction i observe the candidates' policies but only qualified voters vote in the primary election. *Nature* selects the set of qualified voters who can participate in the primary of each party. In the third stage of the game, the candidate who receives the majority of votes in the jurisdiction wins the nomination of his or her party. In the fourth stage, the general election takes place and all voters of the jurisdiction vote from the set of *nominated* candidates to elect a public official. Voting is sincere at the different stages of the electoral contest. In the fifth stage, votes are counted and

the party with a majority of votes wins the general election, forms the local government and implements its proposal on public spending in the jurisdiction.

For the analysis that follows, we define the following concepts: Let G_{jPi}^* be the equilibrium policy of candidate j of party P in jurisdiction i for a democracy with non-integrated parties and democratic decentralization. We also denote G_{0i}^* and G_i^* as the ideal policies of voters participating, respectively, in the primary and in the general election of jurisdiction i and W_{jP0i} and W_{jPsi} are the electoral weights of primary and general election voters in jurisdiction i . For the analysis that follows, W_{jP0i} is the change in the plurality of votes that candidate j of party P expects to receive in the primary election, and $W_{jPi} = \beta_{jP1i} + \beta_{jP2i}$ is the change in the nationwide plurality of votes that candidate j of party P expects to receive in the local general election.

For a formal characterization of the electoral game see definition 4 in the appendix, and Lemma 4 provides a general characterization of local public goods for economies with blanket and closed primary elections. Lemma 5 shows a special case of Lemma 4 in which nominations are conducted through a blanket primary election. Theorem 2 shows that the *strong decentralization theorem* is not satisfied for economies with non-integrated parties and local elections, and Theorem 3 shows that Oates' conventional decentralization theorem holds in the case of blanket primaries but fails to hold in the case of closed primaries.

Lemma 4 *For economies with non-integrated parties and democratic decentralization, a candidate j of party P in jurisdiction i proposes a local public good G_{jPi}^* satisfying the following:*

$$G_{jPi}^* = Y_{jPi} (W_{jP0i} G_{P0i}^* + W_{jPi} G_i^*) \quad (11)$$

Where Y_{jPi} is a proportionality parameter given by:

$$\Upsilon_{jPi} = \frac{1}{(\Theta_{jPi} + \Gamma_{jPi}) - \frac{\psi_{jP0i} + \psi_{jPG Ei}}{q}} \quad (12)$$

And

$$\Theta_{jPi} = \frac{\beta_{jP0i} m_i}{n_i} \quad \text{and} \quad \Gamma_{jPi} = \beta_{jP1i} + \beta_{jP2i} \quad (13)$$

Moreover, ψ_{jP0i} is a weighted covariance of the change in the marginal probability of the vote of candidate j of party P in the local primary election of jurisdiction i , $\frac{\partial F_{jP0hi}}{\partial \varepsilon_{jP0hi}}$, and $\frac{\partial \varepsilon_{jP0hi}}{\partial G_{jPi}}$ which is the change in welfare of each individual voter due to a marginal change in G_{jPi}^* . In addition, $\frac{\partial \Phi_{jPi}}{\partial \rho_{jP0}}$ is the candidate's change in the marginal probability of winning the primary election in jurisdiction i , where

$$\psi_{jP0i} = \frac{\partial \Phi_{jPi}}{\partial \rho_{jP0i}} m_i \text{Cov}_{jP0i} \left(\frac{\partial F_{jP0hi}}{\partial \varepsilon_{jP0hi}}, \frac{\partial \varepsilon_{jP0hi}}{\partial G_{jPi}} \right) \quad (14)$$

Where Cov_{jP0i} is the covariance of the marginal probability of voters in jurisdiction i for candidate j of party P in the primary and the change in welfare of each resident of jurisdiction i due to a marginal change in G_{jPi}^*

And define $\psi_{jPG Ei}$ is a weighted covariance of the change in the marginal probability of the vote of candidate j of party P in the local general election of jurisdiction i , $\frac{\partial F_{jPshi}}{\partial \varepsilon_{jPshi}}$, and $\frac{\partial \varepsilon_{jPshi}}{\partial G_{jPi}}$ is the change in the welfare of voter due to a change in G_{jPi}^* . Moreover, $\frac{\partial \Phi_{jPi}}{\partial \rho_{jPsi}}$ is the candidate's change in the marginal probability of winning the general local election in jurisdiction i , hence

$$\psi_{jPG Ei} = \sum_{s=1}^2 \frac{\partial \Phi_{jPi}}{\partial \rho_{jPsi}} n_i \text{Cov}_{jPsi} \left(\frac{\partial F_{jPshi}}{\partial \varepsilon_{jPshi}}, \frac{\partial \varepsilon_{jPshi}}{\partial G_{jPi}} \right) \quad (15)$$

Where Cov_{jPsi} is the covariance of the marginal probability of voters in jurisdiction i for candidate j of party P in the general election in the electoral state $s=1,2$ and the change in welfare of each resident of jurisdiction i due to a marginal change in G_{jPi}^*

Proof

See the appendix.

Lemma 4 says that sequential elections with a primary and a general election along with democratic decentralization create incentives for candidates of all parties to select a policy G_{jPi}^* that is proportional to a linear combination of the ideal policy of all *participating* primary voters of jurisdiction i determined by G_{P0i}^* and the ideal policy of all residents of the general election of the jurisdiction determined by G_i^* .²³ The parameters $W_{jP0i} > 0$ and $W_{jPi} > 0$ represent the relative electoral influence of primary and general election voters in jurisdiction i . Hence, the higher W_{jP0i} the higher is the change in the number of votes that candidate j of party P expects to receive in the primary of jurisdiction i and candidates have incentives to choose a policy closer to the ideal policy of primary voters of that jurisdiction.

A similar interpretation is given to the weight of electoral influence of voters participating in the general election, given by $W_{jPi} > 0$. Hence, the higher $W_{jPi} > 0$ the higher is the change in the expected number of votes from general election voters in the jurisdiction and candidates have incentives to choose a policy closer to the ideal policy of general election voters of the jurisdiction, which is given by G_i^* . Hence the relative electoral influence of primary versus general election votes are determined by the relative values of W_{jP0i} and W_{jPi} . If $W_{jP0i} > W_{jPi}$ then primary voters of jurisdiction i have a higher electoral weight than general election voters in the jurisdiction and the stronger is the electoral incentive for candidates to propose the ideal policy of the average local primary voter of the jurisdiction. Of particular interest is the case of closed primaries in which primary voters might have significant electoral weight (that is $W_{jP0i} > W_{jPi}$) and therefore, local

²³ The ideal policy of general election voters maximizes the net fiscal incidence from local public goods for the average voter in the jurisdiction, therefore $G_i^* = \sum_{h=1}^{n_i} \alpha_{hi}/q$. And the ideal policy of primary election voters maximizes the net fiscal incidence of public goods for the average primary voter, therefore $G_{P0i}^* = \sum_{h=1}^{m_i} \alpha_{phi}/q$. Note that m_i is the number of participating primary voters and n_i the number of general election voters. In a closed primary $m_i < n_i$ and in a blanket primary $m_i = n_i$.

elections might not lead to middle of the road policies but polarized governments with too much or too little government spending.

As we discussed in our previous section, the equilibrium level of $G_{jp_i}^*$ also depends on a proportionality parameter given by $Y_{jp_i} = \frac{1}{(\Theta_{jp_i} + \Gamma_{jp_i}) - \frac{\psi_{jp_{0i}} + \psi_{jp_{GEi}}}{q}}$ which depends on the cost of producing public goods q and the parameters Θ_{jp_i} and Γ_{jp_i} that explain how the distribution of electoral costs associated with the tax burden are aggregated through the political process. Higher levels of Θ_{jp_i} and Γ_{jp_i} lead to higher electoral costs for primary and general election voters associated with the burden of raising taxes in the jurisdiction and lead to a lower provision of $G_{jp_i}^*$.

In addition, $\psi_{jp_{GEi}}$ and $\psi_{jp_{0i}}$ are weighted covariances between the marginal probability of voting for candidate j of party P in the general and primary election of jurisdiction i and the change in the well-being of each voter from an increase in the provision of the local public good. As we discussed in the previous section, positive and negative values of $\psi_{jp_{GEi}}$ and $\psi_{jp_{0i}}$ tend to move the design of public spending from the center of the distribution of ideal policies of general and primary voters in the jurisdiction towards a policy that is polarized and leading towards too much government spending if $\psi_{jp_{GEi}}$ and $\psi_{jp_{0i}}$ are positive and too little government spending if $\psi_{jp_{GEi}}$ and $\psi_{jp_{0i}}$ are negative.

In what follows, Lemma 5 distinguishes the equilibrium policies for an economy with non-integrated parties and blanket primaries.

Lemma 5 *For economies with non-integrated parties, blanket primaries and democratic centralization $G_{jp_c}^*$ is given by:*

$$G_{jPc}^* = \frac{\hat{g}_1^* + \hat{g}_2^*}{2} \quad (16)$$

Where \hat{g}_1^*, \hat{g}_2^* represent the Pareto efficient local public goods that maximize the social welfare gains from inter-regional policy differentiation.

For economies with non-integrated parties, blanket primaries and democratic decentralization

$G_{jPi}^* = G_i^*$ which is the ideal policy of residents of jurisdiction i and it is given by:

$$G_{jPi}^* = \frac{\sum_{h=1}^{n_i} \alpha_{hi}}{q} = \frac{\alpha_i}{q} \quad \text{for } i = 1, 2 \quad (17)$$

Proof

See the appendix.

Lemma 5 shows the spending policies for an economy with non-integrated parties, blanket primaries and democratic centralization and decentralization. Under these institutions, democratic centralization produces the ideal policy of the nationwide average general election voter, G_{jPc}^* , and under democratic decentralization the ideal policy of the average voter in each jurisdiction, G_{jPi}^* . These outcomes provide several insights: *first*, our model predicts that party integration with a unitary government (see Lemma 2 and condition 3) and party non-integration with blanket primaries and a unitary government (see Lemma 5 and condition 16) produce the same policies. That is $g_c^* = G_{jPc}^*$. However, under local elections, integrated and non-integrated parties with blanket primaries do not produce the same policies, that is, $g_{Li}^* \neq G_{jPi}^*$ (see the equilibrium policies of Lemma 2 and Lemma 5). Hence, the institutions of party integration versus non-integration produce different outcomes under national versus local elections.

Second, in terms of the efficiency properties and the degree of policy differentiation of the equilibrium for economies with non-integrated parties, the implications of Lemma 4 and Lemma 5 are the following: for economies with blanket primaries and democratic centralization, local public goods with and without spillovers are uniform and Pareto efficient (see condition 16 of Lemma 5). That is, blanket primaries create electoral incentives for candidates to recognize the nationwide distribution of benefits and costs of local public goods. Hence, local public goods with and without spillovers are Pareto efficient.

In addition, for economies with non-integrated parties, blanket primaries and democratic decentralization, local public goods with inter-regional spillovers *are not Pareto efficient*; in general, $G_{jpi}^* < \hat{g}_i^*$, for either type of primary, blanket or closed. That is, local sequential elections do not create incentives to internalize the benefits of local public goods across jurisdictions.

Third, if local public goods do not have spillovers, then the political institutions of the primary election determine whether the decentralized provision of local public goods is Pareto efficient or not. In our economy, in blanket primaries all voters participate in the primary and general election. Therefore, the average primary voter is also the average general election voter in each jurisdiction. Since parties select the ideal policy of the average voter in the jurisdiction, local public goods *without* externalities are Pareto efficient. However, for closed primaries, local public goods *are not Pareto efficient* if the electoral weight of general election voters is sufficiently low, that is if $W_{jpi} \rightarrow 0$. In this case, local governments provide a local public good that maximizes the well-being of the average primary voter of jurisdiction i . This means that local spending is focalized to benefit a local minoritarian coalition of voters in the jurisdiction and therefore public goods are not Pareto efficient.

Fourth, the political institutions of primaries also produce different results as to the ability of local governments to maximize the gains from inter-regional policy differentiation. Under blanket primary elections, democratic decentralization produces the ideal policy of the average voter in each jurisdiction and maximizes the gains of the society from inter-regional policy differentiation. If the primary is closed, then the gains from policy differentiation can be maximized only if the electoral weight of primary voters is sufficiently low (that is if $W_{jP0i} \rightarrow 0$, see condition 11 in Lemma 4). If, in contrast, the electoral weight of primary voters is sufficiently high, local governments provide a local public good that maximizes the well-being of the average primary voter of jurisdiction i , which implies that the political process fails to take into account the full extent of preferences of all residents in each jurisdiction. As a result, the extent of inter-regional policy differentiation achieved by local governments is sub-optimal.

10. Non-Integrated Parties with National and Local Elections and Oates' Decentralization Theorem.

In what follows we develop the comparative analysis between national elections and local elections in economies with non-integrated parties. Theorem 2 shows that the *strong decentralization theorem* does not hold, and Theorem 3 shows that Oates' decentralization theorem is satisfied, when primaries are blanket but not when primaries are closed.

Theorem 2 *In democracies with non-integrated parties and blanket primaries, the strong decentralization theorem does not hold but the conventional decentralization theorem holds.*

Proof

See the appendix

Lemma 5 shows that, in systems with non-integrated parties and blanket primaries, parties propose the size of public spending that maximizes a unanimous utilitarian social welfare function of local residents in *each* jurisdiction. That is, parties propose the ideal policy of the average voter of the jurisdiction and, therefore, the resulting policies are Pareto efficient but only for local public goods without spillovers. With spillovers, local spending is not Pareto efficient and, therefore, the *strong decentralization theorem* does not hold for these economies. In addition, under blanket primaries, the local provision of public goods maximizes the welfare gains associated with differentiating local policy to the heterogeneous preferences of voters across jurisdictions.

If the government is democratically centralized and uses blanket primaries, local public goods with and without spillovers are uniform and Pareto efficient. In this case, local public goods reflect the ideal size of public spending for the average voter of all jurisdictions. Lastly, Theorem 2 shows that if local public goods do not display spillovers, democratic decentralization dominates centralization because the Pareto efficient and heterogeneous provision of local public goods in a democratically decentralized system maximizes the welfare gains from policy differentiation, while democratic centralization leads to suboptimal inter-regional policy differentiation. Consequently, the conventional decentralization theorem of Oates (1972) holds for democracies with non-integrated parties and blanket primaries.

Theorem 3 *The strong and the conventional decentralization theorems do not hold in democracies with non-integrated parties and closed primaries.*

Proof

See the appendix.

In general, the *strong* and the conventional decentralization theorems do not hold for a democracy with closed primaries.²⁴ As we mentioned before, sequential elections induce candidates to propose a policy that reflects a compromise between the ideal policy of the average voter in the primary and the ideal policy of the average voter in the general election (see expressions of Lemmas 3 and 4). In the case of closed primaries, it is likely that minority coalitions could have strong electoral influence over local governments that is translated into policies closer to their preferences. In this case, electoral competition might produce extreme policy positions (too much or too little local public spending) instead of the policy that maximizes the net fiscal incidence of local public goods for the society.

For instance, in condition 6 of Lemma 3, if $W_{jP1} = W_{jP2} \cong 0$, then local elections produce a size of public spending in the vicinity of the ideal policy of primary voters in each jurisdiction, instead of the ideal policy of all residents in each jurisdiction.²⁵ Simultaneously, we can find parametric values of the weight of the electoral influence of primary voters—in which democratic centralization averages the political influence of local coalitions of voters—to produce a more

²⁴ A comparison of (2) and (6) shows that the optimal policy for candidates in a nationwide sequential election with closed primaries is, in general, different from the spending policy that maximizes the fiscal surplus for all residents in the economy. A similar conclusion is reached for local elections in a federation (see conditions 2 and 12). As a result, the *strong* and the conventional decentralization theorems do not hold in a system of non-integrated parties and closed primaries.

²⁵ In particular, significantly low values of W_{jP1} and W_{jP2} imply that the marginal change of the joint probability of winning the primary and the general election due to a marginal gain from the plurality of the primary is sufficiently high (maybe because the candidate who wins the primary would win with almost certainty the general election) and therefore candidates will select the ideal policy of a minority of the electorate (i.e. the ideal policy of the average voter among the primary voters).

moderate provision of local public goods that could be welfare superior to the less moderate provision of public goods in the democratically decentralized system of local governments. This is likely true for an economy with many localities in which the political influence of local coalitions is significant over local governments, but the influence of local minority groups fades away in the nationwide election. In this setting, the provision of a uniform public good by a democratically centralized government could be welfare superior to the differentiated but polarized provision produced by a system of democratically decentralized local governments.

11. Comparison of Party Integration versus Blanket and Closed Primaries

In this section we present a summary of the comparative analysis of the different political institutions studied in this paper and discuss some of the main drivers of our results. As Table 1 shows, we find that democratic decentralization and party integration produce the best policy outcomes. This is because local public goods with and without externalities are Pareto efficient and the welfare associated with matching local public goods according to the heterogeneity of voter preferences is maximized. Democratic centralization also produces Pareto efficient public goods, but the degree of policy differentiation is suboptimal. Therefore, under party integration, democratic decentralization dominates centralization and the conventional (compare cells I.A and II.A) and strong (compare cells I.B and II.B) decentralization theorems are satisfied.

Table 1. Summary of Comparative Analysis for Integrated and Non-Integrated Parties

	A. Local Public Goods Without Externalities		B. Local Public Goods With Externalities	
	Are LPGs Differentiated According to Heterogeneous Preferences?	Are LPGs Pareto Efficient?	Are LPGs Differentiated According to Heterogeneous Preferences?	Are LPGs Pareto Efficient?
I. Democratic Decentralization & Party Integration	YES	YES	YES	YES
II. Democratic Centralization & Party Integration	NO	YES	NO	YES
III. Does the Decentralization Theorem Hold?	YES			
IV. Does the <i>Strong</i> Decentralization Theorem Hold?			YES	
V. Democratic Decentralization & Blanket Primaries	YES	YES	NO	NO
VI. Democratic Centralization & Blanket Primaries	NO	YES	NO	YES
VII. Does the Decentralization Theorem Hold?	YES			
VIII. Does the <i>Strong</i> Decentralization Theorem Hold?			NO	
IX. Democratic Decentralization & Closed Primaries	IN GENERAL NO	IN GENERAL NO	IN GENERAL NO	IN GENERAL NO
X. Democratic Centralization & Closed Primaries	IN GENRAL NO	IN GENERAL NO	IN GENERAL NO	IN GENERAL NO
XI. Does the Decentralization Theorem Hold?	IN GENERAL NO			
XII. Does the <i>Strong</i> Decentralization Theorem Hold?			IN GENERAL NO	

Table 1 also compares policy outcomes for economies with non-integrated parties and blanket primaries. For this type of economy, the decentralization theorem holds because local public goods *without* externalities are Pareto efficient and the welfare associated with matching public goods to the heterogeneity of voter preferences is maximized. Democratic centralization also produces Pareto efficient public goods, though the degree of policy differentiation is suboptimal. Therefore, democratic decentralization dominates centralization (compare cells V.A and VIA) but the strong decentralization theorem is not satisfied (compare cells V.B and VIB) since local public goods under democratic decentralization are not Pareto efficient.

Finally, Table 1 shows that if parties are non-integrated with closed primaries then, in general, neither the conventional nor the strong decentralization theorems hold. As we mentioned before, sequential elections with closed primaries induce candidates to propose a policy that reflects a compromise between the ideal policy of the average voter in the primary and the ideal policy of the average voter in the general election. As a result, it is likely that minority coalitions in the electorate could have strong electoral influence over local and central governments that is translated into policies closer to their preferences. In this case, electoral competition might produce extreme policy positions (with too much or too little local public spending) instead of the policy that maximizes the net fiscal incidence of local public goods for the society. Hence, local public goods might not be Pareto efficient and the degree of inter-regional policy differentiation is likely to be suboptimal.

It is important to mention that the main drivers of policy outcomes shown in Table 1 are the incentives of party organization (i.e. whether parties are integrated or non-integrated) and our assumption that parties design policy to maximize votes in the election. Most of the previous literature has overlooked the role of party organization, and the outcomes presented in this paper

could change considerably if instead of seeking votes, parties design policies to maximize the preferences of a representative coalition of voters (or activists) inside parties (for analysis on this line of research see Wittman 1973,1983 and Roemer 2006). Another relevant assumption is that we employ a probabilistic voting model. Under perfect information on the distribution of voters' preferences and how policy platforms of parties translate into votes, parties have incentives to focalize policies towards the median voter. In the case of preference heterogeneity inside each jurisdiction, local public goods provided by local and central governments might not be Pareto efficient and the welfare comparison between democratic decentralization versus centralization could change considerably.

12. Conclusion

In this paper, we analyze how political institutions condition the outcomes of fiscal federalism. We develop a formal extension and refinement of Oates' (1972) decentralization theorem, which has provided the basis for much of the research on fiscal decentralization over the past five decades. In particular, we provide a political economy analysis of local public good provision by incorporating the joint influence of democratic (de)centralization and party (non)integration. Our theory generates a number of new insights. First, we show that for democracies with integrated parties, a system of elected local governments welfare-dominates a centralized government even if local public goods show inter-jurisdictional spillovers. We call this result the *strong decentralization theorem*. When our new theorem is satisfied, the conventional decentralization theorem developed by Oates (1972), where local public goods do not show spillovers, is also satisfied. Indeed, an important implication of our theoretical model is that the combination of democratic decentralization and party integration tends to produce the highest welfare gains from the provision of public goods. Democratic decentralization ensures that local

governments are responsive to the desires of their constituents while party integration incentivizes local leaders to pay for goods that may have spillover benefits.

Second, we show that the *strong decentralization theorem* does not hold for countries with non-integrated parties, whether they have blanket or closed primaries. For countries with non-integrated parties *and blanket primaries*, Oates' conventional decentralization theorem (without spillovers) holds since the political institutions of blanket primaries promote the participation of voters and the representation of their preferences into policies. However, countries with non-integrated parties and closed primaries, even the conventional decentralization theorem does not generally hold. The participation rules of closed primaries induce parties to propose a policy that reflects a compromise between the ideal policy of a weighted average voter of the primary and the ideal policy of a weighted average voter of the general election. This last finding has particularly important implications for much of the past literature. In the case of countries with closed primaries, it is likely that minority coalitions (i.e. primary voters) will have strong political influence over central and local governments that is translated into policies closer to their preferences. When this happens, electoral competition will not necessarily produce Pareto efficient local public goods (even if local public goods do not show spillovers). Parties might adopt extreme policy positions with too much or too little local public spending compared with the policy that maximizes the surplus from the net fiscal exchange for the society as a whole. Another important consequence is that the degree of policy differentiation might be suboptimal and might not exhaust the gains from matching the preferences of voters across jurisdictions if primary elections are closed.

In summary, we show that creating locally elected governments can only be expected to improve public goods allocation either when parties are integrated or when there are no inter-

jurisdictional spillovers. Local governments controlled by non-integrated parties are not likely to provide efficient levels of public goods that spill over into adjacent jurisdictions. Therefore, the outcomes from fiscal and democratic decentralization are closely intertwined with the design of political parties. All of these institutions must be studied jointly if we are to understand better the factors that make local government and fiscal federalism produce superior outcomes.

References

- Bardhan, P. and D. Mookherjee. 2000. Capture and Governance at Local and National levels. *American Economic Review* 90(2), 135-139.
- Bardhan, P. and D. Mookherjee. 2002. Relative Capture of Local and Central Governments: An Essay in the Political Economy of Decentralization, Working Paper, Institute for Economic Development, Boston University, 1-45.
- Besley, T., and S. Coate. 2003. Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach. *Journal of Public Economics* 87: 2611-2637.
- Bordignon, M., L. Colombo, and U. Galmarini. (2008). Fiscal Federalism and Lobbying. *Journal of Public Economics*, 92(12), 2288-2301
- Carey, John M and Matthew Soberg Shugart. 1995. Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas. *Electoral Studies* 14(4): 417-439.
- Chhibber, Pradeep and Ken Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States*. New York: Princeton University Press.
- Coughlin, P.J., 1992. *Probabilistic Voting Theory*. Cambridge University Press.
- Enikolopov, Ruben and Ekaterina Zhuravskaya. 2007. Decentralization and Political Institutions. *Journal of Public Economics* 91: 2261-2290.
- Fabre, Elodie, Bart Maddens, Wilfried Swenden, and Robertas Pogorelis. 2005. Partis politiques nationaux en crise? Organisation des partis et décentralisation. Une comparaison de l'Espagne et du Royaume Uni. *Res Publica* 2005/1.
- Filippov, Mikhail, Peter C. Ordeshook, and Olga Shvetsova. 2004. *Designing Federalism: A Theory of Self-Sustainable Federal Institutions*. New York: Cambridge University Press.
- Keefe, William J. 1998. Parties, politics and public policy in America. Washington DC. CQ Press.

- Haggard, Stephan and Robert R. Kaufman. 1995. *The Political Economy of Democratic Transitions*. Princeton: Princeton University Press.
- Hankla, Charles R., Jorge Martinez-Vazquez, and Raul Ponce Rodriguez. 2019. *Local Accountability and National Coordination in Fiscal Federalism: A Fine Balance*. Edward Elgar Publishers. Series: Studies in Fiscal Federalism and State-Local Finance.
- Hallerberg, Mark and Patrik Marier. 2004. Executive Authority, the Personal Vote, and Budget Discipline in Latin American and Caribbean Countries. *American Journal of Political Science* 48 (3): 571-587.
- Kaufman, K.M. , J.G Gimpel, and A.H. Hoffman 2003. A Promise Fulfilled? Open Primaries and Representation. *The Journal of Politics*, Vol. 65, No. 2, pp. 457-476.
- Lockwood, B. 2015. The Political Economy of Decentralization, in Ahmad, E. and G. Brosio, (Ed), *Handbook of Multilevel Finance*. Edward Elgar Publishing, (pp. 37-65).
- Lockwood, B. 2008. Voting, Lobbying and the Decentralization Theorem. *Economics & Politics*, 20(3), 416– 431.
- McKelvey, R.D and J.W. Patty. 2006. A Theory of Voting in Large Elections. *Games and Economic Behavior*, 57(1):155–180, 20
- Martinez-Vazquez, Jorge. 1982. Fiscal Incidence at the Local Level. *Econometrica* 50 (5): 1207-1218.
- Martinez-Vazquez, Jorge, Santiago Lago-Peñas, and Agnese Sacchi. 2016. The Impact of Fiscal Decentralization: A Survey. *Journal of Economic Surveys*, (on line version) 7 November 2016.
- Myerson, Roger B. 2006. Federalism and Incentives for Success of Democracy. *Quarterly Journal of Political Science* 1: 3-23.
- Nielson, Daniel L. 2003. Supplying Trade Reform: Political Institutions and Liberalization in Middle-Income Presidential Democracies. *American Journal of Political Science* 47 (3): 470-491.
- Oates, Wallace E. 1972. *Fiscal Federalism*. New York: Harcourt, Brace, Jovanovich.
- Oates, Wallace E. 2005. Towards a Second-Generation Theory of Fiscal Federalism. *Journal of International Tax and Public Finance* 12: 349-373.
- O'Halloran, Sharyn. 1994. *Politics, Process, and American Trade Policy*. Ann Arbor: University of Michigan Press.
- Persson, T. And G. Tabellini. 2002. *Political Economics: Explaining Economic Policy*. Cambridge, MA, The MIT Press.
- Persson, T. And G. Tabellini. 2003. *The Economic Effects of Constitutions, Munich Lectures in Economics*. Cambridge, MA, The MIT Press.

- Ponce-Rodriguez, Raul, Charles R. Hankla, Jorge Martinez-Vazquez, and Eunice Heredia-Ortiz. 2018. Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods. *Publius: The Journal of Federalism* 48 (4): 523-558.
- Roemer, J.E., (2006), *Political Competition Theory and Applications*. Harvard University Press, Cambridge Massachusetts.
- Riker, William H. 1987. *The Development of American Federalism*. Boston: Kluwer Academic Publishers.
- Roubini, Nouriel and Jeffrey Sachs. 1989. Government Spending and Budget Deficits: Industrial Democracies. Working Paper No. 2919. National Bureau of Economic Research, Cambridge, MA.
- Seabright, P. 1996. Accountability and Decentralisation in Government: An Incomplete Contracts Model. *European Economic Review*, 40 (1), 61- 89.
- Volden, Craig. 2004. Origin, Operation, and Significance: The Federalism of William H. Riker. *Publius: The Journal of Federalism* 34 (4): 89-107.
- Volkerink, Bjørn and Jakob de Haan. 2001. Fragmented Government Effects of Fiscal Policy: New Evidence. *Public Choice* 109: 221-242.
- Weingast, Barry R. 2009. Second Generation Fiscal Federalism: The Implications of Fiscal Incentives. *Journal of Urban Economics* 65: 279-293.
- Weingast, Barry R. 2014. Second Generation Fiscal Federalism: Political Aspects of Decentralization and Economic Development. *World Development* 53: 14-25.
- Wittman, D.A. (1973). Parties as Utility Maximizers. *American Political Science Review* 67: 490-498.
- _____(1983), Candidate Motivation: A Synthesis of Alternative Theories. *The American Political Science Review* 77(1): 142-157.