International Center for Public Policy

Working Paper 16-03

March 2016

Political Institutions and Federalism: A "Strong" Decentralization Theorem

Raúl A. Ponce-Rodríguez

Charles R. Hankla

Jorge Martinez-Vazquez

Eunice Heredia-Ortiz

Georgia State University | Andrew Young School of Policy Studies

Georgia State University | Andrew Young School of Policy Studies

International Center for Public Policy

Working Paper 16-03

Political Institutions and Federalism: A "Strong" Decentralization Theorem

Raúl A. Ponce-Rodríguez

Charles R. Hankla

Jorge Martinez-Vazquez

Eunice Heredia-Ortiz

March 2016

International Center for Public Policy
Andrew Young School of Policy Studies

The Andrew Young School of Policy Studies was established at Georgia State University with the objective of promoting excellence in the design, implementation, and evaluation of public policy. In addition to two academic departments (economics and public administration), the Andrew Young School houses seven leading research centers and policy programs, including the International Center for Public Policy.

The mission of the International Center for Public Policy is to provide academic and professional training, applied research, and technical assistance in support of sound public policy and sustainable economic growth in developing and transitional economies.

The International Center for Public Policy at the Andrew Young School of Policy Studies is recognized worldwide for its efforts in support of economic and public policy reforms through technical assistance and training around the world. This reputation has been built serving a diverse client base, including the World Bank, the U.S. Agency for International Development (USAID), the United Nations Development Programme (UNDP), finance ministries, government organizations, legislative bodies and private sector institutions.

The success of the International Center for Public Policy reflects the breadth and depth of the in-house technical expertise that the International Center for Public Policy can draw upon. The Andrew Young School's faculty are leading experts in economics and public policy and have authored books, published in major academic and technical journals, and have extensive experience in designing and implementing technical assistance and training programs. Andrew Young School faculty have been active in policy reform in over 40 countries around the world. Our technical assistance strategy is not to merely provide technical prescriptions for policy reform, but to engage in a collaborative effort with the host government and donor agency to identify and analyze the issues at hand, arrive at policy solutions and implement reforms.

The International Center for Public Policy specializes in four broad policy areas:

- Fiscal policy, including tax reforms, public expenditure reviews, tax administration reform
- Fiscal decentralization, including fiscal decentralization reforms, design of intergovernmental transfer systems, urban government finance
- Budgeting and fiscal management, including local government budgeting, performance-based budgeting, capital budgeting, multi-year budgeting
- Economic analysis and revenue forecasting, including micro-simulation, time series forecasting,

For more information about our technical assistance activities and training programs, please visit our website at http://aysps.gsu.edu/isp/index.html or contact us by email at gbotello1@gsu.edu

Political Institutions and Federalism: A "Strong" 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 Martinez-Vazquez
Department of Economics
Georgia State University

Eunice Heredia-Ortiz[1]
Development Alternatives Inc., DAI

In this article, we investigate how differences in the political institutions necessary for implementing decentralization reform may affect the efficiency and welfare properties of decentralization itself. We incorporate insights from political science and economics into a rigorous and formal extension of the influential "decentralization theorem" first developed by Oates in 1972. In our analysis, we go beyond Oates by producing a *strong decentralization theorem* that identifies the political conditions under which democratic decentralization dominates centralization even in the presence of interjurisdictional spillovers. More specifically, we find that beneficial outcomes for public service delivery will obtain when democratic decentralization (i.e. the creation of popularly elected sub-national governments) is combined with party centralization (i.e. the power of national party leaders to nominate candidates for sub-national office). We also find that the participation rules of primaries, whether *closed* or *open*, have important implications for the expected gains from decentralization. Most notably, we find that, when primaries are closed, even Oates' conventional decentralization theorem does not hold. In summary, our theory shows that political institutions matter considerably in determining the welfare gains of decentralization outcomes.
(JEL D61, D72, D78, H73, H75)

[1]The order of authors has been drawn randomly. We would like to thank Roger Myerson, Irfan Nooruddin, Craig Volden and Kyle Hanniman for their helpful feedback. Our thanks also go the International Center for Public Policy at Georgia State's Andrew Young School of Policy Studies for financing the coding. Finally, we would like to thank the members of the Deil Wright Awards committee for naming an earlier version of this work the best paper on federalism presented at the 2011 American Political Science Association Convention in Seattle.

1. Introduction and Background

Decentralization has become a prominent global trend; countries which have engaged in decentralization reforms include China, Indonesia, South Africa, India, the United Kingdom, and many others. 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).

As the process of decentralization has continued apace, however, 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 a benevolent welfare-maximizing government. While this 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. While these problems have been increasingly acknowledged and confronted in the "second generation" research on decentralization, there is still little systematic work concerning which political institutions lead to the social welfare gains expected from Oates' (1972) decentralization theorem.[2] Instead, this second generation of

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

scholarship, which relaxes the assumption of benevolent government, has tended to focus on the problems of assignment and soft budget constraints rather than on the relationship between specific political institutions and the provision of local public goods.

Most of the research that has highlighted the role of institutions in public goods delivery deals with 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 a wide variety of policy outcomes such as free trade, balanced budgets, energy conservation, etc. (see O'Halloran 1994, Nielson 2003, Haggard and Kaufman 1995, Hallerberg and Marier 2004, Roubini and Sachs 1989 and Volkerink and de Haan 2001).

Other scholars have examined the relationship between party organization on the one hand and decentralization on the other. For example, Chhibber and Kollman (2004) make the case that countries devolving more powers 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. Garman, Haggard, and Willis (2001) argue for the same relationship moving in the opposite direction; for them the decentralization of parties is likely to drive greater fiscal decentralization. By contrast, Eaton (2004) and Dickovick (2011) find that the choice to empower sub-national governments can be driven by the incentives of national party leaders.

Perhaps William Riker is the most prominent scholar to have taken up the relationship between decentralization and parties. He argued in his classic 1964 book that party centralization is among the most important determinants of federal centralization as a whole. Extending that argument to the United States, Riker contends in his 1987 book that the American "decentralized

party system is the main protector of the integrity of states in our federalism" (p. 221).[3] 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, arguing that regional and local elections provide opportunities for potential national candidates to prove themselves at the sub-national level.

Recent work on the political economy of fiscal federalism highlights the need for more systematic work on the impact of political institutions on the expected gains from decentralization.[4] For instance, in the legislative model of Besley and Coate (2003), local public goods are not necessarily 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 electoral incentives of politicians. Lockwood (2002) also 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. Bordignon, Colombo and Galmarini (2008) characterize conditions in which lobbying through campaign contributions induces a decentralized provision that is not Pareto efficient when politicians become too greedy.

[3] See also Volden 2004 for an excellent summary of Riker's thought on Federalism.
[4] For a recent survey of the political economy of decentralization see Lockwood (2015).

Despite the value of such work, there has been little focus on the political institutional conditions under which decentralization can deliver on its promises.[5] In this paper, we demonstrate for the first time that the decentralization theorem, which lies at the heart of fiscal federalism, is dependent on the structure of political institutions once the unrealistic assumption of no interjurisdictional spillovers is relaxed. Local elections and certain forms of party institutions must be in place before we can expect decentralization to deliver on its promises.

In particular, in this paper we incorporate into the analysis the role of both democratic (de)centralization and party (de)centralization. Under democratic decentralization local public goods are provided by democratically elected local governments and under democratic centralization by the national government. Party centralization, on the other hand, occurs when national party leaders have the power to nominate candidates for sub-national office. In contrast, nomination under party decentralization involves a political process with sequential elections in which voters select candidates through primaries.

The rest of the paper is structured as follows: Section 2 contains a summary and detailed explanation of the main findings of our theory. Section 3 includes a benchmark model in which local public goods are efficient and match the heterogeneous preferences over public spending from voters. Section 4 incorporates the analysis of party centralization and democratic (de)centralization. Section 5 contains the analysis of party decentralization and democratic (de)centralization. Section 6 concludes.

2. Summary of the Theory

[5] One rare exception is Enikolopov and Zhuravskaya (2007), who argue convincingly that fiscal decentralization produces better outcomes in the presence of party centralization. But they are not interested in political decentralization (i.e. elections) as we are here, and their paper is entirely empirical, with no formal component.

As noted above, we incorporate insights from political science and economics into a rigorous and formal extension of the influential "decentralization theorem," first developed by Oates in 1972. Of particular interest for us is the interaction between democratic (de)centralization and party (de)centralization in the provision of local public goods.

Our notion of what constitutes a centralized party is the same regardless of the electoral system used in a country. For us, a party is centralized when its national party leaders control access to the party name in local elections. For space reasons, we focus our formal analysis in this paper on majoritarian, single-member-district systems, defining decentralized parties as those that hold open or closed primaries (modeled separately) to choose candidates, as opposed to those having national party leaders nominate them. While we understand that many decentralized 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, both the conventional decentralization theorem (which assumes away spillovers) and the strong decentralization theorem (which assumes that public goods show spillovers) hold when parties are centralized. In these cases, the provision of local public goods by a system of elected sub-national governments is welfare superior to the centralized provision even under interjurisdictional spillovers of local public spending. More specifically, we find that these beneficial outcomes for public service delivery will hold when democratic decentralization (i.e. the creation of popularly elected sub-national governments) is combined with party centralization (i.e. the power of national party leaders to nominate candidates for sub-national office). Democratic decentralization ensures that

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

When parties are decentralized, by contrast, the participation rules of primaries create predictable incentives for the provision of public goods, and these in turn explain whether democratic decentralization dominates or is dominated by democratic centralization. For instance, if primaries are open, the strong theorem *is not* satisfied but the conventional decentralization theorem is. The strong decentralization theorem does not generally hold for democracies with party decentralization because primary elections do not create incentives for local politicians to internalize inter-jurisdictional spillovers. However, under open primaries, local public goods with no spillovers are Pareto efficient and a system of local governments maximizes the society's welfare gains from policy differentiation. By contrast, central governments also provide efficient local public goods, but this form of government does not maximize the gains associated with the heterogeneity of preferences.

On the other hand, when primaries are closed, neither the strong nor the conventional theorems are satisfied. For an economy with this type of electoral institution, parties in both structures of government (centralized and decentralized) will not have incentives to provide efficient local public goods. Moreover, local public goods will not maximize the gains from matching public spending with the heterogeneous preferences of voters. These outcomes are explained by the participation rules of closed primaries that give the right to vote only to a segment of co-partisan voters in the electorate. As a result, parties in both structures of government have incentives to target government spending to maximize the utility of a minority of voters. This explains why local public goods are not efficient and do not maximize the gains from matching policy with the heterogeneous preferences of voters.

In a related paper, we also develop a formal model examining the same relationships for democracies that use proportional representation (PR) systems (see Ponce-Rodriguez et al. 2015).[6] In that paper, we find that the benefits of decentralization are also contingent on the specific institutions of PR systems. To be more precise, in party centralized PR systems with closed party lists, the strong decentralization and the conventional decentralization theorems are satisfied. For party decentralized economies with proportional representation and open party lists, by contrast, the strong decentralization theorem is not satisfied while the conventional theorem is.

These findings have very significant implications for the scholarly understanding of decentralization among both economists and political scientists. Most significantly, they show that political institutions mediate the effects of the decentralized provision of public goods to an extent not previously realized. In addition to contributing to the theory of fiscal federalism, our findings help make sense of the mixed results that characterize the empirical scholarship on decentralization while adding clarity and detail to the theoretical literature. 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.

3 The Benchmark and Definitions

We begin by characterizing the set of local public goods that maximize the society's surplus from public goods. This approach to benchmarking our results allows us to compare

[6] Note also that we provide an empirical test of our theory in Ponce et al. (2016a)

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 districts i and $-i$ with $n^i = 1,2 \dots \dots N^i$ individuals in each district. Individuals do not have mobility across jurisdictions. The indirect utility of an individual with endowment e^i in district i is $v^i(e^i, G^i) = Max\, \mu^i(x^i, g^i, g^{-i})$ subject to $a)\ x^i = e^i - t^i$ and $b)\ g^i = N^i t^i\ \ \forall i$. The utility function is $\mu^i(x^i, g^i, g^{-i}) = \phi^i x^i + \psi^i ln(G^i)$ where x^i is a private good and $G^i = g^i + k^{-i} g^{-i}$ is the overall consumption of local public goods provided by districts i and $-i$, g^i, g^{-i}. The terms $\phi^i > 0$ and $\psi^i > 0$ are constants that measure the intensity of preferences for private and public goods.

The parameter $k^{-i} \in [0, 1)\ \forall -i$, measures the extent of inter-regional spillovers of g^{-i} over residents of district i. For local public goods without spillovers $k^{-i} = 0\ \forall -i$, and $k^{-i} = 1$ when local spending in district $-i$ is over a nationwide pure public good. Condition (a) is the individual's budget constraint where t^i is a head tax on residents of district i. The distribution of endowments across districts is given by $e^i \in [\underline{e^i}, \bar{e}^i]: h^i(e^i) > 0\ \ \forall i$ with $\sum_{\forall i,-i} \int_{\forall e^i} h^i(e^i) de^i = 1$. Condition (b), $g^i = N^i t^i\ \ \forall i$, is the constraint that public goods are fully financed by taxes.

The nationwide social welfare for this economy is given by:[7]

$$NSW = \sum_{\forall\, i,-i} \int_{\forall\, e^i} h^i(e^i)\, v^i(e^i, G^i) de^i \tag{1}$$

Let $\hat{\mathbf{g}}^* \in \mathbb{R}^2$: $\hat{\mathbf{g}}^* = [\hat{g}^{*i}, \hat{g}^{*-i}]$ such that $\hat{\mathbf{g}}^* \in\ argmax\ \ NSW$. We assume $\mathbf{H}(NSW)$ is a negative definite Hessian of NSW. For $\partial NSW/\partial g^i = 0\ \ \forall \hat{g}^{*i} > 0, \forall i,$ then $\hat{\mathbf{g}}^*$ is a global maximizer of NSW in the constrained policy set.

[7] Our notation in (1) means $\sum_{\forall\, i,-i} \int_{\forall\, e^i} h^i(e^i)\, v^i de^i = \int_{\forall\, e^i} h^i(e^i)\, v^i de^i + \int_{\forall\, e^{-i}} h^{-i}(e^{-i})\, v^{-i} de^{-i}$.

Proposition 1 characterizes a set of local public goods with and without spillovers that are Pareto efficient and maximize the gains attributed to matching the level of local public spending to the heterogeneous preferences of individuals across districts.

Proposition 1. *The Pareto efficient local public goods* $\hat{\boldsymbol{g}}^* = \left[\hat{g}^{*i}, \hat{g}^{*-i}\right]$ *that match the heterogeneous preferences of individuals across districts satisfy:*

$$\int_{\forall e^i} h^i(e^i)\frac{\partial \mu^i}{\partial G^i}de^i + k^i \int_{\forall e^{-i}} h^{-i}(e^{-i})\frac{\partial \mu^{-i}}{\partial G^{-i}}de^{-i} = \left\{\frac{1}{N}\right\} \int_{\forall e^i} h^i(e^i)\frac{\partial \mu^i}{\partial x^i}de^i \qquad (2)$$

Proof.

Find $\partial \delta_{NSW}/\partial g^i = 0 \ \forall \hat{g}^{*i} > 0, \forall i$ and re-arrange terms to obtain the result in (2).

In (2) local public goods with and without spillovers are provided at the point in which the marginal social gains in both districts from a marginal change in $\hat{g}^{*i} > 0 \ \forall i$ (equivalent to the change in utility of residents of district $i, \int_{\forall e^i} h^i(e^i)\frac{\partial \mu^i}{\partial G^i}de^i$, and due to inter-regional spillovers to the change in utility of residents of district $-i$, $k^i \int_{\forall e^{-i}} h^{-i}(e^{-i})\frac{\partial \mu^{-i}}{\partial G^{-i}}de^{-i}$) is equal to the social marginal costs of financing local spending through taxation $\left\{\frac{1}{N}\right\} \int_{\forall e^i} h^i(e^i)\frac{\partial \mu^i}{\partial x^i}de^i$. At $\hat{g}^{*i} > 0 \ \forall i$ the aggregate surplus from local public goods is maximized. The heterogeneity of preferences and endowments of individuals across districts implies that $\hat{g}^{*i} \neq \hat{g}^{*-i}$.

4 Party Centralization in a Democratically Centralized Government

In this section we introduce a model of electoral competition for a democracy with democratic centralization (voters elect an official to run the central government), and party centralization. We assume that when systems are democratically centralized they are also fiscally centralized. Similarly, democratically decentralized systems, i.e., those with elected local governments, are also fiscally decentralized.

In our economy, 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. In the first stage, candidates announce policies and party leaders nominate the candidate that will run in the general election with the party label. For simplicity, we do not model the decision of citizens to become candidates; instead we assume two candidates in each party look for the nomination from their parties. For an economy with party centralization, party leaders have full command over policy making by nominating only those candidates who adopt the ideal fiscal policy of party leaders. In the second stage of the political process, voters observe the parties' announced policies and elect a public official in the general election.

Two parties, labeled z and $-z$, compete in the election to form the government in a majoritarian electoral system with single member districts. The winning candidate takes all, forms the government, and implements policy. Under a central government, local public goods are provided by a single government that represents voters of all districts. The government finances its expenditures through a uniform tax on residents of all districts. We assume the central government provides uniform local public goods. However, this assumption is just a convenient shortcut of the analysis. As we show in Ponce et al (2016b), our main findings in this

paper are robust if we allow the central government to differentiate local public goods (see our discussion at the end of section 4.1).

During the second stage of the electoral process, individuals vote for the party that advances the policy that is closest to their own preferences on public spending. All individuals vote. Denote $\Psi_c^{zi} = v^{zi}\left(e^i, G^{zi}\right) - v^{-zi}\left(e^i, G^{-zi}\right)$ where Ψ_c^{zi} is the difference in the voter's payoff if party z is elected and implements policies g^{zi} and g^{z-i} in districts i and $-i$ instead of the alternative policies g^{-zi} and g^{-z-i} when party $-z$ is elected. The voter type e^i votes for party z if $\Psi_c^{zi} > 0$; if $\Psi_c^{zi} < 0$ he votes for party $-z$, and the voter flips a fair coin if $\Psi_c^{zi} = 0$.

From the point of view of parties, the individual's vote is uncertain (voting is probabilistic). The probability that a voter type e^i votes for party z in district i is $F_c^{zi}\left(\Psi_c^{zi}\right) = \int_{-\infty}^{\Psi_c^{zi}} f_c^{zi}\left(\Psi_c^{zi}\right) d\Psi_c^{zi}$, where $f_c^{zi}\left(\Psi_c^{zi}\right)$ is a continuous probability distribution over Ψ_c^{zi}. The expected vote of party z in district i is $\phi_c^{zi} = \int_{\forall\, e^i} h^i\left(e^i\right) F_c^{zi}\left(\Psi_c^{zi}\right) de^i$ and the expected vote in both districts is $\phi_c^z = \sum_{\forall i,-i} \phi_c^{zi}$. Define $\pi_c^z : \rho_c^z \to [0,1]$ as a continuous cumulative distribution over the plurality of the party, $\rho_c^z = \phi_c^z - \phi_c^{-z}$, where $\pi_c^z = \int_{-\infty}^{\rho_c^z} w_c^z(\rho_c^z)\, d\rho_c^z$ and $w_c^z(\rho_c^z)$ is the probability distribution over the party's plurality.

The equilibrium provision of local public goods for a democracy with a majoritarian electoral system, party centralization, and a nationwide election to form the central government, g_c^{*z}, is characterized in Lemma 1. Under our assumptions, Downsian parties converge in their fiscal platforms since they maximize a continuous and strictly concave probability of winning the election in the constrained policy set based on a common system of beliefs and strategy policy set.[8] Formally,

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

Proposition 2 *Parties* z *select* $g_c^{*z}\ \forall z, -z$ *for an economy with party centralization and a democratically centralized government such that*

$$g_c^{*z} \in argmax\ \pi_c^z(\rho_c^z)$$

$$subject\ to\ g_c^{zi} = g_c^{z,-i} = g_c^z\ \forall\ z \tag{3}$$

Define $\mathbf{g_c^z} = \left[g_c^{zi}, g_c^{z,-i}\right]$, ξ_c^z and $\delta_c^z(\mathbf{g_c^z}, \xi_c^z) = \pi_c^z + \xi_c^z\left\{g_c^{zi} - g_c^{z,-i}\right\}$. Moreover, we assume $\mathbf{H}(\delta_c^z)$ is a negative definite Hessian of δ_c^z. For the case $\mathbf{g_c^{*z}}$, ξ_c^{*z} it is satisfied that $\partial\delta_c/\partial g_c^{zi} = 0\ \forall g_c^{*zi} > 0\ \forall i$ and $\partial\delta_c/\partial\xi_c^z = 0\ \forall\ \xi_c^{*z} \neq 0$ then $\mathbf{g_c^{*z}}$ is a global maximizer of π_c^z in the constrained policy set.

Lemma 1 *Local public goods with and without spillovers are Pareto efficient for an economy with a majoritarian electoral system, single member districts, a democratically centralized government, and a centralized party system. All parties converge in providing a uniform local public good across districts,* $g_c^{*zi} = g_c^{*z,-i} = g_c^*\ \forall\ z$ *satisfying*

$$\int_{\forall\ e^i} h^i(e^i)\frac{\partial\mu^i}{\partial G^i}de^i + k^i \int_{\forall\ e^{-i}} h^{-i}(e^{-i})\frac{\partial\mu^{-i}}{\partial G^{-i}}de^{-i} = \left\{\frac{1}{N}\right\}\sum_{\forall i,-i}\int_{\forall\ e^i} h^i(e^i)\frac{\partial\mu^i}{\partial x^i}de^i \tag{4}$$

Proof

See the appendix

Lemma 1 says g_c^* is Pareto efficient for all $k^{-i} \in [0,1)\ \forall -i, i$ since a nationwide election provides voting rights to residents of all districts and parties have electoral incentives to

aggregate the individuals' marginal benefits (the left hand side of 4) and marginal costs (the right hand side of 4) from the provision of the uniform local public good in all districts.

4.1 Party Centralization in a Democratically Decentralized Government

In this section we consider the case of party centralization and democratic decentralization (i.e. local elections). Two parties compete in the local election of each district to form the local government. In a federation with a centralized party system, the leaders of nationwide parties face multiple electoral contests and nominate candidates who propose policies that maximize the party's *joint* probability of winning the elections in districts i and $-i$. As mentioned before, in a centralized party system, party leaders have full command on policy making by nominating only those candidates who adopt the ideal policy of party leaders. The winning party in each district takes all, forms the government, and designs policy. Local public goods in each district are chosen by the government of the district and expenditure is financed by a uniform tax on residents of the district.

Anticipating the results below, this section has two main findings. First, we show that party centralization in a system of local governments leads to Pareto efficient local public goods with and without inter-regional spillovers. Second, local public spending is differentiated to match the heterogeneous tax and spending policies demanded by voters across districts. Third, we show that democratic decentralization dominates democratic centralization.

For the analysis that follows we define the joint probability of party z of winning the elections in districts i and $-i$ by $\pi_{cL}^{z} = \pi_{cL}^{z}\left(\rho_{cL}^{zi}, \rho_{cL}^{z,-i}\right)$ where π_{cL}^{z} is a function of the pluralities of the party in both districts, ρ_{cL}^{zi} and $\rho_{cL}^{z,-i}$, where $\rho_{cL}^{zi} = \phi_{cL}^{zi} - \phi_{cL}^{-zi} \;\; \forall z, \forall i$, and $\phi_{cL}^{zi} = \int_{\forall e^i} h^i\left(e^i\right)$

$F_{cL}^{zi}(\Psi_{cL}^{zi})de^i$ is the proportion of votes that party z expects to receive in the local election of district i and $F_{cL}^{zi}(\Psi_{cL}^{zi})$ is the marginal probability that a voter type e^i votes for the party in the district's election (a similar interpretation is given to ϕ_{cL}^{-zi}).

Proposition 3 *In the local election of district i of an economy with party centralization and democratic decentralization, parties z and $-z$ select g_{cL}^{*zi} $\forall\, i$ such that*

$$g_{cL}^{*zi} \in argmax\ \pi_{cL}^{z}(\rho_{cL}^{zi}, \rho_{cL}^{z,-i}) \tag{5}$$

Define $\mathbf{g_{cL}^{z}} = [g_{cL}^{zi}, g_{cL}^{z,-i}]$ and assume $\mathbf{H}(\pi_{cL}^{z})$ is a negative definite Hessian of π_{cL}^{z}. For $\mathbf{g_{cL}^{*z}}$ satisfying $\partial\pi_{cL}^{z}/\partial g_{cL}^{*zi} = 0\ \forall g_{cL}^{*zi} > 0, \forall i$, $\mathbf{g_{cL}^{*z}}$ is a global maximizer of π_{cL}^{z} in the policy set.

On what follows, Lemma 2 characterizes the equilibrium spending policies for this economy and Theorem 1 shows that democratic decentralization dominates democratic centralization.

Lemma 2 *Party centralization in a system of local governments leads to a set of Pareto efficient local public goods with and without spillovers $\mathbf{g_{cL}^{*}} = [g_{cL}^{*i}, g_{cL}^{*-i}]$. At the political equilibrium, g_{cL}^{*i} $\forall i, \forall z$ satisfies the following:*

$$\int_{\forall e^i} h^i(e^i)\frac{\partial\mu^i}{\partial G^i}de^i + k^{-i}\int_{\forall e^{-i}} h^{-i}(e^{-i})\frac{\partial\mu^{-i}}{\partial G^{-i}}de^{-i} = \left\{\frac{1}{N}\right\}\int_{\forall e^i} h^i(e^i)\frac{\partial\mu^i}{\partial x^i}de^i \tag{6}$$

Proof

See the appendix

Lemma 2 says that in each district i parties choose a policy that is equivalent to a policy that maximizes an anonymous Utilitarian nationwide social welfare function subject to the constraint that the local public good of the district is financed by the residents of the district (see the equivalence between the results in expressions 2 and 6 implying $g_{cL}^{*i} = \hat{g}^{*i}\ \forall i$). As a result, the decentralized provision of local public goods with and without spillovers is Pareto efficient. Local public goods are Pareto efficient because the political process is centralized and rational parties recognize that the inter-regional externalities of local public goods create an interdependence between the parties' share of vote in the elections of districts i and $-i$. Thus, parties have electoral incentives to propose spending policies that internalize the inter-regional spillovers that maximize the party's joint probability of winning the elections in districts i and $-i$. In addition, local governments maximize the gains from matching the heterogeneous preferences of voters across districts since $g_{cL}^{*i}\ \forall i$ maximizes the utility of the average voter in each district.

Theorem 1 ***"Strong Decentralization Theorem".*** *The provision of local public goods with and without inter-regional spillovers by a system of local governments welfare-dominates their centralized provision when party systems are centralized.*

Proof

See the appendix

Theorem 1 says that the nationwide welfare of voters is higher when local public goods with and without inter-regional spillovers are provided by a system of local governments as opposed to the welfare associated with the provision from the central government. This is a

stronger version of the decentralization theorem in Oates (1972) in which democratic decentralization dominates democratic centralization only when local public goods do not display 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 centralization.

For reasons of space, we have assumed that the central government is constrained to provide uniform local public goods. However, in Ponce et al (2016b), we relax this assumption and show that our analysis is robust. Decentralization dominates centralization even when the central government can differentiate public goods. This is so because both forms of government provide Pareto efficient local public goods with and without spillovers, but a system of local governments is a superior mechanism to capitalize the welfare gains associated with matching public spending with the heterogeneous preferences of voters.

The intuition behind this result is that optimality conditions require the central government to equalize the marginal political benefits of allocating $1 through public spending and the corresponding marginal costs, of taking away $1 through taxes, across jurisdictions, while local governments do not operate under these constraints. As a result, the degree of policy differentiation by the central government is suboptimal and does not exhaust the gains from matching the preferences of voters across districts. Hence, our assumption here that the central

government provides uniform public goods is essentially a convenient shorthand to take into account that decentralized governments can be more efficient in matching local preferences and needs than the central government.

5 Decentralized Party Systems

We now turn our analysis to the provision of local public goods for a democracy with a majoritarian electoral system, single member districts, and decentralized party systems. For the case of party decentralization, a primary election takes place in which residents first vote to nominate a candidate, while later in the general election voters elect a public official. In this setting, our interest is to analyze how the political institution of primaries creates incentives for parties to represent the interests of a broad set of voters in the electorate versus the preferences of a minority coalition.

Proponents of decentralized 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 open, semi-closed, and closed (see Kaufman, Gimpel and Hoffman 2003). In open primaries voters of any affiliation may vote for any party while in closed primaries only those voters affiliated with a party (probably partisan voters) can vote in the party's primary.

Open 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 provision of local public goods of primary voters. This might be considered socially undesirable because in this case public spending does not maximize the society's net fiscal exchange associated with public goods but it maximizes the net fiscal surplus from public goods for a minority coalition of voters (the primary voters).[9]

The main results of this section are: first, the strong decentralization theorem does not hold for economies with decentralized party systems. We also find that the specific institutions of primaries might (not) lead to the expected benefits of democratic decentralization. In particular, the conventional decentralization theorem (which assumes no spillovers) holds for economies with decentralized party regimes and *open* primaries. However, if primaries are *closed* then the decentralization theorem, in general, does not hold. These are important and novel results. In summary, our theory shows that the political institutions of party decentralized systems may also matter considerably in determining the gains from decentralization.

5.1 Party Decentralization in a Democratically Centralized Government

In this section we analyze a model with party decentralization (there are sequential elections with a primary and a general election) and democratic centralization (voters elect public officials only for the central government). The economy is also fiscally centralized, that is, local public goods are provided by a single government that represents voters of all districts.

[9] The net 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 benefit). On the other hand, higher spending requires higher taxes and lower consumption of private goods (this is the marginal cost). When public spending is low the marginal benefit is likely to outweigh the marginal costs of increasing the provision of local public goods. This guarantees that the fiscal exchange associated with the provision of local public goods entails a non-negative surplus, see Martinez-Vazquez (1982).

The government finances its expenditures through a uniform tax on residents of all districts. As before, two parties compete in a nationwide election to form the central government in a majoritarian electoral system with single member districts. The winning candidate takes all, forms the government, and implements policy.

In the first stage of our model, two individuals, denoted by j and j' in each party seek the nomination of their party by declaring their binding policy platforms over public spending. All voters observe the candidates' policies but only qualified voters vote in the primary election.[10] In a closed primary the right to vote is limited only to voters affiliated with the party, and in open primaries all voters (those affiliated and not affiliated with some party) can vote to nominate a candidate.[11] The candidate who receives the majority of the votes across all districts wins the nomination of her party.

In the second 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.[12] In the first stage of the game, candidates $j, j' \in z$ announce spending platforms to maximize their *joint* probability, Π_d^{jz}, of winning the nationwide primary and general elections. Candidates propose a policy platform that is sequentially rational and therefore their policy platform must consider two different states that might be played in the second stage: candidates might compete in the general election against candidate 1 or 2 of party$-z$.

[10] The assumption that, in the first stage of the electoral contest, 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.

[11] This assumption implies that if the primary is open then all voters in the economy vote in the primaries of parties z and $-z$, while if the primary is closed then only qualified voters vote in the primary of party z or party $-z$.

[12] 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 in the primary election.

Define $\Pi_d^{jz} = \int_{-\infty}^{\tilde{\rho}_0^{jz}} \int_{-\infty}^{\rho_1^{jz}} \int_{-\infty}^{\rho_2^{jz}} w_d^{jz}\left(\tilde{\rho}_0^{jz}, \rho_1^{jz}, \rho_2^{jz}\right) d\tilde{\rho}_0^{jz}\, d\rho_1^{jz}\, d\rho_2^{jz}$ as the joint cumulative mass that candidate $j \in z$ wins both elections where $w_d^{jz} = \partial \Pi_d^{jz} / \partial \tilde{\rho}_0^{jz} \partial \rho_1^{jz} \partial \rho_2^{jz} > 0$, and $\tilde{\rho}_0^{jz}$ is the nationwide plurality in the primary election of candidate $j \in z$ versus that of candidate $j' \neq j : j' \in z$, and ρ_1^{jz} is the nationwide plurality in the general election for the state in which candidate $j \in z$ runs against candidate 1 of party $-z$. A similar interpretation is given toρ_2^{jz}.

For each of the pluralities in the primary $\tilde{\rho}_0^{jz} = \tilde{\phi}_0^{jz} - \tilde{\phi}_0^{j'z}$ and in the general election $\rho_l^{jz} = \phi_l^{jz} - \phi_l^{j-z}\ \forall\, l = \{1,2\},\ \tilde{\phi}_0^{jz} + \tilde{\phi}_0^{j'z} = 1$ and $\phi_l^{jz} + \phi_l^{j-z} = 1\ \forall l$ are both satisfied, where $\tilde{\phi}_0^{jz} = \sum_{\forall i} \tilde{\phi}_0^{jzi}$ is the nationwide expected proportion of the vote in the primary and $\phi_l^{jz} = \sum_{\forall i} \phi_l^{jzi}$ is the nationwide expected proportion of the vote of candidate $j \in z$ in the general election. The expected proportion of the vote of candidate j in the primary of district i is $\tilde{\phi}_0^{jzi} = \int_{\forall\, \tilde{e}^i} \tilde{h}^i\left(\tilde{e}^i\right) \tilde{F}_0^{jzi}\left(\tilde{\Psi}_0^{jzi}\right) d\tilde{e}^i\ \forall i,$ where $\tilde{F}_0^{jzi}\left(\tilde{\Psi}_0^{jzi}\right) = \partial^2 F^{jzi}\left(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi}\right) / \partial \Psi_1^{jzi}\, \partial \Psi_2^{jzi}$ is the continuous marginal probability that a voter type e^i in district *i* votes for $j \in z$ in the primary, and $F^{jzi} = F^{jzi}\left(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi}\right)$ is the joint probability that the voter type e^i votes for $j \in z$ in the primary and the general election. Similarly, the corresponding proportion of the vote for $j \in z$ in district i in the general election is $\phi_l^{jzi} = \int_{\forall\, e^i} h^i\left(e^i\right) F_l^{jzi}\left(\Psi_l^{jzi}\right) de^i$ for $l = \{1\ ,2\}$ where F_l^{jzi} is the marginal cumulative mass of $j \in z$ in district i if he faces candidates $l \in \{1\ \vee 2\} \in -z$ in the general election.

The joint probability that voter type e^i votes for $j \in z$ in the primary and the general election $F^{jzi}\left(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi}\right)$ is continuous and non-decreasing with $\tilde{\Psi}_0^{jzi}$, Ψ_1^{jzi}, and Ψ_2^{jzi}. Recall $\tilde{\Psi}_0^{jzi}$ is the individual's voting calculus and represent the net gain of voting for candidate *j*

of party z instead of candidate j' of party z in the primary, and Ψ_1^{jzi} and Ψ_2^{jzi} are the net gains of voting for candidate j of party z instead of candidates 1 or 2 of party $-z$ in the general election. A voter type e^i in district i votes for $j \in z$ in the nationwide primary and general election if simultaneously $\tilde{\Psi}_0^{jzi} = \tilde{v}^{jzi}\left(\tilde{e}^i, G^{jzi}\right) - \tilde{v}^{j'zi}\left(\tilde{e}^i, G^{j'zi}\right) > 0$, $\Psi_1^{jzi} = v^{jzi}\left(e^i, G^{jzi}\right) - v^{1-zi}\left(e^i, G^{1-zi}\right) > 0$, and $\Psi_2^{jzi} = v^{jzi}\left(e^i, G^{jzi}\right) - v^{2-zi}\left(e^i, G^{2-zi}\right) > 0$.

Proposition 4 *A candidate j of party z selects g_d^{*jz} for an economy with a party decentralized system and democratic centralization such that*

$$g_d^{*jz} \in argmax\ \Pi_d^{jz}\left(\tilde{\rho}_0^{jz},\ \rho_1^{jz}, \rho_2^{jz}\right)$$

$$subject\ to\ \ g_d^{jzi} = g_d^{jz,-i} = g_d^{jz}\ \ \forall j\ , \forall z \tag{7}$$

Define $\mathbf{g_d^{jz}} = \left[g_d^{jzi} g_d^{jz,-i}\right]$, ξ_d^{jz} and $\delta_d^{jz}\left(\mathbf{g_d^{jz}}, \xi_d^{jz}\right) = \Pi_d^{jz}\left(\tilde{\rho}_0^{jz},\ \rho_1^{jz}, \rho_2^{jz}\right) + \xi_d^{jz}\left\{\ g_d^{jzi} - g_d^{jz,-i}\right\}$. Moreover, assume $\mathbf{H}\left(\delta_d^{jz}\right)$ is a negative definite Hessian of δ_d^{jz}. For $\mathbf{g_d^{*jz}}, \xi_d^{*jz}$ satisfying $\partial\delta_d^{jz}/\partial g_d^{jzi} = 0\ \ \forall g_d^{*jzi} > 0$ and $\partial\delta_d^{jz}/\partial\xi_d^{jz} = 0\ \forall \xi_d^{*jz} \neq 0$, then $\mathbf{g_d^{*jz}}$ is a global maximizer of $\Pi_d^{jz}\left(\tilde{\rho}_0^{jz},\ \rho_1^{jz}, \rho_2^{jz}\right)$ in the constrained policy set.

Lemma 3 *For economies with a decentralized party system and democratic centralization, a uniform and Pareto efficient local public good $g^i = g^{-i} = g_d^{*z}\ \forall\ z$ is provided such that it satisfies the following:*

$$\sum\nolimits_{\forall\ i,-i} \alpha^{zi} \int\limits_{\forall\ e^i} h^i\left(e^i\right) v_g^{zi}\left(g_d^{*z}\right) de^i$$

$$= -\gamma^z \left(\sum_{\forall\, i,-i} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \tilde{v}_g^{zi}(g_d^{*z}) d\tilde{e}^i \right) - \sigma_\omega^z \tag{8}$$

Where $v_g^{zi} = \{\partial \mu^{zi}/\partial G^i\} - \left\{\frac{1}{N}\right\} \{\partial \mu^{zi}/\partial x^i\}\ \forall zi$. *Moreover,* $\alpha^{zi} \in (0,1)$:

$$\alpha^{zi} = \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \tag{9}$$

$$\gamma^z = \frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \tag{10}$$

where γ^z *is a weighted rate of substitution between marginal changes in the parties' plurality in the primary and the general election, and*

$$\sigma_\omega^z = \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right) \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \tag{11}$$

where σ_w^z *is a weighted covariance between the marginal probability of voting for party z in the nationwide general election,* $\partial F_l^{zi}/\partial \Psi_l^{zi}$*, and the change in the welfare of voters from an increase in the provision of the local public good* $\partial \Psi_l^{zi}/\partial g_d^z$.

Proof

See the appendix

Lemma 3 says that sequential elections (with a primary and a general election) induce candidates to propose a policy that reflects a compromise between the ideal policy of a nationwide weighted *average voter of the primary* and the ideal policy of a nationwide weighted *average voter of the general election* (see expression 8). The left hand side of (8) is a weighted average of marginal indirect utility gains (or the net fiscal exchange) for voters of all districts

participating in the general election. The parameter $\alpha^{zi} \in (0,1)\ \forall i$ reflects the relative political influence over candidates of residents of district i (the higher α^{zi} the closer is public spending to the ideal policy of the average voter of district i). Moreover, the first term of the right hand side of (8) is the corresponding weighted marginal indirect utility gains of primary voters.

The expression γ^z is a weighted rate of substitution between marginal changes in the parties' plurality in the primary and the general election. This term reflects the parties' incentives to weigh more (less) heavily the preferences of primary versus general election voters. The higher γ^z the higher is the marginal plurality gain of the party in the primary associated with a small increase in g_d^{*z} and the higher is the electoral incentives for party z to produce a policy closer to the ideal policy of the weighted average primary voter. If $\gamma^z = 0$, then parties propose the ideal policy of the weighted average voter of the general election.

Moreover, σ_w^z is a weighted covariance between the marginal probability of voting for party z from individual type e^i in the general election, $\partial F_l^{zi}/\partial \Psi_l^{zi}$, and the change in the well-being of the voter from an increase in the provision of the local public good $\partial \Psi_l^{zi}/\partial g_d^z$. Candidates will have electoral incentives to increase the size of g_d^{*z} when σ_w^z is positive, that is, when the electorate is constituted by voters who simultaneously have higher than average marginal probabilities of voting for the candidate and higher than average values of the net fiscal incidence of local public goods (this term is $\partial \Psi_l^{zi}/\partial g_d^z$).

Lemma 3 also says that the political institutions of the primary election matter to determine whether the centralized provision of local public goods is Pareto efficient or not. To be precise, the provision by the central government of local public goods with and without externalities is Pareto efficient if primary elections are open. However, if primary elections are closed local public goods with and without externalities are Pareto efficient only when γ^z is

sufficiently low. To see this, note that if primary elections are closed and γ^z is sufficiently high, then the central government provides the policy desired by a minority of the electorate (the primary voters) and therefore local public goods with and without externalities are *not* Pareto efficient.

5.2 Party Decentralization in a Democratically Decentralized Government

Let us now proceed to characterize the provision of local public goods for an economy with party decentralization and democratic decentralization (voters also elect public officials at the local level). The structure of the game is easily extended from our previous discussion: local primary and general elections take place in each district. Candidates j, j' in each party, and in each district, announce policy platforms that maximize the joint probability of the candidate $j \in z$ of winning the local primary and general elections $\pi_{dL}^{jzi}\left(\tilde{\rho}_0^{jzi}, \rho_1^{jzi}, \rho_2^{jzi}\right)$. After the local general election takes place, the winner takes all and implements his policy platform. Local public goods are chosen by the local government of the district and expenditure is financed by a uniform tax on residents of the district.

Define the joint probability of candidate j of party z of winning the local primary and general elections by $\pi_{dL}^{jzi} = \int_{-\infty}^{\tilde{\rho}_0^{jzi}} \int_{-\infty}^{\rho_1^{jzi}} \int_{-\infty}^{\rho_2^{jzi}} w_d^{jzi}\left(\tilde{\rho}_0^{jzi}, \rho_1^{jzi}, \rho_2^{jzi}\right) d\tilde{\rho}_0^{jzi}\, d\rho_1^{jzi}\, d\rho_2^{jzi}$ where $w_d^{jzi} = \partial\, \pi_{dL}^{jzi} / \partial\tilde{\rho}_0^{jzi} \partial\rho_1^{jzi} \partial\rho_2^{jzi} > 0$, and $\tilde{\rho}_0^{jzi}$ is the plurality of candidate $j \in z$ versus that of candidate $j' \neq j : j' \in z$ in the local primary, and ρ_1^{jzi} is the plurality for the state in which candidate $j \in z$ runs against candidate 1 of party $-z$ in the local election of district i. A similar interpretation is given to ρ_2^{jzi}.

For each of the pluralities in the local primary $\tilde{\rho}_0^{jzi} = \tilde{\phi}_0^{jzi} - \tilde{\phi}_0^{j'zi}$ and in the local general election $\rho_l^{jzi} = \phi_l^{jzi} - \phi_l^{j-zi}$ it is satisfied that $\tilde{\phi}_0^{jzi} + \tilde{\phi}_0^{j\prime zi} = 1$ and $\phi_l^{jzi} + \phi_l^{j-zi} = 1 \; \forall l = 1,2$. The expected proportion of the vote of candidate j in the primary of district i is $\tilde{\phi}_0^{jzi} = \int_{\forall \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\, \tilde{F}_0^{jzi}(\tilde{\Psi}_0^{jzi}) d\tilde{e}^i \; \forall i,$ where $\tilde{F}_0^{jzi}(\tilde{\Psi}_0^{jzi}) = \partial^2 F^{jzi}(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi}) / \partial\Psi_1^{jzi}\, \partial\Psi_2^{jzi}$ is the continuous marginal probability that a voter type e^i in district *i* votes for $j \in z$ in the primary, and $F^{jzi} = F^{jzi}(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi})$ is the joint probability that the voter type e^i votes for $j \in z$ in the primary and the general elections of district i. Similarly, the corresponding proportion of the vote for $j \in z$ in district i in the general election is $\phi_l^{jzi} = \int_{\forall e^i} h^i(e^i)\, F_l^{jzi}(\Psi_l^{jzi}) de^i$ for $l = \{1,2\}$, where F_l^{jzi} is the marginal cumulative mass of $j \in z$ in district i if he faces candidates $l \in \{1 \vee 2\} \in -z$ in the general election.

Proposition 5 *In the local election of district i of an economy with party decentralization and democratic decentralization, all candidates j, j' of party z select $g_{dL}^{*zi} \; \forall z$ such that*

$$g_{dL}^{*zi} \in argmax \; \pi_{dL}^{zi}(\tilde{\rho}_0^{zi}, \rho_1^{zi}, \rho_2^{zi}) \tag{12}$$

We assume $\mathbf{H}(\pi_{dL}^{zi})$ is a negative definite Hessian of π_{dL}^{zi}. If $g_{dL}^{*zi} \; \forall z$ satisfies $\pi_{dL}^{zi} / \partial g_{dL}^{*zi} = 0 \; \forall g_{dL}^{*zi} > 0, \forall z$, then g_{dL}^{*zi} is a global maximizer of π_{dL}^{zi} in the policy set.

On what follows, Lemma 4 provides a general characterization of local public goods. Theorems 2 shows that the strong decentralization theorem is not satisfied for economies with decentralized party systems, and Theorem 3 shows that the conventional decentralization theorem holds in the case of open primaries and fails to hold in the case of closed primaries.

Lemma 4 *For economies with a decentralized party system and democratic decentralization, local public goods* $g_{dL}^{*zi}\ \forall\, i, -i$ *are provided such that* g_{dL}^{*zi} *satisfies the following:*

$$\int_{\forall\, e^i} h^i(e^i)\, v_{g_{dL}}^{zi}\left(g_{dL}^{*zi}\right) de^i = -\chi^{zi}\left(\int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}}\tilde{v}_{g_{dL}}^{zi}\left(g_{dL}^{*zi}\right) d\tilde{e}^i\right) - \sigma_\omega^{zi} \tag{13}$$

Where

$$\chi^{zi} = \frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i)\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \tag{14}$$

Where χ^{zi} *is a weighted rate of substitution between marginal changes in the party's plurality in the district's primary and the general local election, and*

$$\sigma_\omega^{zi} = \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}}\, \sigma_l^{zi}\left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right) \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i)\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \tag{15}$$

Where σ_ω^{zi} *is a weighted covariance between the marginal probability that voter type* e^i *votes for party z in the local general election in district i,* $\partial F_l^{zi}/\partial \Psi_l^{zi}$*, and the change in the well being of voters from an increase in the provision of the local public good* $\partial \Psi_l^{zi}/\partial g_{dL}^{zi}$*.*

Proof

See appendix

The interpretation of this equilibrium is similar to that given in Lemma 3. In local elections with party decentralization, parties provide a public good that reflects a compromise between the ideal spending policy of the district's weighted average voter in the primary election and the district's weighted average voter in the general election (this is condition 13). The incentives of party *z* to weigh more (less) heavily the preferences over public spending of

primary versus general election voters in district i are given by χ^{zi}. The higher χ^{zi} the higher the electoral incentives for party z to produce a policy closer to the ideal policy of the weighted average voter of the primary in district i. If $\chi^{zi} = 0\ \forall i,$ parties propose the ideal policy of the weighted average voter of the general election of the district. Moreover, the expression, σ_w^{zi}, is a weighted covariance between the marginal probability of voting for party z in the general local election, $\partial F_l^{zi}/\partial \Psi_l^{zi}$, and the change in well being of voters from an increase in the local public good, $\partial \Psi_l^{zi}/\partial g_{dL}^{zi}$. For $\sigma_w^{zi} > 0$, the higher σ_w^{zi}, the higher the provision of the public good in the district.

In terms of the properties of efficiency and policy differentiation of this equilibrium, the implications of Lemma 4 are the following: first, local public goods with inter-regional spillovers are *not* Pareto efficient for either type of primary, open or closed. Second, if local public goods do not show spillovers, then the political institutions of the primary election determines whether the decentralized provision of local public goods is Pareto efficient or not. In open primaries, local public goods without inter-regional spillovers are Pareto efficient for all values of χ^{zi} because parties select the ideal policy of the average voter in the district. However, for closed primaries, local public goods are Pareto efficient only if $\chi^{zi} \longrightarrow 0$. If χ^{zi} is sufficiently high, then local governments provide the local public good in the district that maximizes the well-being of the primary voters of district i (a local minoritarian coalition in the district).

Third, the political institutions of primaries also produce different results on the ability of local governments to maximize the gains from policy differentiation. If primary elections are open then decentralization produces the ideal policy of the average voter in each district and the gains from inter-regional policy differentiation are maximized. If the primary is closed then the gains from policy differentiating can be maximized only if $\chi^{zi} = 0$. As we mentioned before, if

χ^{zi} is sufficiently high, local governments provide a local public good that maximizes the well-being of the average primary voter of district i which implies that the extent of inter-regional policy differentiation achieved by local governments is sub optimal.

Theorem 2 *In majoritarian democracies with a decentralized party system and open primaries, the strong decentralization theorem does not hold but the conventional decentralization theorem holds.*

Proof

See the appendix

On the one hand, Theorem 2 shows that candidates running for local governments in a party decentralized system with open primaries have incentives to propose the size of public spending that maximizes a unanimous Utilitarian social welfare function of local residents in *each* jurisdiction. In other words, parties propose the ideal policy of the average voter of the district and therefore the resulting policies are Pareto efficient but only for local public goods without spillovers. Therefore the strong decentralization theorem does not hold for economies with party decentralized systems and *open* primaries. In addition, the local provision of public goods matches the heterogeneous preferences of voters across districts.

On the other hand, if the government is democratically centralized and there is party decentralization with open primaries, local public goods with and without spillovers are uniform and Pareto efficient. In this case, local public goods reflect the ideal size on public spending of the average voter of all districts.[13] Lastly, Theorem 2 shows that if local public goods do not

[13] In open primaries the distribution of ideal policies of primary and general election voters is the same. For this reason, candidates in local elections have strong incentives to select the ideal policy of the average voter of the

display spillovers then decentralization dominates centralization because the heterogeneous provision of local public goods in a democratically decentralized system maximizes the welfare gains from policy differentiation, while the centralized system leads to suboptimal policy differentiation. Consequently, the conventional decentralization theorem holds for majoritarian democracies with decentralized party systems and open primaries.

Theorem 3 *The strong and the conventional decentralization theorems do not hold in majoritarian democracies with decentralized party systems and closed primaries.*

Proof

See the appendix.

In general, the strong and the conventional decentralization theorems do not hold for a democracy with party decentralization and closed primaries.[14] As we mentioned before, sequential elections induce candidates 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 (see expressions 8 and 13). In the case of closed primaries, it is likely that minority coalitions (such as primary voters) could have strong political influence over local governments that is translated into policies that are 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 surplus from the net fiscal exchange for the society.

district (which means that local public goods without spillovers are Pareto efficient) while candidates competing in a nationwide election choose the ideal policy of the average voter of all districts (which means that local public goods with and without spillovers are efficient).

[14] A comparison of (8) and (2) 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 13 and 2). As a result, the strong and the conventional decentralization theorems do not hold.

For instance, in condition (24) there are parametric values of χ^{zi} in which local elections produce the size of public spending in the vicinity of the ideal policy of primary voters in each district instead of the ideal policy of all residents in each district.[15] Simultaneously, we can find parametric values (i.e. when $\gamma^z \longrightarrow 0 \wedge \chi^{zi} > 0 \; \forall i, \forall z$ are sufficiently high) in which the centralized government averages the political influence of local coalitions to produce a more moderate provision of local public goods that could be welfare superior to the less moderate provision of public goods in the 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 centralized government could be welfare superior to the differentiated but extreme policy positions produced by a system of local governments.

6. Conclusion

In this paper, we develop a formal extension and refinement of the decentralization theorem of Oates (1972), which has provided the basis for much past research in fiscal federalism issues. In particular, we provide a political economy analysis of local public goods by incorporating the joint influence of democratic (de)centralization and party (de)centralization. Our theory provides new insights to the theory of fiscal federalism: first, we show that for democracies with centralized parties, a system of elected local governments welfare-dominates a

[15] In particular, significantly large values of $\chi^{zi} \; \forall i$ 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 party z will select the ideal policy of a minority of the electorate (i.e. the ideal policy of the weighted average voter of the primary in district *i*).

centralized government even if local public goods show inter-jurisdictional spillovers. We call this result the *strong decentralization theorem*. Indeed, an important implication of our theoretical model is that the combination of democratic decentralization and party centralization 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 centralization 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 decentralized parties, whether they have open or closed primaries. For countries with decentralized parties *and open primaries*, the conventional decentralization theorem (without spillovers) holds since the political institutions of open primaries promote the participation of voters and the representation of their preferences into policies. However, for countries with decentralized parties and closed primaries even the conventional decentralization theorem does not generally hold since 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.

To be more specific, 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 consequence is that the degree of policy differentiation in closed primary systems might be suboptimal and might not exhaust the gains from matching the preferences of voters across districts.

In other words, we demonstrate that creating locally elected governments can only be expected to improve public goods allocation either when parties are centralized or when there are no inter-jurisdictional spillovers. Local governments controlled by decentralized parties are not likely to provide efficient levels of public goods that spill over into adjacent constituencies efficiently. To summarize, then, our theory points to the importance of political institutions in determining how efficient decentralization outcomes can be in practice.

References

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.

Coughlin, P.J., 1992. *Probabilistic Voting Theory*. Cambridge University Press.

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.

Dickovick, J. Tyler. 2011. *Decentralization and Recentralization in the Developing World: Comparative Studies from Africa and Latin America.* University Park: The Pennsylvania State University Press.

Eaton, Kent. 2004. Risky Business: Decentralization from above in Chile and Uruguay. *Comparative Politics* 37 (1), 1-22.

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.

Haggard, Stephan and Robert R. Kaufman. 1995. *The Political Economy of Democratic Transitions.* Princeton: Princeton University Press.

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.

Martinez-Vazquez, Jorge. 1982. Fiscal Incidence at the Local Level. *Econometrica* 50 (5): 1207-1218.

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.

Ponce-Rodriguez, Raul, Charles R. Hankla, Jorge Martinez-Vazquez, and Eunice Heredia-Ortiz. 2016a. Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods. Working Paper available from authors.

Ponce-Rodriguez, Raul, Charles R. Hankla, Jorge Martinez-Vazquez, and Eunice Heredia-Ortiz. 2016b. Party Centralization and the Size and Distribution of Local Public Goods. Working Paper available from authors.

Ponce-Rodriguez, Raul, Charles R. Hankla, Jorge Martinez-Vazquez, and Eunice Heredia-Ortiz. 2015. The Provision of Local Public Goods in Proportional Electoral Systems with Closed and Open Party Lists. Working Paper available from authors.

Riker, William H. 1964. *Federalism: Origin, Operation, Significance.* Boston: Little Brown.

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.

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.