2012

Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods

Raul A. Ronce-Rodriguez
Universidad Autonoma de Ciudad Juarez

Charles R. Hankla
Georgia State University

Jorge Martinez-Vazquez
Georgia State University, jorgemartinez@gsu.edu

Eunice Heredia-Ortiz
Development Alternatives, Inc.

INTERNATIONAL CENTER FOR PUBLIC POLICY

International Center for Public Policy
Working Paper 12-27
August 2012

Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods

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 12-27

Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods

Raúl A. Ponce-Rodríguez
Charles R. Hankla
Jorge Martinez-Vazquez
Eunice Heredia-Ortiz

August 2012

International Center for Public Policy
Andrew Young School of Policy Studies
Georgia State University
Atlanta, Georgia 30303
United States of America

Phone: (404) 651-1144
Fax: (404) 651-4449
Email: hseraphin@gsu.edu
Internet: http://aysps.gsu.edu/isp/index.html

Georgia State University | ANDREW YOUNG SCHOOL OF POLICY STUDIES

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 hseraphin@gsu.edu.

Rethinking the Political Economy of Decentralization: How Elections and Parties Shape the Provision of Local Public Goods[1]

Raúl A. Ponce-Rodríguez,* Charles R. Hankla, Jorge Martinez-Vazquez,*** and Eunice Heredia-Ortiz****

*Department of Economics, Universidad Autónoma de Ciudad Juárez
**Department of Political Science, Georgia State University
*** Department of Economics, Georgia State University
****Development Alternatives Inc., DAI

Decentralization is among the most important global trends of the new century, yet there is still no consensus on how to design political institutions to realize its benefits. In this paper, we investigate the political conditions under which decentralization will improve the delivery of public goods. We begin by incorporating insights from political science and economics into a rigorous and formal extension of the "decentralization theorem". Our extension assumes inter-jurisdictional spillovers and suggests that the interaction of democratic decentralization (popularly elected sub-national governments) and party centralization (the power of national party leaders over subnational office-seekers) will produce the best outcomes for public service delivery. To test this argument empirically, we make use of a new dataset of sub-national political institutions created for this project. Our analyses, which allow us to examine educational outcomes in more than 125 countries across more than 25 years, provide support for our theoretical expectations.

[1]The order of authors has been drawn randomly. We would like to thank Craig Volden and Kyle Hanniman for their helpful feedback on earlier drafts of the paper and Gustavo Canavire-Baccarreza for discussions on the empirical approach. Our thanks also to Daniel Kuthy and Vanja Petricevic for their work in coding our dataset, and to 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 Kathleen Hale and 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.

As more and more of the world's states devolve power and resources to sub-national governments, decentralization is emerging as one of the most important global trends of the new century. Yet there is still no consensus concerning the benefits of decentralization and how to design institutions that can realize these benefits. In this paper, we investigate the political conditions under which this trend towards decentralization will improve the delivery of public goods. We begin by incorporating insights from political science and economics into a rigorous and formal extension of the "decentralization theorem". This theorem, which points to the efficiency benefits of sub-national provision of public goods, was first developed by Oates (1972) and has influenced virtually all of the literature over the past four decades.

In his theorem, Oates assumes, among other things, the absence of inter-jurisdictional spillovers in the centralized provision of local public goods. Our extension goes beyond Oates by producing a "strong decentralization theorem" in which the provision of local public goods by a system of sub-national governments can be welfare superior to centralized provision even under spillovers of local public spending. Further, it finds 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 creates the accountability necessary for efficient public goods provision, while party centralization increases local governments' incentives to provide public goods with spillover effects. We also find that democratic decentralization combined with party decentralization and open primaries leads to the hypothesized benefits of decentralization, but that democratic decentralization and party decentralization with closed primaries, in general, does not.

To test these arguments empirically, we make use of a new dataset of sub-national political institutions created for this project. Up to this point, scholars interested in sub-national political institutions have been forced to focus on single cases (especially the United States) or to assume that national level political institutions across countries are replicated at the sub-national level. Our new dataset allows us to examine how the structure of municipal politics influences educational outcomes (our proxy for public

goods provision) in more than 125 countries across more than 25 years. This empirical analysis, to our knowledge the broadest quantitative exploration of sub-national politics in the literature, provides support for our theoretical expectations.

In the scholarly world, our project has the potential to provide a new, interdisciplinary understanding of the implications of decentralization, overcoming some of the gap between work in political science and economics. In addition, we hope that our dataset will be of use to future studies that relate political institutions to outcomes at the sub-national level. Moreover, our project has the potential to influence thinking in the broader community of development practitioners. In the developing world, decentralization reforms have been strongly encouraged and assisted by the efforts of bilateral donor agencies and international financial institutions, many of which have made decentralization promotion a priority. These agencies, in turn, have been spurred on in their efforts by much of the scholarly literature, which emphasizes the beneficial effects of decentralization for the provision of public goods. Our project has the potential to provide a better understanding of which types of political institutions may be necessary for decentralization reforms to produce fuller results.

Review of the Literature

As noted above, decentralization is becoming 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 and bilateral 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 famous theorem, assumed a benevolent, welfare-maximizing government. While this assumption may have been 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 on which political institutions lead to the social welfare gains expected of decentralization by Oates (1972).[2]

We begin our exploration of the existing literature with an examination of what economists and political scientists (sometimes on parallel tracks) have said about decentralization and its efficiency implications. We then move to the broader literature on political institutions, party organization, and public goods provision, nearly all of which has developed with reference to national governments and apart from the study of decentralization. We conclude our review of the literature by highlighting the relative absence of work that considers how specific political institutions such as electoral rules and local representation might mediate the effects of decentralization on efficiency outcomes.

Modern research on decentralization began with Tiebout's landmark 1956 study, which argued that a decentralized system of public service delivery can maximize efficiency by allowing government services to vary according to the preferences of citizens in different jurisdictions. Oates picked up on this idea and qualified it in his 1972 formulation of the decentralization theorem, and scholars have since spent significant time improving or critiquing his argument and testing its empirical merit. For example, Breton (2002), Treisman (2007), Lockwood (2002), and Besley and Coate (2003) have examined whether central

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

governments could themselves target public goods delivery to regional preferences, while Bardhan (2002) and Manor (1999) are skeptical that individuals will move to regions that provide the policies they prefer. Others assert that decentralization increases opportunities for corruption (Treisman 2000, 2007; Tanzi 2002) and can be counter-productive if sub-national civil servants are not sufficiently professionalized (Shah 2003, Manor 1999).

Another group of economists (e.g. Weingast 1995) emphasizes the utility of decentralization because it generates a healthy competition among jurisdictions. Those that are most efficient at public good delivery will find new citizens, they argue, whereas those that govern poorly will find their populations and tax bases shriveled. Of course, like the arguments made by Tiebout and Oates, this argument has attracted its share of critics, especially among those who fear that decentralization could produce an inequitable distribution of goods (Prud'Homme 1995), exacerbate regional enmities (Treisman 1999, von Braun and Grote 2002), or lead to local elite capture (Bardhan and Mookherjee, 2000). Many scholars are also concerned about the possible negative impact of decentralization for a country's fiscal balances. For example, Triesman (2000) and Wibbels (2000) find an empirical connection between federalism and inflation.

Despite the critics, most scholarship in economics and political science, following in the tradition of Tiebout, Oates, and Weingast, has viewed decentralization positively. For example, many scholars believe that decentralization has a salutary effect on corruption by promoting transparency and accountability (Manor 1999, Gurgur and Shah 2002, Crook 2003). Others have highlighted ways in which some of the potential drawbacks of decentralization can be alleviated. For example, Rodden (2006) recommends that central governments follow a no-bailout policy without respect to sub-national authorities, thereby forcing them to internalize the consequences of their fiscal behavior.

On the empirical side, evidence for the proposed link between decentralization and efficiency has been mixed. Among the skeptics, Strumpf et al. (1999) find little evidence that decentralization in Uganda has improved health outcomes, Davoodi and Zou (1998) believe that devolving power to sub-national

governments slows economic growth in developing countries, Schwartz, Guilkey, and Racelis (2002) argue that decentralization reforms increased health expenditures in the Philippines but resulted in a shift away from the provision of public health resources, and Parry (1997) is skeptical that decentralization in Chile has improved educational outcomes. On the positive side, Lewis (1998) associates improved water delivery with decentralization in Kenya, Habibi et al. (2003) point to evidence that strong sub-national government reduced infant mortality in Argentina, and Yilmaz (1999), separating federal and unitary systems, sees a connection between decentralization and economic growth. Studies of Bolivia (Faguet and Sanchez 2008), Argentina (Habibi et al. 2003), and Indonesia (Simatupang 2009), as well as cross-national quantitative analyses (Heredia 2006), also point to improved educational outcomes with decentralization.

A reasonable summary, then, is that most scholars continue to see decentralization as a route to improving the delivery of public goods, but with a number of significant caveats (see Hankla 2009). If the benefits of decentralization are indeed conditional on other factors, something that many scholars are beginning to suspect, it could help account for the mixed empirical findings outlined above. Thus far, however, the literature has spent little time considering how political institutions might matter in mediating the effects of devolving power to sub-national governments. To lay the groundwork for incorporating these institutions into our argument, we turn now to a consideration of the broader literature on institutions and governance in political science.

Political scientists have long investigated the implications of different institutional configurations for the delivery of public goods, although their efforts have focused almost exclusively on national governments. Most scholars in this area agree, at least implicitly, that the political institutions likely to produce positive outcomes are those which expose leaders to popular democratic pressures while insulating them from particularistic interest groups. In making variations of this broad argument, researchers 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 (e.g., free trade, balanced budgets, energy conservation). To take some examples, Rogowski (1987) draws an

association between a proportional electoral system with large electoral districts and free trade policies. Other political scientists have found evidence that strong, democratic executives are more likely to provide public goods such as free trade (O'Halloran 1994, Nielson 2003), economic liberalization (Haggard and Kaufman 1995), and balanced budgets (Hallerberg and Marier 2004) than their less insulated counterparts. In addition, many scholars have associated balanced budgets with low levels of legislative and governmental party fragmentation (e.g. Roubini and Sachs 1989, Volkerink and de Haan 2001).

The impact of political party organization on policy outcomes, one of the central concerns of this article, has been much less thoroughly explored in the literature. The little research that has considered party organization has linked a more centralized structure (with empowered national elites) to public goods provision. Hankla (2006) and Nielson (2003), for example, argue that democracies with centralized political parties are more likely to adopt free trade policies, and Hallerberg and Marier (2004) find a connection between centralized parties and balanced budget in Latin America. Similarly, Hicken and Simmons (2008) argue that that education spending undertaken by decentralized parties is more particularistic and less effective. The link is simply that party centralization shifts power from local elites, who might be tempted to shore-up their support with particularistic goods, to national party leaders, who have electoral incentives to consider the aggregate national interest.

While nearly all of the research relating party structures with public goods delivery concerns the national level, some scholars have investigated the causal relationship between party and party system centralization on the one hand and the empowerment of sub-national governments 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.

Other scholars have addressed the question of whether and how party structures can contribute to (or undermine) the stability of federalism. Undoubtedly William Riker is the most prominent scholar to have taken up this question, arguing in his classic 1964 book that party centralization is among the most important drivers 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.

While all of these scholars have improved our understanding of how partisan and sub-national institutions interact, their focus has not been on connecting particular sub-national political institutions with public goods provision. Indeed, there are very few systematic studies in the literature that make this connection, but it is worth highlighting three influential analyses here. First, Erik Wibbels argues in his 2005 book that the presence of centralized parties facilitates the efforts of national leaders to push sub-national governments into market reforms. Second, R. Douglas Hecock (2006) finds a positive relationship between sub-national political competition and educational spending in Mexico. Third, and perhaps most related to our own work, Rubin Enikolopov and Ekaterina Zhuravskaya (2007) conclude, after a cross-national empirical study, that devolving fiscal authority to sub-national governments is more likely to improve public good (in this case, education) delivery when parties are centralized.

Despite some overlap with our interests here, however, there are a number of significant differences between our argument and those set forth by these scholars. Turning first to Wibbels (2005), his research focuses on party centralization as a means of national control within a decentralized political system, on not on the incentives such structures create for internalizing externalities. In a similar vein, Hecock (2006) is more interested in the level of partisan competition than in the questions of party organization that we study here.

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

Moreover, in contrast to Enikolopov and Zhuravshaya, we consider here the interaction between party centralization and democratic centralization rather than that between party centralization and fiscal centralization, so our theory is significantly different. In addition, we develop our argument formally using the decentralization theorem as a base, an approach which gives us more opportunity to leave our mark on the fundamental theory of decentralization. Finally, we believe that our data, which measures party decentralization more directly (Enikolopov and Zhuravskaya 2007 use party fragmentation in the national legislature) and at the sub-national level, more closely hews to theory. We will also make use of additional control variables for political institutions at the sub-national level and of educational indicators for the dependent variables that pick-up more directly on educational outcomes.

To summarize, then, the purpose of our research is to merge insights from political science with the decentralization literature outlined above, all to identify the political conditions needed for realizing the benefits of decentralization. We turn to developing our theory in the next section.

Development of the Theory

In developing our theory, we begin with the decentralization theorem and expand it formally to address the implications of different institutional configurations for its efficiency predictions. More specifically, we consider four distinct cases:

(1) countries that are **democratically decentralized** (i.e. they have democratically elected sub-national governments) and **party decentralized** (i.e. national leaders lack the power to select candidates for these sub-national elections);

(2) countries that are **democratically centralized** (i.e. they have no elected sub-national governments) but **party decentralized** (i.e. national leaders lack the power to nominate candidates for constituency elections to the national legislature);

(3) countries that are **democratically decentralized** (i.e. they have elected sub-national governments) but **party centralized** (i.e. national party leaders select candidates for sub-national elections); and

(4) countries that are **democratically centralized** (i.e. they have no elected sub-national governments) and **party centralized** (i.e. national party leaders nominate candidates to constituency elections for the national legislature).

In our formal model, we consider decentralized parties to be those that hold open or closed primaries (modeled separately) to choose candidates, as opposed to having national party leaders nominate them. While we understand that many decentralized parties in the world have free candidate access to the ballot (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 our model, we extend the decentralization theorem by incorporating inter-jurisdictional spillovers and find that the benefits of decentralization continue to hold (our "strong decentralization theorem"). Moreover, the results of our model indicate that the combination of democratic decentralization and party centralization yields the most efficient provision of public goods.[4] The intuition is that democratic decentralization produces incentives on the part of politicians to provide citizens with the bundle of public goods that they desire. In other words, as Tiebout (1956) and Oates (1972) point out, polities with elected sub-national governments can better target public services to the desires of their constituents, allowing bundles of goods to vary across constituency. Of course, having these governments be democratically elected is the key to ensuring that they are responsive to citizen desires (see Bird and Vaillancourt 1998, Manor 1999).

Party centralization, on the other hand, has the benefit of increasing the chances that the externalities of local public goods will be internalized. A common concern about democratic and fiscal decentralization is that local governments will under-provide public goods whose benefits spillover beyond their constituencies (e.g., Bird et al. 2003). This is because these governments are unable to internalize and

[4] Note, however, that our results indicate that democratic decentralization, when combined with open primaries, also yields positive outcomes. However, closed primaries are more problematic. As we are unable to differentiate cross-nationally between open and closed primaries in our empirical estimation, we plan to compare centralized and decentralized parties in the aggregate. It should also be noted that all of our predictions are based on specific assumptions that we will highlight in our models below.

profit from the political benefits of providing these goods optimally. The rational policy is instead to ignore the benefits that arise in other jurisdictions and/or to free-ride on the expenditures of neighboring districts; in either case the production of public goods will not be optimal. A number of basic public services, such as primary health, general education, water treatment and environmental protection, are likely to generate spillover effects and may not receive sufficient financing from local governments.

When parties are centralized, however, sub-national elected leaders do have some incentives to provide goods that may spillover into neighboring constituencies. National party leaders will be interested in generating optimal levels of public goods with spillover effects because they are concerned with their party's prospects in the country as a whole. In centralized parties, these national leaders have significant powers, not least of which is nomination power, over sub-national politicians and can push them to supply these goods.

To summarize, then, sub-national leaders in these systems have two masters whose interests are sometimes in competition, namely party chiefs in the national capital and local voters in their constituencies. Without the former, these leaders cannot be nominated and without the latter they cannot be elected. These competing loyalties produce incentives both to provide differentiated local public goods and to spend more money on goods with spillover effects. As a result, our argument is that systems that mix democratic decentralization with party centralization will have the best outcomes, other things equal. Systems that are centralized in both ways lack sufficient incentives to differentiate and target goods to local preferences, and systems that are decentralized in both have little incentive to generate optimal levels of public goods with geographical externalities.

A final question is whether countries that mix centralization and decentralization in the reverse way, those with no locally elected governments but with decentralized parties, might be able to produce the same beneficial tension. We think not. Even if politicians elected to the national legislature from local districts have incentives to concern themselves with local preferences, their ability to force the central government to differentiate tax and spending bundles for their constituents will be limited. Moreover, theoretical

models developed by Lockwood (2002) and Besley and Coate (2003) indicate that, while central governments may provide different constituencies with different bundles of public goods, a more decentralized approach to decision-making is likely to produce more efficient differentiation.

The Formal Model

In this section, we present our full formal expansion of the decentralization theorem.

The Benchmark

We begin by characterizing the set of local public goods that maximize the aggregate public good surplus which is our criteria for the comparative analysis of the welfare properties of the fiscally centralized and decentralized provision of local public goods. This benchmark allows us to compare our results with the normative analysis of Oates (1972) and recent political economy analysis (see Besley and Coate 2003 and Lookwood 2008).

Consider an economy constituted by districts i and $-i$ with $n^i = 1,2 \dots \dots N$ individuals in each district. Households do not have mobility across jurisdictions. The preferences of a household with an endowment e^i in district i is $v^i(e^i, G^i) = Max\ \mu^i(x^i, g^i, g^{-i}) = x^i G^i$ subject to $a)\ x^i = e^i - t^i$ and $b)\ g^i = Nt^i\ \ \forall i$, where $v^i(e^i, G^i)$ is the individual´s indirect utility, $\mu^i(x^i, g^i, g^{-i})$ are his preferences over a private good x^i, $G^i = (g^i + k^{-i}g^{-i})$ is the overall consumption of local public goods provided by district i, g^i, and by district $-i$, g^{-i}, and where t^i is a head tax on residents of district i

The parameter $k^{-i} \in [0, 1) \forall -i, i$, measures the extent of inter-regional spillovers of g^{-i} over residents of district i. For local public goods without spillovers $k^{-i} = 0\ \forall -i, i$, 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. The distribution of heterogeneous 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), $g^i = Nt^i \ \ \forall i$, is the constraint that public goods are financed by taxes.[5]

The nationwide welfare for this economy is given by:

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

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.

Proposition 1 characterizes a set of local public goods with and without spillovers that are Pareto efficient and exhaust the gains attributed to matching the size of local public spending according to the heterogeneous preferences of households over public goods.

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

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

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

[5] The government´s budget constraints say that g^i is financed by a head tax applied only to residents of the district. This configuration allows us to eliminate any possible gains of economies of scale in the provision of local public goods by the central government over sub-national governments. We impose this condition to evaluate whether the Decentralization Theorem of Oates (1972) holds in modern democracies once we introduce political institutions and incentives instead of governments controlled by benevolent social planners.

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 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, $\left\{\frac{1}{N}\right\} \int_{\forall e^i} h^i(e^i) \frac{\partial \mu^i}{\partial x^i} de^i$, of financing local spending through taxation. At $\hat{g}^{*i} > 0 \; \forall i$ the aggregate surplus from local public goods is maximized.

Party Centralization in a Single Unit of Government

In this section we develop a two stage model of electoral competition for a democracy with party centralization. The parties' problem is to aggregate the heterogeneous and conflicting preferences of voters for 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. We do not model the decision of citizens to become candidates, we simply assume two candidates in each party look for the nomination of their parties. For an economy with party centralization, party leaders have full command on 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' policies and elect a public official in the general election. All individuals vote.

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 designs policy. Under a central government, local public goods are provided by a single government that represents voters of all districts. The government finances its budget expenditures through a uniform tax on residents of all districts. We follow the literature by assuming that local public goods provided by the central government are uniform across districts (see Oates 1972, 1995).

In the second stage of the electoral process, individuals vote for the party that advances the spending policy that is closest to their own views on public spending. Denote $\Psi_c^{zi} = v^{zi}(e^i, G^{zi}) - v^{-zi}(e^i, G^{-zi})$ where Ψ^{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 choice of the vote is uncertain. The probability that a voter type e^i votes for party z in district i is $F_c^{zi}(\Psi^{zi}) = \int_{-\infty}^{\Psi_c^{zi}} f_c^{zi}(\Psi_c^{zi})\, d\Psi_c^{zi}$, where $f_c^{zi}(\Psi_c^{zi})$ is a continuous probability distribution over Ψ_c^{zi}. The expected vote of party z in district i is $\phi_c^{zi} = \int_{\forall e^i} h^i(e^i)\, F_c^{zi}(\Psi_c^{zi}) de^i$ and the expected vote in both districts is $\phi_c^z = \sum_{\forall i,-i} \phi_c^{zi}$. Define $\pi_c^z : \rho_c^z = \phi_c^z - \phi_c^{-z} \rightarrow [0,1]$ as a continuous cumulative distribution over the plurality of the party, ρ_c^z, where $\pi_c^z = \int_{-\infty}^{\rho^z} w_c^z(\rho_c^z)\, d\rho_c^z$ 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 proposition 2. Under our assumptions, Downsian parties converge in their fiscal platforms since they maximize a continuous an strictly concave probability of winning the election in the constrained policy set based on a common system of beliefs and strategy policy set (for a formal proof of convergence in probabilistic models with homogeneous parties see Coughlin 1992). Formally,

Proposition 2 Parties z select $g_c^{*z}\ \forall z, -z$ for an economy with party centralization and a single unit of 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 $\mathbf{g_c^z} = [g_c^{zi}, g_c^{z,-i}]$, ξ_c^z and $\delta_c^z(\mathbf{g_c^z}, \xi_c^z) = \pi_c^z + \xi_c^z\{g_c^{zi} - g_c^{z,-i}\}$. 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} satisfies $\partial\delta_c/\partial g_c^{zi} = 0\ \forall g_c^{*zi} > 0$ 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 are Pareto efficient for an economy with a majoritarian electoral system, single member districts, a single unit of government, a centralized party system, and for* $k^{-i} \in [0,1) \forall -i, i$. *All parties converge in providing a uniform local public good across districts,* $g_c^{*zi} = g_c^{*z,-i} = g_c^* \ \forall\, z$ *satisfying*

$$\sum_{\forall\, i,-i} \left(1 + k^{-i}\right) \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial G^i} de^i \ = \left\{\frac{1}{N}\right\} \sum_{\forall i,-i} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial x^i} de^i = 0 \qquad (4)$$

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 households' benefits (the left hand side of 4) and the parties' electoral costs (the right hand side of 4) from the provision of the uniform local public good in all districts.

Local Governments and Party Centralization

In this section we consider an economy with party centralization in which local public goods are provided by a system of local governments. 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 $-i$.[6] As mentioned above, in a centralized party system, party leaders have full command on policy making by nominating only those candidates who adopt the ideal fiscal policy of party leaders. The winning party in each district takes all, forms the government,

[6] In this paper we only consider the, empirically relevant, case in which parties have presence in all of the jurisdictions and we ignore the case in which some parties have only a regional presence.

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.

The findings of this section are the following: First, in lemma 2 we show that party centralization in a system of local governments leads to Pareto efficient local public goods with and without inter-regional spillovers. Moreover, in a system of local governments public spending is differentiated to match the heterogeneous preferences of voters across districts.

We also show that the provision of local public goods by a system of local governments is welfare superior to the unitary provision of local public goods with and without spillovers if parties are centralized, party leaders seek to maximize votes, the preferences over public goods are heterogeneous, and the unitary provision does not show economies of scale.

To prove our claims, we define the joint probability of party z of winning the elections in districts i and $-i$ by $\pi_{cL}^{z} = \pi_{cL}^{z}\left(\rho_{cL}^{zi}, \rho_{cL}^{z,-i}\right)$ as 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 $\phi_{cL}^{zi} = \int_{\forall\, e^i} h^i\left(e^i\right) F_{cL}^{zi}\left(\Psi_{cL}^{zi}\right) de^i$ is the proportion of votes that party z expects to receive in the local election of district i, and $F_{cL}^{zi}\left(\Psi_{cL}^{zi}\right)$ is the marginal probability that a voter type e^i votes for the party in the district´s election.

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

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

Define $\mathbf{g_{cL}^{z}} = \left[g_{cL}^{zi}, g_{cL}^{z,-i}\right]$, 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 constrained policy set.

On what follows, Lemma 2 characterizes the equilibrium spending policies for this economy and Theorem 1 shows the main result of this section.

Lemma 2 *Party centralization in a system of local governments leads to a set of Pareto efficient local public goods* $\mathbf{g^*_{cL}} = \left[g^{*i}_{cL}, g^{*-i}_{cL}\right]$ for all $k^{-i} \in [0,1)\ \forall -i, i, \forall z$. *At the political equilibrium,* $g^{*i}_{cL}\ \forall i, \forall z$ *satisfies the following:*

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

Proof

See the appendix.

Lemma 2 says that in each jurisdiction, parties in district i 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 2 and 6 implying $g^{*i}_{cL} = \hat{g}^{*i}\ \forall i$) . As a result, the spending policies in a system of local governments are Pareto efficient for local public goods with and without spillovers. Even though a local election in district i does not provide voting rights to voters of other jurisdictions, 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 local elections of districts i and $-i$. Thus, parties have electoral incentives to propose spending policies that internalize the inter-regional spillovers to maximize the party's joint probability of winning local elections in districts i and $-i$.

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 the centralized provision.*

Proof

See the appendix.

Theorem 1 says that the nationwide welfare of voters is higher under the provision of local public goods with and without inter-regional spillovers by a system of local governments relative the nationwide wellbeing derived by the provision of local public goods provided by the central government. This is a stronger case for the decentralized provision of local public goods than the Decentralization Theorem in which a system of local governments is welfare superior to a unitary provision only for the case in which local public goods do not display inter-regional spillovers (see Oates 1972).

Note, first that, local public goods with and without inter-regional spillovers are Pareto efficient under the provision of both a central government and a system of local governments. Second, local elections induce parties to select the size of public spending that maximizes a unanimous utilitarian nationwide social welfare without the constraint that local public goods must be uniform as it is the case of the unitary provision in our economy. By matching the heterogeneous preferences of households for public spending across districts, Pareto efficient local public goods with and without regional spillovers in a federation exhausts the nationwide welfare benefits to be gained from policy differentiation. Since the central government does not differentiate local public goods to local preferences then a system of local governments is welfare superior to the unitary option of public spending.

Fiscal Federalism and Party Decentralization

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

public goods. In the latter case, the government seeks to target the benefits of the collective action in this economy.

Proponents of decentralized party systems argue that primary elections promote the political participation of voters and the representation of their interests in the policies implemented by the government. However, the participation rules of primaries might actually limit both the voters' participation in elections and their political influence on policy design.

Primary elections can be open, semi-closed, and closed (see Kaufman and Gimpel 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. 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 fiscal exchange associated with the provision of local public goods for all residents of this economy but it maximizes the fiscal surplus from public goods for a minority coalition of voters in the electorate (the primary voters).[7]

The main results of this section are: first, we show that local public goods without inter-regional spillovers are Pareto efficient in decentralized political regimes for all types of primary elections and all structures of government. In our economy, the participation rules of primaries do not affect the efficiency but affect the distribution of gains across voters from public spending by modifying the political influence of primary versus general election voters in determining the level of public spending.

[7] The fiscal exchange of local public goods reflects the following tradeoff: on the one hand, an increase of public spending leads to higher utility of 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). At low taxes, public spending is also low which implies that the marginal benefit is likely to outweigh the marginal costs of increasing the provision of local public goods. This guarantees that at the equilibrium, the fiscal exchange associated with the provision of local public goods entails a non-negative surplus. See, for example, Martinez-Vazquez (1982).

Second, we also show that the Decentralization Theorem holds for economies with majoritarian electoral systems, single member districts, and decentralized party regimes with open primaries. However, if the primaries are closed the decentralization Theorem, in general, does not hold.

Primary and General Elections in a Single Government and a System of Local Governments

Consider first the case in which parties compete for a single national government in a sequential electoral contest. 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, and all voters observe the candidates' policies but only qualified voters vote in the primary election.[8] 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. [9] The candidate who receives the majority of the votes across all districts wins the nomination of his 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.[10] The winner of the general election takes all, forms the government, and implements his policy platform.

In the first stage of the game, candidates $j, j' \in z$ announce their fiscal platforms on public spending to maximize the 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: the candidates might compete in the general election against candidate 1 or 2 of party $-z$.

[8] 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 election to please, respectively, primary and general election voters.

[9] 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 $-z$.

[10] The assumption that voters vote sincerely also 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 $j \in z$ wins both elections where $w_d^{jz} = \partial\Pi_d^{jz}/\partial\tilde{\rho}_0^{jz}\partial\rho_1^{jz}\partial\rho_2^{jz} > 0$, $\tilde{\rho}_0^{jz}$ is the nationwide plurality of candidate $j \in z$ versus that of candidate $j' \neq j: j' \in z$ in the primary election, and ρ_1^{jz} is the nationwide plurality for the state in which candidate $j \in z$ runs against candidate $1 \in -z$ in the general election. A similar interpretation is given to ρ_2^{jz}.[11]

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

The joint probability $F^{jzi}(\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}, \Psi_2^{jzi})$ is continuous and non-decreasing with $\tilde{\Psi}_0^{jzi}, \Psi_1^{jzi}$, and Ψ_2^{jzi}. Sincere voting implies that a voter type e^i in district i votes for $j \in zi$ in the nationwide primary and general election if $\tilde{\Psi}_0^{jzi} = \tilde{v}^{jzi}(\tilde{e}^i, G^{jzi}) - \tilde{v}^{j'zi}(\tilde{e}^i, G^{j'zi}) > 0$,

$\Psi_1^{jzi} = v^{jzi}(e^i, G^{jzi}) - v^{1,-zi}(e^i, G^{1,-zi}) > 0$, $\Psi_2^{jzi} = v^{jzi}(e^i, G^{jzi}) - v^{2,-zi}(e^i, G^{2,-zi}) > 0$.

[11] On what follows, and for convenience of our analysis, we will distinguish voters who participate in the primary election (from voters who participate in the general election) with a tilde.

Proposition 4 A candidate j of party z selects g_d^{*jz} for an economy with a single unit of government and party decentralization 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.

Proposition 5 *The electoral equilibrium for our economy with party decentralization for all structures of government is: all candidates of the same party select the same fiscal platform. However, if the primary election is open the fiscal policies of parties z and –z converge and if the primary is closed the parties´ policies diverge.*

Lemma 3 *For economies with a decentralized party system and a central government representing voters of all districts, 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:*

$$\sum\nolimits_{\forall\ i,-i} \alpha^{zi} \int\limits_{\forall\ e^i} h^i\left(e^i\right) v_g^{zi}(g_d^{*z}) de^i$$
$$= -\gamma^z \left(\sum\nolimits_{\forall\ i,-i} \int\limits_{\forall\ \tilde{e}^i} \tilde{h}^i\left(\tilde{e}^i\right) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \tilde{v}_g^{zi}(g_d^{*z}) d\tilde{e}^i\right) - \sigma_\omega^z \tag{8}$$

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

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

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

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

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

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 wellbeing 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.

In this economy, the parties' policies converge when the primary is open and diverge in sequential elections with closed primaries. The intuition of this result is straightforward: in sequential elections with open primaries parties share a common system of beliefs over the citizens' voting behavior, there are no restrictions on the voting rights for citizens (voters can vote in all primaries and the general election), the policy strategy set is the same for both parties, and candidates are not otherwise differentiated. Moreover, parties seek to maximize a function that is continuous and strictly concave in the constrained policy space. In this case, parties converge in their fiscal policies by selecting the ideal policy of the average voter of the general election.

However, if primaries are closed then parties propose a uniform 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 expression 8). In closed primaries, the voting rights are only granted to affiliated voters and therefore the distribution of the ideal policies of primary voters is different for each party. Since the distribution of ideal policies of primary and general election voters that parties aggregate are different, the parties' spending policies diverge.

Lemma 3 characterizes the first order condition for the problem of public spending design for candidates in decentralized party systems. In this Lemma, the left hand side of (8) is a weighted average of marginal indirect utility gains due to a marginal change in the public good for all voters participating in the general election while the first term of the right hand side of (8) is the corresponding weighted marginal indirect utility gain 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 over public spending 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 the electoral incentives for party z to produce a policy closer to the ideal policy of the weighted average primary voter.

Moreover, σ_w^z is a weighted covariance between the marginal probability of voting for party z for the individual type e^i, $\partial F_l^{zi}/\partial \Psi_l^{zi}$, 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 the uniform local public good when voters with higher than average marginal probabilities of voting for the candidate have also higher than average values of $\partial \Psi_l^{zi}/\partial g_d^z$.

Party Decentralization and Local Elections

Now we proceed to characterize the provision of local public goods in a system of local governments with party decentralization. A detailed description of the electoral game for an economy with

party decentralization in a system of local governments is omitted to save space. However, the structure of the game is easily extended from our previous discussion: a local primary and general election take place. Candidates j, j' in each party, and in each district, announce policy platforms that maximize the joint probability of the candidate, say $j \in z$, of winning the primary and general elections in their district, $\pi_{dL}^{zi}\left(\tilde{\rho}_0^{zi},\ \rho_1^{zi}, \rho_2^{zi}\right)$. Only qualified residents of the district vote in the local primary while all the residents of the district can vote in the local general election. Voting is sincere in both elections. The local general election takes place, the winner takes all, and implements his policy platform.

Proposition 6 In the local election of district i of an economy with party decentralization, all candidates j, j' of party z select $g_{dL}^{*zi}\ \forall z$ such that

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

We assume $\mathbf{H}\left(\pi_{dL}^{zi}\right)$ 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 constrained policy set.

On what follows, Lemma 4 provides a general characterization of local public goods for federations with majoritarian systems, single member districts, and party decentralized regimes. Also on what follows, Theorems 2 and 3 show that the decentralization theorem holds in open primaries and fails to hold in closed primaries.

Lemma 4 *For economies with party decentralization and a system of local governments representing voters in each district, 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

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

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

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

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 the appendix.

The interpretation to this equilibrium is similar to that given in Lemma 3. In local elections, parties provide a public good that reflects a compromise between the ideal spending policies of the district's weighted average voter of the primary and the district´s weighted average voter of the general election. The parties´ incentives to weigh more (less) heavily the preferences over public spending of primary versus general election voters in district i are given by χ^{zi} and σ_w^{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 primary voter in district i. Moreover, the expression, σ_w^{zi}, is a weighted covariance between the marginal probability of voting for party z in the general election of the district, $\partial F_l^{zi}/\partial \Psi_l^{zi}$, and the change in well being of voters from an increase in the provision of 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.

Theorem 2 *The Decentralization Theorem holds in Majoritarian democracies with a decentralized party system, open primaries, and single member districts.*

Proof

See the appendix

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. The resulting policies are Pareto efficient for local public goods without spillovers. Moreover, public goods are matched according to the heterogeneous preferences of voters across districts. In the case candidates compete to form a nationwide government, local public goods with and without spillover are uniform and Pareto efficient, and the size of the local public good reflects the ideal size on public spending from the nationwide average voter. Lastly, if local public goods do not display regional spillovers then a system of local governments supplying local public goods $g_{dL}^{*i} > 0$ in districts i and $-i$ is welfare superior to a uniform provision $g_d^* > 0$ by a central government. Consequently, the Decentralization Theorem holds for majoritarian democracies with decentralized party regimes and open primaries.

Theorem 3 *The Decentralization Theorem does not hold in Majoritarian democracies with single member districts and decentralized party systems with closed primaries.*

Proof

See the appendix

In general, the Decentralization Theorem does not hold for a democracy with party decentralization and closed primaries.[12] To see this, consider the case in which minority coalitions (i.e. primary voters) have strong political influence over local governments that is translated into local public goods that are closer to their preferences. In this case, the electoral competition might produce extreme policy positions (too much or too little local public spending) instead of moderate policies.

For instance, in condition (13) there are parametric values of χ^{zi} and σ_w^{zi} in which local elections produce the size of public spending in the vicinity of the ideal policy of primary voters in each district (a minority of the society) instead of the ideal policy of all residents in each district (the utilitarian measure of social welfare). Simultaneously, we can find parametric values in the equilibrium of the election to form a central government that 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. Therefore, in this setting, the provision of a uniform public good by a single government could be welfare superior to the set of differentiated but extreme provision of local public goods in a system of local governments.[13]

Empirical Analysis

The key expectation stemming from our formal model is that the combination of democratic decentralization and party centralization will lead to the best delivery of public goods, other things equal. We test this hypothesis with a series of quantitative models of all electorally competitive countries from 1976 to 2007, contingent on data availability. Our most expansive model considers 2078 observations and

[12] A comparison of (8) and (2) shows that the electorally optimal policy for Downsian candidates in a nationwide sequential election with closed primaries are, in general, different to the size of public spending that maximizes the fiscal surplus from public spending 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 Decentralization Theorem does not hold.

[13] It is also simple to characterize conditions in which a system of local governments would produce public goods that are Pareto superior to those provided by a single nationwide government.

133 countries, to our knowledge the broadest examination of sub-national political institutions in the literature.

We employ seven different measures of educational outcomes to operationalize our dependent variable, the provision of public goods at the sub-national level. Educational outcome measures are often used in the empirical literature to denote public goods provision at the sub-national level (see, for example, Enikolopov and Zhuravskaya 2007 and Faguet and Sanchez 2008). Our specific measures come from World Bank (2010) and include *Primary School Completion Rate*, *Primary School Teachers* as a proportion of youth population, *Children out of School* as a proportion of youth population, *Preprimary Enrollment Rate*, *Primary Enrollment Rate*, *Secondary Enrollment Rate*, and *Tertiary Enrollment Rate*. We discuss how these seven measures are coded and provide some summary statistics in Table 1.

We select educational outcomes as our dependent variable because they allow us to examine both allocative efficiency gains (i.e. differentiation based on local preference – the main hypothesized benefit of democratic decentralization) and the degree to which public goods are provided in the face of spillovers (the hypothesized benefit of centralized parties). Why should this be so? First, our indicators of educational provision are subject to strong extra-jurisdictional spillover effects. All inhabitants of a country benefit from the educational attainment of their fellow citizens — in general, greater knowledge accumulation leads to reduced crime and economic improvements that spill outside the limits of any single jurisdiction. Another source of spillover effects occurs with population mobility – local residents may move outside a jurisdiction after receiving their education, and residents of neighboring jurisdictions may sometimes register for schools (especially tertiary schools) not provided in their own locales. For these reasons, the provision of education, as measured by our indicators, can be associated with a greater willingness on the part of local leaders to provide public goods in the presence of spillovers.

Our educational indicators also capture allocative efficiency effects for at least two reasons. First, improved political accountability resulting from democratic decentralization provides decentralized governments greater incentives to act in accordance with the needs and preferences of their constituents.

Table 1: Summary Statistics – The Dependent Variable (Differenced and Lagged)

Variable	Computation Method and Source	Mean	Range	Expectations
Primary School Completion Rate (Differenced)	The differenced ratio of total primary school graduates to the total population of the relevant age. (Source: World Bank)	.873	-34.19 to 46.19	Dependent Variable
Children Out of School (Differenced)	The differenced ratio of primary aged children not enrolled in school to the total population under age 14. (Source: World Bank)	-.220	-9.44 to 4.88	Dependent Variable
Primary School Teachers (Differenced)	The differenced ratio of primary school teachers to the total population under age 14. (Source: World Bank)	.027	-.776 to .614	Dependent Variable
Preprimary School Enrollment (Differenced)	The differenced ratio of total enrollment in preprimary school, regardless of age, to the total population of preprimary school age. (Source: World Bank)	1.27	-45.5 to 46.2	Dependent Variable
Primary School Enrollment (Differenced)	The differenced ratio of total enrollment in primary school, regardless of age, to the total population of primary school age. (Source: World Bank)	.618	-27.0 to 28.7	Dependent Variable
Secondary School Enrollment (Differenced)	The differenced ratio of total enrollment in secondary school, regardless of age, to the total population of secondary school age. (Source: World Bank)	1.01	-51.8 to 38.2	Dependent Variable
Tertiary School Enrollment (Differenced)	The differenced ratio of total enrollment in tertiary school, regardless of age, to the total population of tertiary school age. (Source: World Bank)	1.21	-11.8 to 26.2	Dependent Variable
Primary School Completion Rate (Lagged)	The lagged ratio of total primary school graduates to the total population of relevant age. (Source: World Bank)	80.7	0 to 125	Negative
Children Out of School (Lagged)	The lagged ratio of primary aged children not enrolled in school to the total population under age 14. (Source: World Bank)	4.57	0 to 30.1	Negative
Primary School Teachers (Lagged)	The lagged ratio of primary school teachers to the total population under age 14. (Source: World Bank)	1.57	.110 to 4.34	Negative
Preprimary School Enrollment (Lagged)	The lagged ratio of total enrollment in preprimary school, regardless of age to the total population of preprimary school age. (Source: World Bank)	48.5	.129 to 122	Negative
Primary School Enrollment (Lagged)	The lagged ratio of total enrollment in primary school, regardless of age, to the total population of primary school age. (Source: World Bank)	98.7	16.8 to 159	Negative
Secondary School Enrollment (Lagged)	The lagged ratio of total enrollment in secondary school, regardless of age, to the total population of secondary school age. (Source: World Bank)	70.1	2.40 to 162	Negative
Tertiary School Enrollment (Lagged)	The lagged ratio of total enrollment in tertiary school, regardless of age, to the total population of tertiary school age. (Source: World Bank)	26.8	.108 to 98.0	Negative

While most constituencies will prize superior educational outcomes, different sorts of practices are likely to produce these outcomes in different locales. For example, in one constituency, limited resources might best be channeled into increasing the number of teachers, whereas in another improved educational materials might be a focus. As a result, we believe that superior educational provision likely reflects (other things equal) an ability on the part of officials to consider local preferences and conditions.

Second, two recent papers (Faguet and Sanchez 2005 and Solé-Ollé and Esteller-Moré 2005) conclude that decentralization leads to better adjustment between investment patterns and local demands (in Bolivian municipalities in the first instance and Spanish provinces in the second). Similarly, in a more recent paper, Arze del Granado, Martinez-Vazquez, and McNab (2012) analyze the effects of decentralization on the composition of public expenditures for a large panel of countries and conclude that decentralization of public goods delivery is usually accompanied by an increase in educational expenditures. This finding asserts that decentralization, via greater responsiveness of public officials and preference matching, can increase allocative efficiency by altering the composition of public expenditures.

On the right side of the equation, our theory requires that we consider both the existence of elected sub-national governments and the level of party decentralization at the sub-national level. We develop an original dataset of sub-national political institutions to capture both of these measures, which we code for all countries between 1975 and 2010, where data are available. As part of this dataset, we code for the presence of elections, the structure of legislative-executive relations, the electoral system, the extent to which the national party system is replicated, and the centralization of parties at both the highest sub-national level and the municipal level (defined as the lowest level of sub-national government).[14]

To operationalize party centralization, we make use of an indicator first developed by Carey and Shugart (1995) and labeled the "Ballot" variable. This variable measures what is perhaps the most significant power that national party leaders can wield – the power to nominate candidates for office. The ballot variable is coded "0" when national party leaders have full nomination powers, "1" when they nominate a list but voters have the power to change its order (as in open list or alternate vote systems), and "2" when they have little power over nomination (as in primary systems or systems where candidates get on the ballot by collecting signatures or paying a fee). Carey and Shugart coded their variable for a small number of representative cases, and Hankla (2006) later added all democracies from 1975 to 2004. We

The dataset, along with a source list, is available upon request. Note that another comprehensive dataset of sub-national indicators has recently been produced by Ivanyna and Shah (2012), but their focus is not on the specific political institutions that we code.

expand these datasets by capturing not just the power of national party leaders to nominate candidates for the national legislature, but also whether these national party leaders can nominate candidates to regional and local assemblies.

Yet, simply coding party centralization is not enough to measure the power of national party leaders over sub-national and municipal elections. We also need to know whether national parties are competing and winning in these elections. If only 10% of a municipal council's seats are held by national parties, or if, for example, 50% of its seats are chosen in non-partisan elections, national party influence cannot be significant even if parties are centralized as defined above.

Drawing on these considerations, we create one key independent variable from our dataset to include in the model—a dummy variable, labeled *Democratic Decentralization, Party Centralization*, coded "1" when (1) there are municipal elections, and (2) more than 75% of municipal council seats are held by national parties, and (3) national party leaders exercise centralized power over municipal party nomination (i.e. party centralization is coded "0" above).[15] To our knowledge, this article is the first to consider party system nationalization and party centralization simultaneously in a large empirical model.

Additionally, we control for potentially confounding political factors by including two additional variables, also coded as part of our original dataset. The first of these is *Municipal Plurality,* coded "1" when municipal council elections are held using a plurality (as opposed to a proportional or mixed) electoral system. The second is *Municipal Directly Elected Executive,* coded "1" when municipal executives are directly elected and not removable (except through impeachment or election recall) by the municipal councils. While there is little research on the impact of these institutions at the municipal level, previous research on electoral and executive institutions at the national level indicates that they may matter

[15] Note that the extent of local party system nationalization can vary by municipality within individual countries, and that party centralization can vary by party within individual countries. There tends, all the same, to be a great deal of commonality in party system nationalization and party structure within a country, allowing for relatively straightforward coding in most cases. That said, when we encountered mixed cases, we went with what appeared to be the most common institutional structure in the country. Those country-years coded "1" on the primary dummy variable are listed in Table 3, and we are very open to feedback by country experts on the accuracy of the coding.

for public goods outcomes.[16] For example, there is reason to believe that strong unitary executives (elected and subject to reasonable legislative oversight) may produce better public outcomes than dominant legislatures (see, for example, Mukherjee 2003; Egger, Koethenbuerger, and Smart 2010; Sabatini 2003). And, although this finding is still quite contested, proportional electoral systems may have certain benefits over simple plurality systems (see, for example, Lijphart 1977). For all three political variables, of course, democratically centralized systems are among those coded "0".

We include in the model a series of economic and social control variables, namely GDP per capita in purchasing power parity, fertility rate, population density, urban population, world region dummies, and expenditure and revenue decentralization. Countries with a higher GDP per capita can be expected to have more resources to pour into education. Likewise, countries with lower fertility rates will tend to have more resources to devote to each child. Higher levels of population density and urban population should also be associated with better educational outcomes, as they facilitate access to teachers and school resources. And greater fiscal decentralization, both on the expenditure and revenue sides, should empower local governments to take a greater role in education service delivery.

In addition to these variables, we include dummies for world regions in all of the models to control for cultural or geographically specific effects. And as our theoretical model assumes elections, we also restrict our models to countries that are minimally electorally competitive using the *Legislative Index of Electoral Competitiveness* from the Database of Political Institutions (Beck et al. 2001). This index ranges from one to seven, and we only include country-years coded at least a six, indicating that more than one party holds seats in the national legislature. We choose this more expansive definition of electoral competition in preference to a measure of democracy (such as *Polity)* in order to maximize our observations and because we believe that electoral competition but not necessary full-fledged democracy must be present to observe our theorized effects.

[16] For a summary of this literature as it relates to local political institutions, see Hankla and Downs (2010).

Table 2: Summary Statistics – The Independent Variables

Variable	Computation Method and Source	Mean	Range	Expected Impact on Public Goods
Democratic Decentralization, Party Centralization (Lagged)	Coded "1" when (1) there are municipal elections, and (2) more than 75% of municipal council seats are held by national parties, and (3) national party leaders control party nomination in municipal elections. (Source: Original Dataset)	.505	Dummy	Positive (negatively related to Children Out of School, positively related to all other Y)
Municipal Plurality	Coded "1" when (1) there are municipal elections, and (2) a plurality system is used to elect the municipal assembly. (Source: Original Dataset)	.262	Dummy	Uncertain (Negative?)
Municipal Directly Elected Executive	Coded "1" when (1) there are municipal elections, and (2) the municipal mayor or other executive is directly elected and cannot be removed by the municipal council. (Source: Original Dataset)	.391	Dummy	Uncertain (Positive?)
Fertility (Differenced)	Differenced average births per woman (Source: World Bank)	-.053	-.340 to .335	Negative
Fertility (Lagged)	Lagged average births per woman (Source: World Bank)	3.24	1.08 to 7.74	Negative
Logged GDP per capita (Differenced)	Differenced Logged GDP per capita ppp (Source: World Bank)	.021	-.703 to .599	Positive
Logged GDP per capita (Lagged)	Lagged Logged GDP per capita ppp (Source: World Bank)	8.54	5.61 to 10.8	Positive
Logged Population Density (Differenced)	Differenced logged people per square kilometer (Source: World Bank)	.014	-.038 to .124	Positive
Logged Population Density (Lagged)	Lagged logged people per square kilometer (Source: World Bank)	4.01	.367 to 8.76	Positive
Urban Population (Differenced)	Differenced urban population as a percent of total (Source: World Bank)	.340	-.480 to 3.04	Positive
Urban Population (Lagged)	Lagged urban population as a percent of total (Source: World Bank)	53.5	7.02 to 100	Positive
Expenditure Decentralization (Differenced)	Differenced Percentage of Government Expenditures Undertaken at the Sub-National Level (Source: IMF)	.0008	-.210 to .196	Uncertain (Positive?)
Expenditure Decentralization (Lagged)	Lagged Percentage of Government Expenditures Undertaken at the Sub-National Level (Source: IMF)	.251	.016 to .605	Uncertain (Positive?)
Revenue Decentralization (Differenced)	Differenced Percentage of Government Revenues Collected at the Sub-National Level (Source: IMF)	.0009	-.224 to .208	Uncertain (Positive?)
Revenue Decentralization (Lagged)	Lagged Percentage of Government Revenues Collected at the Sub-National Level (Source: IMF)	.255	.008 to .617	Uncertain (Positive?)
Regional Dummies	Dummy variables for seven world regions, with Sub-Saharan Africa as the reference category	N/A	N/A	N/A
Legislative Electoral Competitiveness (Lagged)	Lagged Legislative Index of Electoral Competitiveness (Source: Database of Political Institutions)	N/A	1 to 7	Used to restrict dataset to countries with multiple parties in the national legislature (scoring 6 or 7)

Table 3: Countries Coded "1" on Democratic Decentralization, Party Centralization
(Note: Only electorally competitive country-years included; coded "1" for 1975-2006 unless otherwise stated)

Albania (1992-2006), Argentina (1975, 1984-1990), Austria, Azerbaijan (2000-2006), Benin (2002-2006), Bolivia (1987-2006), Bosnia (2003-2006), Botswana, Bulgaria (1991-2006), Burkina Faso (1995-2006), Burundi (2005-2006), Cambodia (2002-2006), Colombia (1975-1990), Costa Rica, Croatia (1993-2006), Cyprus, Denmark, Dominican Republic, Ecuador (1980-96), Egypt (1996-2006), El Salvador (1983-2006), Equatorial Guinea (1995-2006), Estonia (1993-2006), Fiji (1975-1997), Finland, France, Gabon (1997-2006), Gambia (1975-1994, 2002-2006), Georgia (1993-2006), Greece, Guatemala (1995-2006), Guinea (2005-2006), Guyana (1994-2006), Haiti (1991, 1995-2006), Honduras (1992-2006), Indonesia (2005-2006), Israel, Italy, Ivory Coast (2001-2006), Jamaica, Japan, Kazakhstan (1997-2006), South Korea (1991-2006), Kirgizia (1996-1998), Latvia (1994-2006), Lebanon (1998-2006), Liberia (1986-1989), Lithuania (1995-2006), Macedonia (1996-2006), Madagascar (1995-2006), Malawi (2000-2005), Mali (1993-2006), Mauritania (1992-2005), Mexico, Moldova (1995-2006), Moldova (1995-2006), Mongolia (2001-2006), Morocco (1992-2006), Mozambique (1998-2006), Namibia (1992-2006), Nepal (1992-2001), Nicaragua (1990-2006), Niger (2004-2006), Panama (1985-2006), Paraguay (1991-2006), Peru (1981-2006), Portugal (1977-2006), Romania (1992-2006), Russia (2000-2006), Senegal (1996-2006), Sierra Leone (2004-2006), Slovak Republic (1993-2006), Slovenia (1994-2006), Spain (1979-2006), Sri Lanka (1975-1986), Sweden, Taiwan (1994-2006), Tajikistan (2000-2006), Tanzania (1996-2006), Thailand (1976, 1980-2006), Trinidad, Tunisia (1995-2006), Ukraine (1998-2006), United Kingdom, Uzbekistan (2000-2006), Venezuela, Yemen (2001-2006), Yugoslavia (1993-2001), Zambia (1992-2006), Zimbabwe (1984-2006)

We do not control for factors, such as teacher pay, that are related directly to educational outcomes, as these are potentially part of the causal mechanism connecting political institutions with educational outcomes. The data for GDP per capita, fertility rate, urban population, and population density are from World Bank (2010), and the fiscal decentralization indicators are from International Monetary Fund (2004). Summary statistics on all of our variables are presented in Table 2, and a complete listing of all country-years coded "1" on our primary *Democratic Decentralization, Party Centralization* variable is presented in Table 3.

For our empirical analyses, we estimate error correction models with random effects and regional dummies.[17] Granger causality tests show that no endogeneity exists in the models, so an instrumental variable approach is not necessary in this context. However, Dickey-Fuller tests show that our dependent variables are non-stationary in many of our country cases. It is this fact that motivates the use of error correction models for our analysis, as the differenced dependent variable corrects for the unit root problem. In addition, an error correction framework has the benefit of allowing us to separate the short-term from the long-term effects of our independent variables. Each independent variable is included in the model in both differenced and lagged form, with the differenced variables measuring short-term effects and the lagged

[17] For more on error correction models, see Baltagi 2000.

variables measuring long-term effects. We do not difference our primary independent variables as they change only slowly across time, and so differencing would produce a large number of zeros.

For our primary models, we select a random effects framework. While a Hausman test shows that we cannot reject fixed country effects in our models, the low level of cross-temporal variable in our primary independent variables means that a fixed effects model will wash out our most interesting variation. By contrast, our random effects models allow us to deal with panel effects in a way that does not eliminate cross-country variation. For robustness, we also run our primary models with a pooled OLS estimator and country-clustered standard errors, as OLS coefficients have been shown to be quite robust in a variety of settings (see Beck and Katz 1995). For reasons of space we do not present these results, but we discuss them in the next section.

As a robustness test, we re-estimate each of our seven models after restricting the dependent variables to within two standard deviations in each direction from their means. We do this to ensure that our results are not dependent on the somewhat large outliers to be found in the differenced dependent variables. We make use of robust standard errors in each of our models, and we include a lagged dependent variable, as is customary in error correction models. We present the results of our seven primary models in Table 4, of those same models without outliers in Table 5, and of the models with fiscal decentralization controls in Table 6.

Table 4: Results of the Primary Models
(Error Correction Models with Random Effects, Robust Standard Errors, and Regional Dummies)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	Y= Diff. Primary Completion Rate (N=1290, 120 countries)	Y= Diff. Primary School Teachers (N=1523, 130 countries)	Y= Diff. Preprimary Enrollment Rate (N=1795, 125 countries)	Y= Diff. Primary Enrollment Rate (N=2078, 133 countries)	Y= Diff. Secondary Enrollment Rate (N=1808, 130 countries)	Y= Diff. Tertiary Enrollment Rate (N=1612, 127 countries)	Y= Diff. Children Out of School (N=1293, 117 countries)
Lagged Dependent Variable	-0.078*** (0.019)	-.006 (.007)	-.024*** (.007)	-.065 (.014)	-.035 (.008)	.001 (.006)	-0.083*** (0.022)
Democratic Decentralization, Party Centralization (Lagged)	0.452** (0.225)	.008* (.005)	.376* (.226)	.427 (.364)	.358** (.152)	.102 (.151)	-0.177 (0.163)
Municipal Plurality (Lagged)	-0.028 (0.313)	-.004 (.005)	-.019 (.295)	-.161 (.470)	-.319* (.172)	.051 (.216)	0.150 (0.100)
Municipal Directly Elected Executive (Lagged)	-0.086 (0.240)	.009 (.006)	.100 (.222)	.289 (.374)	-.146 (.169)	.421*** (.151)	-0.035 (0.128)
Fertility (Differenced)	-3.08 (2.20)	-.051 (.039)	3.00 (1.88)	-.564 (1.99)	-1.75 (2.20)	.801 (1.21)	-0.019 (0.574)
Fertility (Lagged)	-0.649** (0.285)	-.002 (.003)	-.457*** (.154)	-.688*** (.234)	-.665*** (.158)	-.197* (.110)	0.183** (0.086)
Logged GDP per capita (Differenced)	1.951 (2.10)	.085* (.049)	5.65** (2.27)	2.46 (2.32)	-.859 (2.11)	4.11** (1.69)	-1.07 (0.888)
Logged GDP per capita (Lagged)	-0.097 (0.225)	.001 (.004)	.303 (.219)	-.798** (.341)	.041 (.136)	.197 (.177)	0.107 (0.139)
Logged Population Density (Differenced)	5.53 (21.2)	.054 (.341)	18.1 (17.4)	46.9* (25.1)	11.7 (15.0)	.405 (15.5)	-17.0* (9.70)
Logged Population Density (Lagged)	-0.065 (0.100)	-.001 (.001)	.035 (.092)	.241 (.151)	-.070 (.063)	-.118** (.056)	-0.057 (0.042)
Urban Population (Differenced)	-0.058 (0.439)	.008 (.006)	.310 (.305)	-.076 (.313)	.137 (.197)	-.267* (.138)	0.119 (0.117)
Urban Population (Lagged)	0.007 (0.011)	-.00003 (.0002)	.0005 (.009)	-.006 (.014)	.005 (.007)	.007 (.007)	0.004 (0.007)
R^2	.079	.020	.037	.111	.045	.129	.141

***p<.01, **p<.05, *p<.10. All tests are 2-tailed. Robust standard errors are in parenthesis

Table 5: Results of the Primary Models with Outliers Eliminated
(Error Correction Models with random effects, robust standard errors, and regional dummies, and with the dependent variable limited to two standard deviations from its mean.)

Variable	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
	Y= Diff. Primary Completion Rate (N=1232, 120 countries)	**Y= Diff. Primary School Teachers (N=1450, 130 countries)**	**Y= Diff. Preprimary Enrollment Rate (N=1721, 124 countries)**	**Y= Diff. Primary Enrollment Rate (N=1981, 131 countries)**	**Y= Diff. Secondary Enrollment Rate (N=1743, 130 countries)**	**Y= Diff. Tertiary Enrollment Rate (N=1538, 126 countries)**	**Y= Diff. Children Out of School (N=1228, 113 countries)**
Lagged Dependent Variable	-.053*** (.010)	-.010*** (.006)	-.008* (.004)	-.028*** (.006)	-.015*** (.005)	.008** (.003)	-.044*** (.011)
Democratic Decentralization, Party Centralization (Lagged)	.450** (.203)	.007** (.004)	.325* (.175)	.313 (.220)	.313** (.141)	.016 (.103)	-.099 (.074)
Municipal Plurality (Lagged)	.010 (.237)	-.002 (.005)	-.231 (.223)	-.053 (.252)	-.221 (.156)	-.010 (.136)	.012 (.052)
Municipal Directly Elected Executive (Lagged)	-.057 (.191)	.006 (.004)	.081 (.204)	.275 (.205)	-.023 (.169)	.257** (.114)	-.078 (.068)
Fertility (Differenced)	-2.12 (1.86)	-.044 (.028)	2.13* (1.18)	-.006 (1.21)	-1.96 (1.47)	.327 (1.05)	-.465 (.360)
Fertility (Lagged)	-.220 (.173)	-.005* (.003)	-.134 (.121)	-.123 (.128)	-.474*** (.127)	-.195** (.094)	.063 (.047)
Logged GDP per capita (Differenced)	.412 (1.90)	.055 (.033)	4.30** (1.77)	2.02 (1.31)	.330 (1.65)	2.67** (1.28)	-.118 (.366)
Logged GDP per capita (Lagged)	-.036 (.189)	.003 (.005)	.159 (.169)	-.511*** (.168)	-.082 (.140)	.091 (.120)	.050 (.050)
Logged Population Density (Differenced)	-3.48 (19.4)	.126 (.258)	6.16 (13.3)	12.7 (12.1)	5.71 (12.0)	-1.77 (11.2)	-5.97** (2.78)
Logged Population Density (Lagged)	.010 (.067)	.001 (.002)	-.106 (.081)	-.008 (.077)	-.084 (.055)	-.089* (.046)	.016 (.019)
Urban Population (Differenced)	-.005 (.345)	.001 (.005)	.577** (.265)	-.168 (.223)	.120 (.181)	-.127 (.112)	-.022 (.043)
Urban Population (Lagged)	.006 (.009)	-.0001 (.0002)	.003 (.006)	.002 (.007)	-.001 (.005)	.001 (.003)	.002 (.003)
R^2	.085	.020	.071	.111	.051	.182	.117

***p<.01, **p<.05, *p<.10. All tests are 2-tailed. Robust standard errors are in parenthesis

Results

The quantitative results provide strong support for our hypothesis. Beginning with the primary models in Table 4, the *Democratic Decentralization, Party Centralization* variable is significant in four of the seven models and in the expected direction but not significant in the other three. Note that we have estimated models with all of the theoretically relevant education variables that are available for a wide

swath of countries and found in World Bank (2010). Note also that the various models differ significantly in the number of N's included, the specific observations, and in the nature of the dependent variable. For these reasons, the consistency of the results across the seven models is quite striking and indicates that countries enjoying both democratic decentralization and party centralization tend to produce better educational outcomes. Moreover, when the models are rerun using pooled OLS and country-clustered standard errors, the primary results are largely unchanged. The only significant difference that arises with the new estimator is that the *Democratic Decentralization, Party Centralization* variable comes closer to statistical significance in the *Children Out of School* Model (p=.114). And the size of these predicted effects is also worthy of consideration. Model 1, for instance, estimates that the combination of democratic decentralization and party centralization causes an increase in differenced primary school enrollment of nearly half a percent, or a little more than one-tenth of a standard deviation.

Table 5 presents the results of the primary models with their outliers dropped, i.e. restricted only to observations where the dependent variable is within two standard deviations above or below its mean. The results of the primary independent variable hold fully, and the variable's statistical significance is in fact strengthened in the *Primary School Teachers* model.[18] Indeed, of the seven dependent variables considered in our models, it is only in one case – *Primary Enrollment Rate* – that our primary independent variable never achieves or comes close to achieving statistical significance. Why might that be the case? We

[18]In Table 6, located in our online appendix for space reasons, we also present models with the expenditure and revenue decentralization variable included. We estimate these models separately because the fiscal decentralization variables are missing for many observations and severely restrict our N's. Here again, the *Democratic Decentralization, Party Centralization* variable is in the expected direction in all models, but it retains significance in two of the seven (including, for the first time, *Tertiary Enrollment Rate)*. In our judgment, the fact that the primary variable remains significant in two of the models and is always in the expected direction, despite the greatly restricted sample size, represents at least moderately strong confirmatory evidence for our argument. The most striking finding related to the fiscal decentralization variables is that the impact of expenditure decentralization on educational provision tends to be negative (especially in the short-run), while the impact of revenue decentralization tends to be positive. Perhaps this is an indication that the devolution of expenditure powers without the ability to raise money independently does not produce the expected gains; under these conditions, local governments struggle in the short run to develop the capacity to manage devolved funds. The presence in only one model of a negative long-term relationship between expenditure decentralization and educational outcomes is evidence that these problems often are eventually resolved. On the other hand, the ability of local governments to raise revenue is perhaps a better measure of their genuine independence and responsibility, and therefore more associated with a more accountable public goods provision. Indeed, a number of scholars have argued that the hypothesized benefits of decentralization only accrue when sub-national governments have the power to raise their own revenues (see, for example, Weingast (2009) and Rodden (2003).

suspect that the greater (although far from perfect) universality of primary education means that political factors such as the ones we consider will impact its provision less dramatically that the provision of, say, preprimary or secondary education.

Which control variables matter for educational outcomes? Perhaps the most interesting findings are that (1) municipalities with plurality electoral systems tend to see weaker improvements in the provision of educational public goods, while (2) municipalities with directly elected executives (mayors) tend to see stronger improvements.[19] These relationships do not appear in all of the models, but they are common enough to suggest some support for the prior literature on the benefits of direct election and electoral proportionality.

In evaluating the economic and social control variables, it is important to remember that the lagged variables represent long-term effects on the change in the dependent variable while the differenced variables represent short term effects. Turning to the *Fertility* variables, higher levels of fertility are associated with a long-term trend toward lower educational provision, as expected. The short-term impact of *Fertility* is much less in evidence, which stands to reason given the time that it takes shifts in fertility to impact educational demand. The results also indicate that increases in differenced GDP per capita are associated with improvements in education. It is worth noting, however, that the lagged (long-term) measure of GDP per capita is associated in several models with weaker improvement in public goods provision. This combination of results likely indicates that short-term improvements in economic growth will translate into improved educational outcomes, but that, over the long term, rich countries will stop seeing educational improvements after they have achieved a certain level. In other words, poor but growing economies can expect more dramatic educational improvements than rich but stagnant ones. It is also worth noting that the only model where long-term high GDP per capita improves outcomes is the *Tertiary Enrollment Rate*. This finding is not surprising, given that countries are less likely to achieve full

[19] Note that a negative sign in the *Children Out of School* models indicates a better public goods outcome. Note also that the negative effect of plurality elections is most apparent in the models in Table 6, found in the appendix.

enrollment in tertiary education after a certain level of development than they are in primary or secondary education.

There is evidence of the short-term benefits of higher *Population Density* on improving educational outcomes, as expected. This effect is enhanced because increases in population are likely translate (albeit indirectly) into a larger denominator for some of the dependent variables. In the long-run, however, high *Population Density* seems to have a negative impact on tertiary enrollment. This result could be because countries that are already quite dense experience weaker improvements in tertiary education. The short-term impact of *Urban Population* is surprising, where it shows up in one *Preprimary Enrollment* model as positive, and in one *Tertiary Enrollment Rate* model as negative. This result may indicate that more increasingly urbanized societies are more likely to provide the somewhat unusual service of preschool education, but this growth in urbanization can also indicate economic and social weaknesses that may inhibit other forms of educational provision. But in the final analysis, differenced *Urban Population* is not significant in enough models to draw a clear conclusion.

What can we say to summarize the results? The benefits of combining democratic decentralization with party centralization are well borne out in our empirical analysis. The *Democratic Decentralization, Party Centralization* variable is statistically significant in eight of the fourteen models, and it is always in the expected direction. Its lack of significance in some of the models may give pause in making final judgments. That said, the indicator shows up as significant across models using two different estimators, it is able to explain at least five very different measures of educational attainment, and it is robust to vastly different combinations of country-years. Given the difficulty of measuring educational outcomes, particularly in the developing world, and given the complexity and specificity of local politics in different countries, the robustness of the results provided here are notable. That said, our theory and others that draw general lessons from sub-national political institutions will continue to require refinement in the light of ongoing empirical tests.

Conclusion

In this paper, we examine which types of political institutions may be necessary to deliver the gains from decentralization predicted by much of the literature. We begin by developing a formal extension and refinement of the decentralization theorem of Oates (1972), which has provided the basis for much past research. Assuming inter-jurisdiction spillovers (something Oates did not do), we develop a "strong decentralization theorem" which indicates the potential superiority of sub-national governance even in the presence of externalities. An important implication of our theoretical model is that the combination of democratic decentralization and party centralization tends to produce the most efficient 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.

To test our argument empirically, we create a large dataset on sub-national political institutions and use it to estimate a series of cross-national empirical analysis of educational outcomes. Our dataset is, to our knowledge, the first to compile measures of sub-national political institutions across a large set of countries. Our empirical findings provide strong support for our hypotheses. They show that the combination of municipal elections and party centralization tends to improve educational outcomes.

Our ultimate goal in this paper is to understand better how the growing influence of sub-national fiscal decentralization may impact the everyday lives of citizens around the world depending on the arrangements used for political decentralization. There is much additional research to be done, not least in refining our theory on the role of decentralized political institutions and applying it to new policy areas. This paper also shows the potential of merging political science and economics into broader approaches to explore the inter-connected dynamics of decentralized governance.

Works Cited

Aldrich, J.H., and R.M., Alvarez. 1994. Issues and the presidential primary voter. *Journal of Political Behavior*, Vol. 16, No. 3, pp. 289-317.

Arze del Granado, F. Javier, Jorge Martinez-Vazquez, and Robert M. McNab. 2012. Decentralized Governance and Preferences for Public Goods. Working paper, International Center for Public Policy, Andrew Young School, Georgia State University.

Baltagi, Badi H., ed. 2000. *Nonstationary Panels, Panel Cointegration, and Dynamic Panels.* New York: Elsevier Science.

Bardhan, Pranab. 2002. Decentralization of governance and development. *Journal of Economic Perspectives* 16 (4): 185–205.

Bardhan, Pranab and D. Mookherjee. 2000. Capture and governance at local and national levels. *The American Economic Review* 90 (2): 135-139.

Barro, Robert J and Jong-Wha Lee. 2010. A New Data Set of Educational Attainment in the World, 1950-2010. NBER Working Paper 15902.

Beck, Nathaniel and Jonathan N. Katz. 1995. What to do (and not to do) with Time-Series Cross-Section Data. *American Political Science Review* 89 (3): 634-647.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. New Tools in Comparative Political Economy: The Database of Political Institutions. *World Bank Economic Review* 15(1): 165-176.

Besley, T., and S. Coate, (2003), Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach, *Journal of Public Economics* 87, pp. 2611-2637.

Bird, Richard M. and Francois Vaillancourt. 1998. Fiscal Decentralization in Developing Countries: An Overview. In Richard M. Bird and Francois Vaillancourt, eds. *Fiscal Decentralization in Developing Countries*. New York: Cambridge University Press.

Bird, Richard, Bernard Dafflon, Claude Jeanrenaud, and Gebhard Kirchgässner. 2003. Assignment of responsibilities and fiscal federalism. In *Federalism in a changing world: Learning from each other*, ed. Raoul Blindenbacker, and Arnold Koller. Ithaca: McGill-Queens University Press.

Breton, Albert. 2002. An introduction to decentralization failure. In *Managing fiscal decentralization*, ed. Ehtisham Ahmad, and Vito Tanzi. New York: Routledge.

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.

Crook, Richard Charles. 2003. *Decentralisation and Good Goverance. In Federalism in a changing world: Learning from each other, ed.* Raoul Blindenbacker, and Arnold Koller. Ithaca: McGill-Queens University Press.

Davoodi, Hamid, and Heng-fu Zou. 1998. Fiscal Decentralization and Economic Growth: A Cross-Country Study. *Journal of Urban Economics* 43: 244–257.

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.

Dixit, A., and J. Londregan. 1998. Fiscal Federalism and Redistributive Politics. *Journal of Public Economics 68*, pp. 153-180

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

Egger, Peter, Marko Koethenbuerger, and Michael Smart. 2010. Electoral Rules and Incentive Effects of Fiscal Transfers: Evidence from Germany. Documents de Treball de l'Institut d'Economia de Barcelona 2010/44.

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.

Faguet, Jean-Paul, and Fabio Sanchez. 2008. Decentralization's effects on educational outcomes in Bolivia and Columbia. *World Development*, 36 (7): 1294-1316

Ferejohn, J. 1974. *Pork Barrel Politics: Rivers and Harbors Legislation 1947– 1968*. Stanford University Press, Stanford, CA.

Filippov, Mikhail, Peter C. Ordeshook, and Olga Shvetsova. 2004. *Designing Federalism: A Theory of Self-Sustainable Federal Institutions*. New York: Cambridge University Press.

Garman, Christopher, Stephen Haggard, and Eliza Willis. 2001. Fiscal Decentralization: A Political Theory with Latin American Cases. *World Politics* 53, 205-36.

Gurgur, Tugrul, and Anwar Shah. 2002. Localization and corruption: Panacea or pandora'sbox? In *Managing Fiscal Decentralization*, ed. Ehtisham Ahmad, and Vito Tanzi. New York: Routledge.

Habibi, Nadir, Cindy Huang, Diego Miranda, Victoria Murillo, Gustav Ranis, Mainak Sarkar, and Frances Stewart. 2003. Decentralization and Human Development in Argentina. *Journal of Human Development and Capabilities* 4 (1): 73-101.

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.

Hankla, Charles R. 2006. Party Strength and International Trade: A Cross-National Analysis. *Comparative Political Studies* 39 (9): 1133-1156.

Hankla. Charles R. 2009. "When is Fiscal Decentralization Good for Governance?" *Publius: The Journal of Federalism* 39 (4): 632-650

Hankla, Charles R. and William M. Downs. 2010. "Decentralization, Governance, and the Structure of Local Political Institutions: Assessing the Evidence" *Local Government Studies* 36 (6): 759-783.

Hecock, R. Douglas. 2006. Electoral Competition, Globalization, and Subnational Education Spending in Mexico, 1999-2004. *American Journal of Political Science* 50 (4): 950-961.

Heredia-Ortiz, Eunice. 2006. The Impact of Education Decentralization on Education Output: A Cross-Country Study. Dissertation, Department of Economics, Georgia State University.

Hicken, Allen and Joel W. Simmons. 2008. The Personal Vote and the Efficiency of Education Spending. *American Journal of Political Science* 52 (1): 109-124.

International Monetary Fund. 2001. *Government Financial Statistics Manuel.*

Ivanyna, Maksym and Anwar Shah. 2012. Is your government closer to the people? Worldwide Indicators on Localization and Decentralization. Working Paper.

Kaufman, K.M. and J.G Gimpel. 2003. A Promise Fulfilled? Open Primaries and Representation. *The Journal of Politics*, Vol. 65, No. 2, pp. 457-476.

Lewis, Blaine. 1998. The impact of public infrastructure on municipal economic development: Empirical results from Kenya. *Review of Urban and Regional Development Studies* 10 (2): 142–456.

Lijphart, Arend. 1977. *Democracy in Plural Societies*. New Haven: Yale University Press.

Lockwood, B. 2002. Distributive politics and the benefits of decentralization. *Review of Economic Studies* 69, No. 2, 313– 338.

Manor, James. 1999. *The Political Economy of Democratic Decentralization*. Washington, DC: The World Bank.

Marshall, Monty G. and Keith Jaggers. 2000. Polity IV Project: Political Regime Characteristics and Transitions, 1800-1999. *Dataset Users Manuel.* College Park: University of Maryland, College Park.

Martinez-Vazquez, Jorge. 1982 "Fiscal Incidence at the Local Level." *Econometrica* 50, no. 5 (September):1207-1218.

Mukherjee, Bumba. 2003. Political Parties and the Size of Government in Multiparty Legislatures: Examining Cross-Country and Panel Data Evidence. *Comparative Political Studies* 36 (6): 699-728.

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.

Organization for Economic Cooperation and Development. 2004. *Education at a Glance: OECD Indicators 2004*. Paris and Washington, DC: OECD.

O'Halloran, Sharyn. 1994. *Politics, process, and American trade policy*. Ann Arbor: University of Michigan Press.

Parry, Taryn. 1997. Achieving balance in decentralization: A case study of education decentralization in Chile. *World Development* 25 (2): 211–225.

Prud'Homme, Re´my. 1995. The dangers of decentralization. *World Bank Research Observer* 10: 201–220.

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

Riker, William H. 1987. *The Development of American Federalism.* Boston: Kluwer Academic Publishers.

Rodden, Jonathan. 2003. Reviving Leviathan: Fiscal Federalism and the Growth of Government. *International Organization* 57: 695-729.

Rodden, Jonathan. 2006. *Hamilton's Paradox: The Promise and Peril of Fiscal Federalism*. New York: Cambridge University Press.

Rogowski, Ronald. 1987. Trade and the Variety of Democratic Institutions. *International Organization* 41 (2): 202-223.

Roubini, Nouriel and Jeffrey Sachs. 1989. Government Spending and Budget Deficits: Industrial Democracies. Working Paper No. 2919. National Bureau of Economic Research, Cambridge, MA.

Sabatini, Christopher. 2003. Decentralization and Political Parties. *Journal of Democracy* 14 (2): 138-150.

Schwartz, B. J , David Guilkey, and Rachel Racelis. 2002. Decentralization, allocative efficiency and health service outcomes in the Philippines. In *MEASURE Evaluation*. Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill.

Shah, Anwar. 2003. Fiscal decentralization in transition economies and developing countries. In *Federalism in a Changing World: Learning From Each Other*, ed. Raoul Blindenbacker,and Arnold Koller. Ithaca: McGill-Queens University Press.

Simatupang, Renata. 2009. Evaluation of Decentralization Outcomes in Indonesia: Analysis of Health and Education Sectors. Dissertation, Department of Economics, Georgia State University.

Solé-Ollé, A. and Esteller-Moré, A. 2005. Does Decentralization improve the efficiency in the allocation of public investment? Evidence from Spain, Working Paper 2005/5, Institut d'Economia de Barcelona.

Strumpf, Koleman, Paul Hutchinson and John S. Akin. 1999. *Decentralization and Government Provision of Public and Private Goods: The Public Health Sector in Uganda.* Working Paper, *MEASURE Evaluation* Project. Chapel Hill, NC: University of North Carolina at Chapel Hill.

Tanzi, Vito. 2002. Pitfalls on the road to fiscal decentralization. In *Managing fiscal decentralization*, ed. Ehtisham Ahmad, and Vito Tanzi. New York: Routledge.

Tiebout, C.M. 1956. A Pure Theory of Local Expenditure. *Journal of Political Economy* 64: 416-424.

Treisman, Daniel. 1999. Political decentralization and economic reform: A game theoretical analysis. *American Journal of Political Science* 43 (2): 488–517.

Treisman, Daniel. 2000. Decentralization and Inflation: Commitment, Collective Action, or Continuity? *American Political Science Review* 94(4): 837-858.

Treisman, Daniel. 2007. *The Architecture of Government: Rethinking Political Decentralization.* New York: Cambridge University Press.

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.

Von Braun, Joachim and Ulrike Grote. 2002. Does Decentralization Serve the Poor? In Ehtisham Ahmad and Vito Tanzi, eds. *Managing Fiscal Decentralization*. New York: Routledge.

Weingast, Barry R. 1995. The economic role of political institutions: Market-preserving federalism and economic growth. *Journal of Law, Economics, and Organization* 11: 1–31.

Weingast, Barry R. 2009. Second Generation Fiscal Federalism: The Implications of Fiscal Incentives. *Journal of Urban Economics 65: 279-293.*

Wibbels, Erik. 2000. Federalism and the Politics of Macroeconomic Policy and Performance. *American Journal of Political Science* 44(4): 687-702.

Wibbels, Erik. 2005. *Federalism and the Market: Intergovernmental Competition and Economic Reform in the Developing World.* New York: Cambridge University Press.

World Bank. 2010. *World Development Indicators.* Washington, DC: The World Bank.

Yilmaz, Serdar. 1999. The impact of fiscal decentralization on macroeconomic performance. Paper prepared for the 92nd Annual Conference on Taxation, Atlanta, Georgia.

Appendix

Lemma 1 *Local public goods are Pareto efficient for an economy with a majoritarian electoral system, single member districts, a single unit of government, a centralized party system, and for* $k^{-i} \in [0,1) \forall -i, i$. *All parties converge in providing a uniform local public good across districts,* $g_c^{*zi} = g_c^{*z,-i} = g_c^* \ \forall\, z$ *satisfying*

$$\sum_{\forall\, i,-i} \left(1 + k^{-i}\right) \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial G^i} de^i = \left\{\frac{1}{N}\right\} \sum_{\forall i,-\, i} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial x^i} de^i = 0 \qquad (16)$$

Proof

The parties' problem is $Max\ \delta_c^z(\mathbf{g_c^z}, \xi_c^z) = \pi_c^z(\rho_c^z)$ subject to $g_c^{zi} = g_c^{z,-i} = g_c^z$. We impose the equality restriction in the objective function of party z. The first order conditions is $\frac{\partial \delta_c^z}{\partial g_c^z} = \frac{\partial \phi_c^z}{g_c^z} = \sum_{\forall\, i,-i} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_c^{zi}(\Psi_c^{zi})}{\partial \Psi_c^{zi}} \frac{\partial v^{zi}}{\partial G^{zi}} \frac{\partial G^{zi}}{\partial g_c^z} de^i = 0 \ \ \forall g_c^{*z} > 0$.The parties' policies converge, $g_c^{*z} = g_c^{*,-z} = g_c^*$, in probabilistic voting models with homogeneous parties (see Coughlin 1992) hence $f_c^{zi}(0) = f_c^{z,-i}(0) \in \mathbb{R}_+ \ \forall\, i, \forall\, z$. Therefore the first order condition becomes $\sum_{\forall\, i,-i} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial v^i}{\partial G^i} \frac{\partial G^i}{\partial g_c} de^i = 0$. Use $\frac{\partial v^i}{\partial G^i} \frac{\partial G^i}{\partial g_c} = \left(1 + k^{-i}\right) \frac{\partial \mu^i}{\partial G^i} - \left(\frac{1}{N}\right) \frac{\partial \mu^i}{\partial x^i}$ and $\frac{\partial v^i}{\partial G^{-i}} \frac{\partial G^{-i}}{\partial g_c} = \left(1 + k^i\right) \frac{\partial \mu^{-i}}{\partial G^{-i}} - \left(\frac{1}{N}\right) \frac{\partial \mu^{-i}}{\partial x^{-i}}$ to show that the uniform local public good g_c^* satisfies

$$\sum_{\forall\, i,-i} \left(1 + k^{-i}\right) \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial G^i} de^i = \left\{\frac{1}{N}\right\} \sum_{\forall i,-\, i} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial x^i} de^i \qquad (17)$$

■

Lemma 2 *Party centralization in a system of local governments leads to a set of Pareto efficient local public goods* $\mathbf{g_{cL}^*} = \left[g_{cL}^{*i}, g_{cL}^{*-i}\right]$ for all $k^{-i} \in [0,1)\ \forall -i, i, \forall z$. *At the political equilibrium,* $g_{cL}^{*i}\ \forall i, \forall z$ *satisfies the following:*

$$\int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial G^i} de^i + k^{-i} \int_{\forall\, e^{-i}} h^{-i}\left(e^{-i}\right) \frac{\partial \mu^{-i}}{\partial G^{-i}} de^{-i} = \left\{\frac{1}{N}\right\} \int_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial \mu^i}{\partial x^i} de^i \qquad (18)$$

Proof

In the local election of district i party z selects $g_{cL}^{*zi} \in \arg\max\, \pi_{cL}^z = \pi_{cL}^z\left(\rho_{cL}^{zi}, \rho_{cL}^{z,-i}\right)$. The first order condition for an interior maximizer with $g_{cL}^{*zi} > 0$ is $\frac{\partial \pi_{cL}^z}{\partial \rho_{cL}^{zi}} \frac{\partial \rho_{cL}^{zi}}{\partial g_{cL}^{zi}} + \frac{\partial \pi_{cL}^z}{\partial \rho_{cL}^{z,-i}} \frac{\partial \rho_{cL}^{z,-i}}{\partial g_{cL}^{zi}} = 0$. By definition $\rho_{cL}^{zi} = \phi_{cL}^{zi} -$

ϕ_{cL}^{-zi} and the sum of the expected proportion of the votes for parties z and $-z$ is 1, that is, $\phi_{cL}^{zi} + \phi_{cL}^{-zi} = 1 \ \forall \ i$ therefore $\frac{\partial \rho_{cL}^{zi}}{\partial g_{cL}^{zi}} = 2\frac{\partial \phi_{cL}^{zi}}{\partial g_{cL}^{zi}}$ for $g_{cL}^{*zi} \ \forall i, \forall z$. Then $\frac{\partial \phi_{cL}^{zi}}{\partial g_{cL}^{zi}} = \int_{\forall \ e^i} h^i(e^i) f_{cL}^{zi} \left\{\frac{\partial \mu^i}{\partial G^i} - \left(\frac{1}{N}\right)\frac{\partial \mu^i}{\partial x^i}\right\} de^i$ and $\frac{\partial \phi_{cL}^{z,-i}}{\partial g_{cL}^{zi}} = k^i \int_{\forall \ e^{-i}} h^{-i}(e^{-i}) f_{cL}^{z,-i} \frac{\partial \mu^{z,-i}}{\partial G^{zi}} de^{-i}$. The convergence of the parties' policies $g_{cL}^{*zi} = g_{cL}^{*-zi} = g_{cL}^{*i}$, $\forall \ z, -z \in i$ implies $f_{cL}^{zi}(0) = f_{cL}^{z,-i}(0) \in \mathbb{R}_+ , \rho_{cL}^{zi} = \rho_{cL}^{z,-i} = 0 \ \forall \ i, \forall \ z$. Define $\Theta^i = \frac{\partial \pi_{cL}^{z}(0)}{\partial \rho_{cL}^{z,-i}} \Big/ \frac{\partial \pi_{cL}^{z}(0)}{\partial \rho_{cL}^{zi}} = 1, \forall i$ therefore $\frac{\partial \pi_{cL}^{z}}{\partial \rho_{cL}^{zi}}\frac{\partial \rho_{cL}^{zi}}{\partial g_{cL}^{zi}} + \frac{\partial \pi_{cL}^{z}}{\partial \rho_{cL}^{z,-i}}\frac{\partial \rho_{cL}^{z,-i}}{\partial g_{cL}^{zi}} = 0$ is equivalent to

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

■

Theorem 1 *The provision of local public goods with and without inter-regional spillovers by a system of local governments welfare-dominates the centralized provision.*

Proof

It is simple to verify that conditions (2), (4) and (6) imply $\hat{\mathbf{g}}^* = \mathbf{g}_{\mathbf{cL}}^* \neq \mathbf{g}_{\mathbf{c}}^*$ for $\hat{\mathbf{g}}^*, \mathbf{g}_{\mathbf{cL}}^*, \mathbf{g}_{\mathbf{c}}^* \in \mathbb{R}^2$: $\hat{\mathbf{g}}^* = \left[\hat{g}^{*i}, \hat{g}^{*,-i}\right]$, $\mathbf{g}_{\mathbf{cL}}^* = \left[g_{cL}^{*i}, g_{cL}^{*-i}\right]$ and $\mathbf{g}_{\mathbf{c}}^* = \left[g_c^{*i}, g_c^{*-i}\right]$. By the strict concavity of $NSW = \sum_{\forall \ i,-i} \int_{\forall \ e^i} h^i(e^i) v^i(e^i, G^i) de^i$ on the constrained policy space ∃ feasible $\mathbf{g^o} = \left[g^{i0}, g^{-i0}\right], \mathbf{g^1} = \left[g^{i1}, g^{-i1}\right], \Omega \in [0,1]$: $\Omega = \frac{1}{g^{i1} - g^{io}} = \frac{1}{g^{-i1} - g^{-io}}$ and $\mathbf{g^\Omega} = \Omega\mathbf{g^1} + (1-\Omega)\mathbf{g^o}$ such that $NSW(\mathbf{g^\Omega}) > \{\Omega NSW(\mathbf{g^1}) + (1-\Omega)\text{NSW}(\mathbf{g^o})\}$. Note $\Omega NSW(\mathbf{g^1}) + (1-\Omega)\text{NSW}(\mathbf{g^o}) = \text{NSW}(\mathbf{g^o}) + \Omega\{NSW(\mathbf{g^1}) - \text{NSW}\mathbf{go}$ which implies $NSW\mathbf{g\Omega} - NSW\mathbf{go} > \Omega NSW\mathbf{g1} - \text{NSW}\mathbf{go}$. We take $Limgi1 - gio \forall i \rightarrow 0$ $\Omega\{NSW(\mathbf{g^1}) - \text{NSW}(\mathbf{g^o})\} = \frac{\partial NSW(\mathbf{g})}{\partial g^i} + \frac{\partial NSW(\mathbf{g})}{\partial g^{-i}}$ to show

$$NSW(\hat{\mathbf{g}}^*) - NSW(\mathbf{g}) > \frac{\partial NSW(\mathbf{g})}{\partial g^i} + \frac{\partial NSW(\mathbf{g})}{\partial g^{-i}} \tag{19}$$

Without of loss of generality, set $\mathbf{g^\Omega} = \hat{\mathbf{g}}^* = \left[\hat{g}^{*i}, \hat{g}^{*,-i}\right]$, where $\hat{g}^{*i}$ and $\hat{g}^{*,-i}$ are global maximizers of NSW and $\mathbf{g^o}$ is some feasible $\mathbf{g} = \left[g^i, g^{-i}\right]$: $\mathbf{g} \neq \hat{\mathbf{g}}^*$ then it is satisfied that $\frac{\partial NSW(\hat{\mathbf{g}}^*)}{\partial g^i} = \frac{\partial NSW(\hat{\mathbf{g}}^*)}{\partial g^{-i}} = 0 \ \forall \ \hat{g}^{*i} > 0 \ \forall i \implies NSW(\hat{\mathbf{g}}^*) > NSW(\mathbf{g}) \ \ \forall \ g^i \neq \hat{g}^{*i}, \forall i$. Therefore conditions (2), (4) and (6) mean $\mathbf{g}_{\mathbf{cL}}^* = \hat{\mathbf{g}}^*$ and $\mathbf{g}_{\mathbf{c}}^* \neq \hat{\mathbf{g}}^*$ and hence $NSW(\mathbf{g}_{\mathbf{cL}}^*) > NSW(\mathbf{g}_{\mathbf{c}}^*)$.

■

Lemma 3 *For economies with a decentralized party system and a nationwide single government representing voters of all districts, 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:*

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

$$= -\gamma^z \left(\sum\nolimits_{\forall\, i,-i} \int\limits_{\forall\, \tilde{e}^i} \tilde{h}^i\left(\tilde{e}^i\right) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \tilde{v}_g^{zi}(g_d^{*z}) d\tilde{e}^i \right) - \sigma_\omega^z \qquad (20)$$

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

$$\alpha^{zi} = \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (21)$$

$$\gamma^z = \frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (22)$$

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

$$\boldsymbol{\sigma_\omega^z} = \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right) \Bigg/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (23)$$

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

Proof

For an economy with party decentralization and a nationwide government, party z designs public spending to maximize $\Pi_d^z(\tilde{\rho}_0^z,\ \rho_1^z, \rho_2^z)$ subject to $g_d^{zi} = g_d^{z,-i} = g_d^z$. We impose the equality restriction in the objective function of party z. The first order condition for the party's problem is

$$\frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \frac{\partial \tilde{\phi}_0^z}{\partial g_d^z} + \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \frac{\partial \phi_l^z}{\partial g_d^z} = 0 \quad \forall\ g_d^{*z} > 0, \forall z, \forall i \qquad (24)$$

By definition $\tilde{\phi}_0^z = \sum_{\forall i,-i} \tilde{\phi}_0^{zi}$, $\phi_l^z = \sum_{\forall i,-i} \phi_l^{zi}$ for $l = 1{,}2$. The sum of the expected votes for parties z and $-z$ in both the primary and the general election is one then $\tilde{\phi}_0^z + \tilde{\phi}_0^{-z} = 1$, and $\phi_l^z + \phi_l^{-z} = 1\ \forall l$. Since

$g_d^{*zi} = g_d^{*z,-i} = g_d^{*z}$ then $\frac{\partial \tilde{\phi}_0^{zi}}{\partial g_d^z} = \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_d^z} de^i \; \forall i$ and $\frac{\partial \phi_l^{zi}}{\partial g_d^z} = \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \frac{\partial \Psi_l^{zi}}{\partial g_d^z} de^i$ for $l = 1,2, \forall i$. It follows that the first order condition is given by

$$\sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \left\{ \sum_{\forall i,-i} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \frac{\partial \Psi_l^{zi}}{\partial g_d^z} de^i \right\} = -\frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \left\{ \sum_{\forall i,-i} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_d^z} de^i \right\} \quad (25)$$

From the definition of the covariance between A and B, $\sigma(A,B) = E[AB] - E[A]E[B]$. Re-define $A = \left\{ \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}}, \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \right\}$ and $B = \left\{ \frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_d^z}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right\}$ for $l = 1,2$ to find:

$$\int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \frac{\partial \Psi_l^{zi}}{\partial g_d^z} de^i =$$

$$\sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right) + \left\{ \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \right\} \left\{ \int_{\forall\, e^i} h^i(e^i) \frac{\partial \Psi_l^{zi}}{\partial g_d^z} de^i \right\} \quad (26)$$

Define $\frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_d^z} = \tilde{v}_g^{zi} = \frac{\partial \tilde{\mu}^{zi}}{\partial G_d^z} - \frac{\partial \tilde{\mu}^{zi}}{\partial \tilde{x}^{zi}} \left(\frac{1}{N} \right)$ and $\frac{\partial \Psi_l^{zi}}{\partial g_d^z} = v_g^{zi} = \frac{\partial \mu^{zi}}{\partial G^z} - \frac{\partial \mu^{zi}}{\partial x^{zi}} \left(\frac{1}{N} \right)$ $\forall l$ and use (26) into (25) to show

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

$$-\sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right) - \frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \left\{ \sum_{\forall i,-i} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \tilde{v}_g^{zi} de^i \right\} \quad (27)$$

Define $\alpha^{zi} \in (0,1): \sum_{\forall i,-i} \alpha^{zi} = 1$where

$$\alpha^{zi} = \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \Big/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \quad (28)$$

And

$$\gamma^z = \frac{\partial \Pi_d^z}{\partial \tilde{\rho}_0^z} \Big/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \quad (29)$$

And

$$\sigma_\omega^z = \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_d^z} \right) \Big/ \sum_{\forall i,-i} \sum_{l=\{1,2\}} \frac{\partial \Pi_d^z}{\partial \rho_l^z} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \quad (30)$$

Use (28), (29), and (30) into (27) to obtain the expression in (20)[20]

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

$$= -\gamma^z \left(\sum\nolimits_{\forall\, i,-i} \int\limits_{\forall\, \tilde{e}^i} \tilde{h}^i\left(\tilde{e}^i\right) \frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}} \tilde{v}_g^{zi}(g_d^{*z}) d\tilde{e}^i \right) - \sigma_\omega^z \qquad (31)$$

■

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

$$\int\limits_{\forall\, e^i} h^i\left(e^i\right) v_{g_{dL}}^{zi}\left(g_{dL}^{*zi}\right) de^i = -\chi^{zi} \left(\int\limits_{\forall\, \tilde{e}^i} \tilde{h}^i\left(\tilde{e}^i\right) \frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}} \tilde{v}_{g_{dL}}^{zi}\left(g_{dL}^{*zi}\right) d\tilde{e}^i \right) - \sigma_\omega^{zi} \qquad (32)$$

Where

$$\chi^{zi} = \frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (33)$$

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

$$\sigma_\omega^{zi} = \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}}\, \sigma_l^{zi} \left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}} \right) \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int\limits_{\forall\, e^i} h^i\left(e^i\right) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (34)$$

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, $\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}$, and the change in the well being of voters from an increase in the provision of the local public good $\frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}$.

Proof

In a federation with party decentralization, the spending policy of party z in district i is $g_{dL}^{*zi} \in argmax\ \pi_{dL}^{zi}\left(\tilde{\rho}_0^{zi},\ \rho_1^{zi}, \rho_2^{zi}\right)$.[21] The first order condition of the party´s problem is

$$\frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \frac{\partial \tilde{\phi}_0^{zi}}{\partial g_{dL}^{zi}} + \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \frac{\partial \phi_l^{zi}}{\partial g_{dL}^{zi}} = 0 \qquad (35)$$

[20] Equivalent to expression (16) of Lemma 3 in the paper.

[21] Since the policies of candidates $j \in zi$ and $j' \in zi$ converge, we re-define $g_{dL}^{*jzi} = g_{dL}^{*j'zi} = g_{dL}^{*zi}$.

From definition, $\tilde{\phi}_0^{zi} = \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\, \tilde{F}_0^{zi}(\tilde{\Psi}_0^{zi}) d\tilde{e}^i$ and $\phi_l^{zi} = \int_{\forall\, e^i} h^i(e^i)\, F_l^{zi}(\Psi_l^{zi}) de^i$ for $l = \{1,2\}$. Obtain $\partial\tilde{\phi}_0^{zi}/\partial g_{dL}^{zi}$, $\partial\phi_l^{zi}/\partial g_{dL}^{zi}$ $\forall l$, and re-arrange terms to re-express (35) as follows

$$\sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}} de^i = -\frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_{dL}^{zi}} de^i \qquad (36)$$

From the definition of the covariance between A and B, $\sigma(A,B) = E[AB] - E[A]E[B]$. Re-define $A = \left\{\frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}}, \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}\right\}$ and $B = \left\{\frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_{dL}^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right\}$ for $l = 1{,}2$ to express the following

$$\int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}} de^i =$$

$$\sigma_l^{zi}\left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right) + \left\{\int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i\right\}\left\{\int_{\forall\, e^i} h^i(e^i) \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}} de^i\right\} \qquad (37)$$

Where $\sigma_l^{zi}\left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right)$ is the covariance between $\partial F_l^{zi}/\partial \Psi_l^{zi}$ and the voter´s marginal well being, $\partial \Psi_l^{zi}/\partial g_{dL}^{zi}$, due to a change in the provision of the local public good in district i. Use conditions $\frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_{dL}^{zi}} = \tilde{v}_{g_{dL}}^{zi} = \frac{\partial \tilde{\mu}^{zi}}{\partial G_{dL}^{zi}} - \frac{\partial \tilde{\mu}^{zi}}{\partial \tilde{x}^{zi}}\left(\frac{1}{N}\right)$ and $\frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}} = v_{g_{dL}}^{zi} = \frac{\partial \mu^{zi}}{\partial G_{dL}^{zi}} - \frac{\partial \mu^{zi}}{\partial x^{zi}}\left(\frac{1}{N}\right)$ $\forall l$ and substitute (37) into (36) to establish

$$\int_{\forall\, e^i} h^i(e^i)\, v_{g_{dL}}^{zi} de^i = \frac{-\frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \tilde{\Psi}_0^{zi}} \frac{\partial \tilde{\Psi}_0^{zi}}{\partial g_{dL}^{zi}} de^i - \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \sigma_l^{zi}}{\sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i} \qquad (38)$$

Define

$$\chi^{zi} = \frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}} \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (39)$$

And

$$\sigma_\omega^{zi} = \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}}\, \sigma_l^{zi}\left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right) \Bigg/ \sum_{l=\{1,2\}} \frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}} \int_{\forall\, e^i} h^i(e^i) \frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}} de^i \qquad (40)$$

To express condition (38) as follows

$$\int_{\forall e^i} h^i(e^i)\, v_{g_{dL}}^{zi}(g_{dL}^{*zi})de^i = -\chi^{zi}\left(\int_{\forall \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}}\tilde{v}_{g_{dL}}^{zi}(g_{dL}^{*zi})d\tilde{e}^i\right) - \sigma_\omega^{zi} \qquad (41)$$

■

Theorem 2 *The Decentralization Theorem holds in Majoritarian democracies with a decentralized party system, open primaries, and single member districts.*

Proof

For the economy with party decentralization and a system of local governments, the politically optimal policy for party z in each district is $g_{dL}^{*zi} > 0\ \forall i, -i$ satisfying:

$$\int_{\forall e^i} h^i(e^i)\, v_g^{zi}(g_{dL}^{*zi})de^i = -\chi^{zi}\left(\int_{\forall \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}}\tilde{v}_g^{zi}(g_{dL}^{*zi})d\tilde{e}^i\right) - \sigma_\omega^{zi} \qquad (42)$$

Under open primaries, the distribution of preferences over policy of voters voting in the primary is the same as the distribution of preferences of voters voting in the general election. Therefore, $h^i(e^i) = \tilde{h}^i(\tilde{e}^i)$ and $v_{g_{dL}}^{zi}(g_{dL}^{*zi}) = \tilde{v}_{g_{dL}}^{zi}(g_{dL}^{*zi})\ \forall z$. In this case, the parties' policies converge, therefore $\widetilde{\Psi}_0^{zi} = \widetilde{\Psi}_0^{-zi} = 0$, $\Psi_l^{zi} = \Psi_l^{-zi} = 0\ \forall z, \forall l$, $\frac{\partial \tilde{F}_0^{zi}(0)}{\partial \widetilde{\Psi}_0^{zi}} = c_0$ and $\frac{\partial F_l^z(0)}{\partial \Psi_l^{zi}} = c_1\ \forall z, \forall l$ where c_0, c_1 are positive constants. Hence $\sigma_l^{zi}\left(\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}, \frac{\partial \Psi_l^{zi}}{\partial g_{dL}^{zi}}\right) = 0\ \forall\ l\ \Rightarrow\ \sigma_\omega^{zi} = 0$ and $\int_{\forall\ \tilde{e}^i} \tilde{h}^i(\tilde{e}^i)\frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}}\tilde{v}_g^{zi}(g_{dL}^{*zi})d\tilde{e}^i = \int_{\forall\ e^i} h^i(e^i)\,\tilde{v}_{g_{dL}}^{zi}(g_{dL}^{*zi})de^i$. Moreover, $\tilde{\rho}_0^{zi} = 0$ and $\rho_l^{zi} = 0$ for $l = \{1,2\}$ leading to $\frac{\partial \pi_{dL}^{zi}(0)}{\partial \tilde{\rho}_0^{zi}} = \frac{\partial \pi_{dL}^{zi}(0)}{\partial \rho_l^{zi}} = c_2 \in \mathbb{R}_+\ \forall z$ and the following conditions hold:

$$\int_{\forall e^i} h^i(e^i)\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}de^i = c_1 \int_{\forall e^i} h^i(e^i)\,de^i = c_1 \in \mathbb{R}_+\ \forall z, \qquad (43)$$

Since $\int_{\forall\ e^i} h^i(e^i)\,de^i = 1$ measures the proportion of voters in the local election, and

$$\chi^{zi} = \frac{\partial \pi_{dL}^{zi}}{\partial \tilde{\rho}_0^{zi}}\Big/\sum_{l=\{1,2\}}\frac{\partial \pi_{dL}^{zi}}{\partial \rho_l^{zi}}\int_{\forall\ e^i} h^i(e^i)\frac{\partial F_l^{zi}}{\partial \Psi_l^{zi}}de^i = c_2/2c_2c_1 = constant \qquad (44)$$

Use (43) and (44) into (42) and re-arrange terms to show that the first order condition for $g_{dL}^{*zi}\ \forall i, -i$ satisfies the following condition:

$$\int_{\forall\, e^i} h^i(e^i)\, v_g^{zi}(g_{dL}^{*i})de^i = 0 \implies g_{dL}^{*zi} = g_{dL}^{*i} \;\; \forall\, z, \forall i \tag{45}$$

Now consider an economy with party decentralization in open primaries and a nationwide government. By (16) in Lemma 3 (which is equivalent to condition 20), the characterization of the first order condition for this economy with $g_d^{*z} > 0$ is given by:

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

$$= -\gamma^z \left(\sum\nolimits_{\forall\, i,-i} \int_{\forall\, \tilde{e}^i} \tilde{h}^i(\tilde{e}^i) \frac{\partial \tilde{F}_0^{zi}}{\partial \widetilde{\Psi}_0^{zi}} \tilde{v}_g^{zi}(g_d^{*z})d\tilde{e}^i \right) - \sigma_\omega^z \tag{46}$$

In this equilibrium, the parties' policies also converge then $g_d^{*z} = g_d^*$ and following similar steps as those shown above imply that the first order condition in (46) can be expressed as follows:

$$\int_{\forall\, e^i} h^i(e^i)\, v_g^{zi}(g_d^*)de^i = -\int_{\forall\, e^{-i}} h^{-i}(e^{-i})\, v_g^{z,-i}(g_d^*)de^{-i} \tag{47}$$

Moreover, recall from condition (2) that $\hat{g}^{*i} > 0\; \forall\, i$: $\hat{g}^{*i} \in \arg\max \int_{\forall\, e^i} h^i(e^i)\, v^i(e^i, G^i)de^i$ is is the global maximizer of the aggregate well being of residents of district i such that it is satisfied[22]

$$\int_{\forall\, e^i} h^i(e^i)\, v_g^i(\hat{g}^{*i})de^i = 0 \quad \forall i \;\; \hat{g}^{*i} > 0 \tag{48}$$

Therefore

$$\int_{\forall\, e^i} h^i(e^i)\, v^i(\hat{g}^{*i})de^i \geq \int_{\forall\, e^i} h^i(e^i)\, v^i(g^i)de^i \quad \forall\, g^i \neq \hat{g}^{*i} \tag{49}$$

Without loss of generality, the heterogeneity of preferences of voters for public goods means $g^{*i} \geq g^{*-i}$.[23] Moreover, conditions (45) and (2) imply that $g_{dL}^{*i} = \hat{g}^{*i}\;\; \forall i$, and the expressions in (47) and (48) imply $g_{dL}^{*-i} \leq g_d^* \leq g_{dL}^{*i}$. These outcomes and condition (49) means

[22] See that condition (45) is equivalent to condition (2).

[23] The heterogeneity of preferences means that, in general, $g^{*i} \geq g^{*-i}$ or $g^{*i} \leq g^{*-i}$. For the purpose of our analysis, and without loss of generality, we assume $g^{*i} \geq g^{*-i}$.

$$\int_{\forall e^i} h^i(e^i)\, v^i(g_{dL}^{*i})de^i + \int_{\forall e^{-i}} h^{-i}(e^{-i})\, v^{-i}(g_{dL}^{*-i})de^{-i} \geq \sum\nolimits_{\forall i,-i} \int_{\forall e^i} h^i(e^i)\, v^i(g_d^*)de^i \quad (50)$$

■

Theorem 3 *The Decentralization Theorem does not hold in Majoritarian democracies with single member districts and decentralized party systems with closed primaries.*

Proof

A party z seeking to form a central government in politically decentralized regimes with closed primaries selects $g_d^{*z} \in argmax\, \Pi_d^z(\tilde{\rho}_0^z,\, \rho_1^z, \rho_2^z)$ subject to $g_d^{*zi} = g_d^{*z,-i} = g_d^{*z}$. By proposition 2, $\hat{g}^{*i} \in \arg\max \int_{\forall e^i} h^i(e^i)\, v^i(e^i, G^i)de^i \;\; \forall i$ where $\hat{g}^{*i}$ is the policy that maximizes the nationwide surplus from the fiscal exchange associated with a local public good in district i. Condition (20) of Lemma 3 and condition (2) from proposition 2 imply that, in general, $g_d^{*z} \neq \hat{g}^{*i}\; \forall i$. Similarly, in a system of local governments with party decentralization and closed primaries, party z selects $g_{dL}^{*zi} \in argmax\; \pi_{dL}^{zi}\; \forall i$. Condition (32) from Lemma 3 and condition (2) from proposition 2 show that, in general, $g_{dL}^{*zi} \neq \hat{g}^{*i}\; \forall i$. As a result, in general, the nationwide aggregate wellbeing of voters satisfies the following

$$\int_{\forall e^i} h^i(e^i)\, v^i(g_{dL}^{*zi})de^i + \int_{\forall e^{-i}} h^{-i}(e^{-i})\, v^{-i}(g_{dL}^{*z,-i})de^{-i} \underset{>}{\overset{<}{\lessgtr}} \sum\nolimits_{\forall i,-i} \int_{\forall e^i} h^i(e^i)\, v^i(g_d^{*z})de^i \quad (51)$$

■

Table 6: Results of the Primary Models with Fiscal Decentralization Controls
(Error Correction Models with random effects, robust standard errors, and regional dummies)

Variable	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20	Model 21
	Y= Diff. Primary Completion Rate (N=348, 48 countries)	Y= Diff. Primary School Teachers (N=391, 50 countries)	Y= Diff. Preprimary Enrollment Rate (N=653, 57 countries)	Y= Diff. Primary Enrollment Rate (N=669, 57 countries)	Y= Diff. Secondary Enrollment Rate (N=660, 54 countries)	Y= Diff. Tertiary Enrollment Rate (N=618, 54 countries)	Y= Diff. Children Out of School (N=406, 42 countries)
Lagged Dependent Variable	-.192** (.083)	-.031** (.012)	-.049*** (.013)	-.076*** (.022)	-.024* (.013)	.015 (.011)	-.151*** (.036)
Democratic Decentralization, Party Centralization (Lagged)	.522 (.595)	.002 (.010)	.763** (.304)	.151 (.194)	.352 (.338)	.449* (.257)	-.104 (.082)
Municipal Plurality (Lagged)	-.528 (.427)	-.001 (.012)	-1.17** (.472)	-.326* (.194)	-.578* (.343)	-.455 (.341)	.107 (.081)
Municipal Directly Elected Executive (Lagged)	-.436 (.510)	-.009 (.011)	.049 (.346)	.282 (.196)	.023 (.311)	.111 (.319)	-.210** (.087)
Fertility (Differenced)	-2.99 (4.58)	.017 (.100)	2.91 (3.23)	.319 (1.74)	-.097 (4.97)	-.801 (1.62)	-.011 (.696)
Fertility (Lagged)	-.335 (.506)	-.008 (.012)	-1.00** (.422)	.038 (.236)	-.171 (.298)	-.392* (.220)	.003 (.109)
Logged GDP per capita (Differenced)	1.71 (4.46)	.018 (.080)	5.19 (3.77)	6.43 (5.03)	1.42 (3.77)	6.36** (2.91)	.459 (1.01)
Logged GDP per capita (Lagged)	.583 (.555)	.006 (.011)	.626 (.473)	-.428* (.227)	.588 (.394)	.529** (.253)	-.071 (.093)
Logged Population Density (Differenced)	-64.6 (40.2)	-.518 (.788)	32.8 (35.4)	5.41 (22.2)	15.1 (28.4)	-28.2 (21.1)	-.364 (9.48)
Logged Population Density (Lagged)	-.086 (.278)	.011* (.006)	-.049 (.197)	-.067 (.078)	-.098 (.119)	-.237** (.110)	-.026 (.033)
Urban Population (Differenced)	-1.52* (.887)	.002 (.026)	-1.06 (.663)	.206 (.542)	.193 (.595)	-.493 (.310)	.073 (.135)
Urban Population (Lagged)	.002 (.035)	.0003 (.0007)	-.003 (.019)	.021** (.008)	.020 (.020)	-.002 (.010)	-.005 (.004)
Expenditure Decentralization (Differenced)	-3.55 (10.2)	-.243 (.287)	3.42 (11.2)	-12.2* (6.56)	-22.8** (11.2)	-11.5* (6.60)	-.820 (1.68)
Expenditure Decentralization (Lagged)	-18.8 (14.7)	.015 (.183)	4.25 (3.80)	3.19 (2.50)	-9.87** (4.09)	-3.21 (2.90)	.212 (.601)
Revenue Decentralization (Differenced)	12.5 (14.2)	.292 (.287)	-10.3 (11.2)	13.6 (8.39)	18.2* (9.74)	14.1** (6.93)	-4.03* (2.25)
Revenue Decentralization (Lagged)	18.4 (15.4)	.057 (.184)	-7.90* (4.05)	-3.50 (2.50)	10.8** (4.57)	1.17 (3.13)	-.116 (.634)
R^2	.145	.041	.091	.051	.050	.145	.155

***p<.01, **p<.05, *p<.10. All tests are 2-tailed. Robust standard errors are in parenthesis