Universidad del Rosario
Facultad de Economía

1960 - 2010
50 Años
Facultad de Economía

PARTIAL DECENTRALIZATION AS A SAFEGUARD AGAINST FAVORITISM

Darwin Cortés

SERIE DOCUMENTOS DE TRABAJO

No. 89

Septiembre 2010

Partial Decentralization as a Safeguard against Favoritism

Darwin Cortés[*]

September 28, 2010

Abstract

In this paper I investigate the optimal level of decentralization of tasks for the provision of a local public good. I enrich the well-known trade-off between internalization of spillovers (that favors centralization) and accountability (that favors decentralization) by considering that public goods are produced through multiple tasks. This adds an additional institutional setting, partial decentralization, to the classical choice between full decentralization and full centralization. The main results are that partial decentralization is optimal when both the variance of exogenous shocks to electorate's utility is large and the electorate expects high performance from politicians. I also show that the optimal institutional setting depends on the degree of substitutability/complementarity between tasks. In particular, I show that a large degree of substitutability between tasks makes favoritism more likely, which increases the desirability of partial decentralization as a safeguard against favoritism.

JEL classification: D62, D72, D82, H10
KEYWORDS: Accountability, Multitask, Internalization of spillovers, Substitution.

[*]Universidad del Rosario, Bogotá, Colombia. E-mail: darwin.cortes@urosario.edu.co. I am grateful to Giorgio Brunello, Georges Casamatta, Guido Friebel, Gerhard Glomm and Darío Maldonado for their useful comments. I also thank participants at PET05, LACEA07 and Rosario seminar for their insights to previous versions. I gratefully acknowledge financial support from the Colombian Central Bank and the European Commission (EEEEPE - Marie Curie Training Network - University of Padua), as well as the hospitality of the Economics Department of the University of Padua. The findings, recommendations, interpretations and conclusions expressed in this paper are those of the authors and not necessarily reflect the view of the Department of Economics of the Universidad del Rosario.

1 Introduction

The provision of public goods entails different tasks. For instance, the provision of basic education involves, among other tasks, personnel management, organization of instruction, planning and resources allocation. When one looks at how these tasks are assigned to different levels of government across countries, it rarely follows a pattern where all tasks are assigned to either the local government or the central government. Some countries like Turkey and Malaysia assign almost all tasks to the central government. Other countries like Sweden and Hungary assign almost all tasks to schools. Most countries assign tasks to different tiers of government and the school (OECD, 1998).

What can explain the cross-country variation in assigning tasks to different levels of government? The paper provides an answer to this question based on the multitask nature of the provision of public goods and the low-powered incentives that citizens may provide to politicians through re-election. Assuming that local governments are more accountable while the central government can internalize spillovers, I show that Partial Decentralization (a government with two tiers) is optimal when the central government under full centralization is likely to favor a region. The rationale is as follows: the optimal level of decentralization depends on the risk born by politicians of being fired from office. When the risk is low, it is better to assign tasks to local governments. In this case, ex-post citizens' welfare is almost *determined* by politicians effort, such that the central government would be able to maximize chances of re-election with reduced effort (probably by coordinating policies across regions). When the risk is large, it is better to assign tasks to the central government, since internalization of externalities increases benefits from effort. However, if the risk is very large, the central government will have incentives to favor a region (or regions) to keep the office. Partial Decentralization cuts down these incentives. The cost of no favoritism is the creation of free-riding among the levels of government. When the risk of no re-election is very large, the benefit from no favoritism is larger than the cost of free-riding. In addition, I show that when the degree of substitutability of tasks is large, which makes politicians' effort more costly, the desirability of Partial Decentralization increases as a safeguard against favoritism.

As in Seabright (1996) and a subsequent body of literature, the basic trade-off in this paper is between accountability and internalization of spillovers. On the one hand, local governments are more accountable than the central government in the sense that citizens from a region may directly re-elect or oust their local politician. However, local governments do not internalize spillovers. On the other hand, central government may internalize externalities but is less accountable than the local governments. Seabright (1996) shows that centralization has an effect on accountability, the *reduced pivot probability effect*: A loss of accountability due to the reduced probability that the voters of a region have to determine the result of election. Thus, central government may be re-elected even if some regions want to fire it.

I formalize these ideas in a simple political agency model[1] with multitask. In the

[1]See Besley (2005) or Persson and Tabellini (2000) for a comprehensive presentation of political agency models.

model citizens live in one of two regions. Two tasks are inputs for the provision of one local public good. The public good provided in one region has positive spillovers on the other. Tasks may be assigned to two levels of government: the central government and the local governments. Effort in tasks is costly to politicians but generates utility to citizens. Not only does citizens' utility depend on politicians' effort but it also depends on a regional-specific shock.

Re-election is the mechanism available to citizens to provide incentives to politicians[2]. With this mechanism citizens cannot sign complete contracts based on verifiable information on politicians' performance. Instead, citizens establish a utility threshold that must be achieved in order for citizens to re-elect the incumbent politician. To capture the *reduced pivot probability effect* of centralization in the model, re-election rules are defined as follows: a local government is re-elected directly by citizens from the region. Central government is re-elected when at least one region wants to re-elect it. It means that the national politician may be re-elected even when one of the regions wants to fire it. This rule solves ties in favor of the incumbent politician and allows to capture the reduced accountability of the central government with a model with two regions.

I compare the social welfare provided by three institutional settings. Either local governments are in charge of both tasks in their own region (full decentralization), or the central government carries out both tasks in both regions (full centralization), or each tier of government carries out one task in each region (partial decentralization). The question is: What circumstances (if any) make partial decentralization to provide the largest welfare?

Since the central government is less accountable, it may favor one region. In single-task models, favoritism increases the case for full decentralization. The main contribution of this paper is to show that partial decentralization may do better for social welfare than full decentralization when *regional favoritism* is likely.

More precisely, I define three strategies that the central government can follow. The first one, as usual, defines the symmetric solution. Since regions are equal ex-ante , the central government may exert the same effort in both regions. The other two strategies define two types of regional favoritism. In the first type, the central government exerts effort in a single region. This effort is larger than the one in the symmetric solution because the central government uses externalities to decrease chances of no re-election. In the second type, the central government will coordinate effort in both regions, with the aim of being re-elected in the favored region. Besides the difference in the equilibrium efforts, these types of favoritism differ in how the central government assumes accountability. In the first type the central government follows the rule of re-election defined above. In the second type, the central government takes advantage of the reduced pivotal effect to bet on being re-elected by a single region with no care about the other region. In this simple model, it means that the central government behaves as if it was the local government of the favored region. For this reason, I will call the first type of favoritism, *weak favoritism* and the second one, *strong favoritism.*

[2] This paper models what the literature on voting calls *retrospective* voting, that is, the idea that voters assess incumbents actions while in office.

In terms of the model, the risk of not being re-elected depends on the variance of the exogenous shock that affects citizens' utility and the utility threshold used in re-election. This threshold may be interpreted as an exogenous citizens' expectation of the politicians' performance.

I show that partial decentralization does better than the other regimes, when both the variance of exogenous shock and the utility threshold are large. In principle, it would be better having a government that internalizes spillovers than having a government more accountable. However, the best strategy for the central government under full centralization is to *strongly* favor one region. With partial decentralization, each tier of government "specializes" in one task. Since local governments are more accountable they exert positive effort, which reduces the risk of no re-election (and the incentives for strong favoritism) of the central government. The cost of no favoritism with partial decentralization is that the central government free-rides on local governments' effort. At the equilibrium, the central government exerts a lower effort (due to free-riding) in both regions. This result extends to the case in which the central government chooses to *weakly* favor a region when citizens are risk averse to favoritism.

Furthermore, I investigate how task substitutability (complementarity) affects the previous results. This issue is at the core of the mainstream multitask problem with complete contracts (Holmstrom and Milgrom (1991)). I introduce task substitution in the cost function of politicians. Tasks are substitutes (complements) if an increase of effort in one task makes effort in the other tasks more (less) costly. I show that full centralization is better than full decentralization to deal with task substitutability. However, increasing task substitution also make the central government more prone to strong favoritism, which increases the case for partial decentralization.

The approach adopted in this paper steps on the seminal papers by Barro (1973) and Ferejohn (1986). These authors propose a principal - agent relationship between politicians and citizens, focusing on moral hazard problems in government. Seabright (1996) uses the main features of Ferejohn's model to analyze the appropriate level of decentralization of a local policy variable. Besides the *reduced pivot probability effect*, the author stresses a *rent-scale effect* of centralization that gathers the idea that central government may have more benefits from reelection than local governments. Hindriks and Lockwood (2004) add an adverse selection problem in a fiscal decentralization model, showing that decentralizing fiscal policy may increase the probability of firing bad incumbents. This paper abstracts from adverse selection issues. Selection effects of re-election would provide interesting insights but a pure moral hazard setting illustrates the point on partial decentralization and favoritism. My model departs from Seabright's in the issues arising from the multitask nature. In particular, multitask allows us to analyze a two-tier institutional setting - partial decentralization. This regime is not possible with Seabright's nor Hindriks and Lockwood's models.

To the best of my knowledge, there are two papers that integrate multitask to political agency models. Alesina and Tabellini (2004) use differences in incentives between politicians and bureaucrats to derive normative conclusions on whether tasks should be allocated to one or another. Padro i Miquel (2004) present a one-politician-one citizen model with multitask issues. They show that multitask has an adverse effect on the politician's equilibrium effort, in particular to implement interior solutions. This result resembles my result on weak favoritism. However, in their case it is due to non-concavity issues and in mine due to the reduced accountability of central government.

The paper also relates to other political economy models of fiscal decentralization. Lockwood (2002) and Besley and Coate (2003) present models of legislatures in which the central government is conceived as a central instance with regional delegates voting over agendas of regional projects. Without assuming spending uniformity there still is a trade-off between centralization and decentralization because centralization may be inefficient to select projects. These papers do not consider partial decentralization. More closely related, **?** shows that a two-tier legislature may alleviate the inefficiencies of the central legislature even if there are no externalities provided that projects of national interest come along with projects of local interest. Local legislatures may undertake projects discarded by the national legislature that have regional benefits.

In addition, Bordignon, Colombo, and Galmarini (2003) show that partial decentralization may do better than both full decentralization and full centralization in presence of lobbying. In their model two regional firms may lobby either to increase the provision of a regional public good or to monopolize a regional market. Assigning the regulatory policy (the number of firms in each market) to the central government and the provision of the public good to the local governments may reduce the effectiveness of lobbying for the markets.

Finally, in the context of complete contracts, **?** proposes a model of adverse selection to analyze regional favoritism. The national politician has more information about the citizens preferences of the largest region. Decentralization arises as a mechanism that may alleviate the downward distortion of production in the minority region.

The rest of the paper is organized as follows: the next section outlines the model. The third section presents a benchmark in which tasks are separable and there is no favoritism. The following section introduces favoritism and shows how it affects the welfare provided by the institutional settings at study. It also discusses the case in which citizens are risk averse. The fifth section deals with the analysis of non separable tasks. The last section concludes.

2 The Model

There are two regions $i = 1, 2$. In each region, two tasks should be done to provide a local public service. Three institutional settings are studied. With *full decentralization* one politician in each region (the local governments) is in charge of the respective tasks for his/her region. With *full centralization* one politician (the central government) is in charge of all tasks. Finally, with *partial decentralization* there are three politicians: the central government, who is in charge of one task in each region, and the local governments, who are in charge of the other task in their own regions.

In each region, citizens are identical and the social welfare depends on the local public service delivered in both regions. Moreover, citizens' welfare is also affected by a regional shock, y_i. Citizens' welfare from region i is given by

$$U_i(x_i, x_j) + y_i = \frac{x_i + \varepsilon x_j}{1 + \varepsilon} + y_i \tag{1}$$

where ε is positive and smaller than 1 (I assume positive externalities) and $i, j = 1, 2$, $i \neq j$ denote the regions. In line with Besley (2005), this assumption means that the citizens do have a common interest in achieving some outcome, such as, increasing the quality of public education, increasing the efficiency of public spending, improving public health, etc. Moreover, the analysis presupposes the existence of a well-functioning legal system[3].

The public service in region i, x_i, is provided carrying out tasks a_i and b_i. The technology is linear and is expressed by $x_i(a_i, b_i) = a_i + b_i$. As it was mentioned above, tasks a and b are delegated either both to the central government, both to local governments or one to each level of government. Performing tasks is costly for politicians since they have to exert effort on them. With *full centralization* the cost function of the central government is

$$C(a, b) = \sum_{i=1}^{2} \left[\frac{a_i^2}{2} + \frac{b_i^2}{2} + k^C a_i b_i \right] \tag{2}$$

where $a = (a_1, a_2)$ and $b = (b_1, b_2)$.

Similarly, with *full decentralization* the cost function of the local governments is

$$C_i(a_i, b_i) = \frac{a_i^2}{2} + \frac{b_i^2}{2} + k^L a_i b_i, \text{ for } i = 1, 2. \tag{3}$$

Finally, when tasks are split (*partial decentralization*), each local government, in charge of task a, has cost

$$C_i(a_i) = \frac{a_i^2}{2} + k^{PD} a_i b_i, \text{ for } i = 1, 2 \tag{4}$$

and the central government, in charge of task b, has cost

$$C(b) = \sum_{i=1}^{2} \left[\frac{b_i^2}{2} + k^{PD} a_i b_i \right]. \tag{5}$$

Notice that if $k^L = k^C = k^{PD}$, the total cost borne by the politicians is the same across the institutional forms when they exert equal efforts. In this case, the total cost of local governments under full decentralization (which is twice the cost in equation (3)) is equal to the cost borne by the central government under full centralization (equation (2)), and equal to the total cost of politicians under partial decentralization (sum of equations (4) and (5)). In other words, there are no economies of scale. Moreover, notice that since cost is quadratic, the marginal cost is increasing with effort.

[3]For instance, politicians cannot bribe or intimidate citizens and there are legal sanctions for using offices corruptly.

I also assume that k^L, k^C and k^{PD} belong to the interval $(-1,1)$ so that equations (2), (3), (4) and (5) are strictly convex. Following the multitask agency literature, I will say that efforts are *substitutes* when additional effort on one task makes more costly the effort on the other task, i.e., *task substitution* increases the marginal cost of effort. Conversely, efforts are *complements* when additional effort in one task reduces the marginal cost of effort in the other task. Formally, efforts are substitutes (complements) when the cross-partial derivatives of $C_i(.)$ and $C(.)$ are positive (negative), that is, when the *coefficients of effort substitution*, k^L, k^C and k^{PD} are positive (negative).

Citizens provide incentives for politicians to exert effort through re-election. When the national politician is re-elected, earns W; otherwise, earns zero. When the local politicians are re-elected they earn pW; otherwise, they earn zero. Coefficient p is a proportional coefficient, $p \in (0,1]$, that allows for the national politician to have a larger benefit from reelection than the local politicians. This is congruent with the intuition that national politicians may have more power than local politicians.

Re-election rules are given. The local government from region i is re-elected if the realized welfare[4], $U_i(x_i, x_j) + \widetilde{y}_i$, is higher than or equal to an exogenous level, T. The exogenous level T can be interpreted as the welfare that is expected from a rival political party or as a minimum level of welfare that citizens expect from politicians' actions. For the sake of simplicity, T is assumed to be the same in both regions and for both tiers of government.

The central government is re-elected when one of the two regions obtains a realized welfare higher than or equal to T. This re-election rule solves ties in favor of the incumbent politician. Along with the assumption on the number of regions, it allows the model to be tractable while keeping interesting features. In particular, this re-election rule captures the idea that the central government is less accountable than the local governments. With the central government the citizens in either region have a lower probability to determine the outcome of the election. The literature has called this effect the *reduced pivot probability effect* of centralization. The local governments are more accountable than the central government, since citizens can fire the local government when the realized welfare is smaller than T. Instead, the re-election of the central government depends also on the realized welfare in the other region. It can happen that a region wants but cannot fire the central government because the other region wants to re-elect him.

I further assume that the regional shocks, y_i, follow an $i.i.d$ uniform distribution with support $\left[-\frac{Y}{2}, \frac{Y}{2}\right]$. Let s_i be the event that region i is satisfied with its local government, i.e., $U_i(x_i, x_j) + y_i \geq T$. Let S^1 be the event that at least one region is satisfied with the central government. The probabilities of the events s_i and S^1 boil down to

$$pr(s_i) = \frac{\frac{Y}{2} - T + U_i(x)}{Y}$$

and

$$pr\left(S^1\right) = \frac{Y^2 - \left(\frac{Y}{2} + T - U_1(x)\right)\left(\frac{Y}{2} + T - U_2(x)\right)}{Y^2}$$

[4] $\widetilde{y}_i$ denotes a realization of y_i.

where subindexes 1, 2 in the last equation denote the regions.

The timing is as follows: The economy has a form of government; either full decentralization, full centralization or partial decentralization. In stage one, government(s) choose(s) a and b. Efforts are not observable to citizens. Afterwards, the regional shocks, y_i, are realized. Realizations are not observable to citizens neither. In stage two, the population welfare is realized and citizens choose whether to re-elect their governments.

Notice that citizens have a limited role in the game since they can just decide to re-elect or fire the incumbent politicians. I will characterize the equilibrium effort of politicians in each institutional setting, in order to compare the expected social welfare generated.

3 Benchmark: Separable tasks and no favoritism

As usual in this type of games, I solve backwards. In Stage 2 citizens realize their welfare, compare it with T and vote. If the actual welfare is larger than T for both regions, citizens re-elect the incumbent politician(s) in all the institutional settings. If the actual welfare is lower than T for both regions, citizens fire the incumbent politician(s) in all the institutional settings. If the actual welfare is lower than T for one region but higher than T for the other, there are differences in the re-election outcomes across the forms of government. With full decentralization the local government of the 'non satisfied' region is fired, while the local government of the 'satisfied' region is re-elected. With full centralization, the central government is re-elected. With partial decentralization, both the local government of the 'satisfied' region and the central government are re-elected. But, the local government of the 'non satisfied' region is fired.

The benchmark is characterized by the assumption that $k^L = k^C = k^{PD} = 0$, that is, politician's effort in one task does not affect the marginal cost of exerting effort in the other. Since I want to stress the effect of regional favoritism on the optimal institutional setting, the benchmark is also characterized by the no favoritism solution. As we will see below the central government may implement three strategies in order to maximize the chances of being re-elected. In the remaining of this section I present the politicians' choices under the benchmark and compare the citizens' welfare under the three forms of government considered. The analysis of favoritism is left to the Section 4.

3.1 The behavior of politicians

In Stage 1, politicians choose effort a and b in order to maximize their chances to be re-elected. Under my assumptions, local governments exert effort in the assigned tasks whenever the expected gain from re-election is larger than zero. They exert no effort, otherwise. On the other hand, since the central government deals with policies in both localities and its accountability is reduced, there are three strategies that the national politician can follow. The first strategy is to exert effort in both regions, that is, the symmetric solution. The other two strategies are asymmetric and represent different types of regional favoritism. This section deals with the symmetric solution (the benchmark).

3.1.1 Full Centralization

When the central government does not favor any region, it chooses a and b to maximize

$$
\begin{aligned}
Wpr\left(S^1\right) - C(a,b) = & \\
& W\left[\frac{Y^2 - \left(\frac{Y}{2} + T - U_1(x)\right)\left(\frac{Y}{2} + T - U_2(x)\right)}{Y^2}\right] \\
& - \sum_{i=1}^{2}\left[\frac{a_i^2}{2} + \frac{b_i^2}{2}\right],
\end{aligned} \tag{6}
$$

where $a = (a_1, a_2)$, $b = (b_1, b_2)$ and $x = (x_1, x_2)$.

At the equilibrium the central government exerts the same effort in the four tasks,

$$
a_i^C = b_i^C = \frac{(2T + Y)\,W}{2\left(Y^2 + 2W\right)}, \text{ for } i = 1, 2. \tag{7}
$$

Notice that the equilibrium efforts do not depend on the externality coefficient, ε, since the central government manage to fully coordinate the tasks among regions. The equilibrium efforts depend on the exogenous welfare expectation, T, the variability of the exogenous shocks, Y, and the benefit from re-election W [5]. An increase in T increases the probability of the central government to be fired and thus increases equilibrium effort. Both the benefit of re-election and the variability of the shocks have two opposite effects due to the coordination of externalities. The total effects are shown in the following Lemma (Proof in the appendix),

Lemma 1 *At the equilibrium, an increase in the benefit of re-election, W, increases effort of the central government. An increase of the variability of the shocks, Y, increases effort if $Y < -2T + \sqrt{4T^2 + 2W}$ and reduces effort if $Y > -2T + \sqrt{4T^2 + 2W}$.*

The first result is the expected. The second result says that the regional shocks make the incumbent politician to face a risk of not being re-elected. When the variance of y_i is small, the politician reacts to an increase in the variance by increasing effort. In this way he reduces the risk of not being re-elected. Somehow, when the shocks' variability is small, the central government can determine or control the final result on social welfare. On the other hand, when the variance is large, the increase in the probability of re-election due to an additional unit of effort is small and, thus, an increase of effort following an increase of the variance of y_i is no longer optimal.

[5] Under my assumptions, the regional shocks have mean equal to zero and variance equal to $\frac{Y^2}{12}$.

3.1.2 Full Decentralization

With *full decentralization*, the local government of each region i chooses a_i and b_i to maximize

$$pWpr(s_i) - C_i(a_i, b_i) = pW\left[\frac{\frac{Y}{2} - T + U_i(x)}{Y}\right] - \frac{a_i^2}{2} - \frac{b_i^2}{2} \tag{8}$$

The first order conditions give immediately the equilibrium efforts,

$$a_i^L = b_i^L = \frac{pW}{(1+\varepsilon)Y}, \text{ for } i = 1, 2. \tag{9}$$

The benefit from re-election increases the equilibrium efforts. The variance of the shock reduces the probability of being re-elected provided by an additional unit of effort. Thus, the equilibrium efforts decrease when the variance of the regional shock increases. Moreover, the local governments do not coordinate their policies. They are able to reduce their own effort by $\frac{1}{1+\varepsilon}$ due to the external effects among regions.

3.1.3 Partial Decentralization

With *partial decentralization*, the local government of region i chooses a_i to maximize

$$pWpr(s_i) - C_i(a_i) = pW\left[\frac{\frac{Y}{2} - T + U_i(x)}{Y}\right] - \frac{a_i^2}{2}, \text{ for } i = 1, 2 \tag{10}$$

and the central government chooses b to maximize

$$Wpr(S^1) - C(b) = W\left[\frac{Y^2 - \left(\frac{Y}{2} + T - U_1(x)\right)\left(\frac{Y}{2} + T - U_2(x)\right)}{Y^2}\right] - \frac{b_1^2}{2} - \frac{b_2^2}{2} \tag{11}$$

Local governments maximize equation (10) when

$$a_1^{PD} = a_2^{PD} = \frac{pW}{(1+\varepsilon)Y} \tag{12}$$

and the central government maximizes equation (11) when

$$b_1^{PD} = b_2^{PD} = \frac{W}{W+Y^2}\left[T + \frac{Y}{2} - \frac{pW}{(1+\varepsilon)Y}\right]. \tag{13}$$

With partial decentralization each tier of government specializes in one task or policy. Politicians may free-ride on each other. At the equilibrium local governments free-ride on each other as they do under full decentralization. Besides, the central government free-rides on the local governments' efforts. But, local governments cannot free-ride on the central government. This is so because local governments are more accountable than the central government. That is, with partial decentralization there is a vertical externality between the central government and the local governments. The effect of this externality at the equilibrium is gathered by the negative term in equation (13), which corresponds to the local governments' effort weighted by the probability for the central government of being re-elected.

3.2 Institutional Comparison

As the previous section shows, politicians are subject to risk. They are able to deal with this risk in different ways across the institutional settings. Accountability and different externalities have an effect on the ability of politicians to deal with this risk and to generate welfare. In this section I establish some conditions on the size of the variability of the shock that make politicians to exert more effort in one or another institutional setting.

I compare the institutional settings using the expected social welfare generated in each of them. The social welfare is defined as the average of the expected welfare in each region, $\frac{1}{2}U_1 + \frac{1}{2}U_2 + E[y_i]$. Since $E[y_i] = 0$, the average welfare boils down to the first two terms of the previous expression.

3.2.1 Full Centralization Versus Full Decentralization

The expected social welfare with full centralization is

$$\frac{1}{2}U_1^C + \frac{1}{2}U_2^C = \frac{W(2T+Y)}{Y^2+2W} \tag{14}$$

where U_1^C and U_2^C are the social welfare of each region, when the central government exerts efforts given in equation (7).

On the other hand, the expected social welfare with full decentralization is

$$\frac{1}{2}U_1^L + \frac{1}{2}U_2^L = \frac{2pW}{(1+\varepsilon)Y} \tag{15}$$

where U_1^L and U_2^L indicate the social welfare of each region when the efforts are those given in equation (9).

The main trade-off is as follows: on the one hand, the central government may deal with risk coordinating policies across regions. On the other hand, local governments are more accountable. The following lemma shows the results (Proof in the Appendix),

Lemma 2 *Let* $\underline{Y}_1 \equiv \frac{W}{T}$, $\underline{Y}_2 \equiv \frac{-T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{1+\varepsilon-2p}$, $\underline{Y}_3 \equiv \frac{T(1+\varepsilon)-\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ *and* $\overline{Y}_3 \equiv \frac{T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$.

With no favoritism, full decentralization provides larger social welfare than full centralization with the following conditions:

1. *If* $1+\varepsilon = 2p$*, when* $Y < \underline{Y}_1$
2. *If* $1+\varepsilon > 2p$*, when* $Y < \underline{Y}_2$
3. *If* $1+\varepsilon < 2p$*, when* $Y < \underline{Y}_3$ *or* $Y > \overline{Y}_3$

Otherwise, full centralization provides larger social welfare.

Three cases should be considered when comparing full decentralization and full centralization. These cases are defined according to the relative benefits of assigning tasks to local governments instead of to the central government. When $1+\varepsilon > 2p$ (Case 2), full decentralization is relatively inefficient to generate welfare. The benefits the local governments receive from re-election are relatively low with respect to the proportional diminution of efforts due to external effects. On the contrary, when $1+\varepsilon < 2p$ (Case 3), the benefits the local governments receive are relatively high with respect to the negative effect of externalities and, thus, full decentralization is relatively efficient to generate welfare. The case in which $1+\varepsilon = 2p$ (Case 1) is an intermediate case between the previous cases.

In all the cases, the social welfare with full decentralization is higher than the social welfare with full centralization when Y is lower a given threshold, that is, when the economy is relatively stable. The rationale behind this result is the following: when Y is very low the politicians are more able to determine the social welfare actually realized. In these conditions, the central government is more able to reduce efforts than the local governments for two reasons: first, the central government can coordinate efforts across regions, and second, it is less accountable than the local governments. Hence, when the variance of regional shocks is small, it is better to assign both tasks to the more accountable government.

In Case 3, the more favorable to local governments, full decentralization also provides the largest welfare when the economy is very unstable. When Y is very large, the increase in the probability to be re-elected due to an additional unit of effort is very low for both tiers of government. In this case, accountability outweighs externality coordination. Indeed, full centralization provides more welfare than full decentralization only if T is larger than $\sqrt{\frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}}$. Otherwise, real roots $\underline{Y}_3$ and $\overline{Y}_3$ do not exist and full decentralization always provides larger welfare.

3.2.2 Full Centralization Versus Partial Decentralization

With partial decentralization, the expected social welfare amounts to

$$\frac{1}{2}U_1^{PD} + \frac{1}{2}U_2^{PD} = \frac{W(2(1+\varepsilon)T + (1+\varepsilon+2p)Y)}{2(1+\varepsilon)(Y^2+W)} \tag{16}$$

where U_1^{PD} and U_2^{PD} indicate the social welfare of each locality when efforts are those in equations (12) and (13).

I summarize the results of comparing the expected social welfare given in equations (14) and (16) in the following Lemma (Proof in the Appendix),

Lemma 3 *With no favoritism, partial decentralization provides larger social welfare than full centralization under conditions (1.-3.) defined in Lemma 2. Otherwise, full centralization provides larger social welfare.*

Lemma 3 shows that Partial Decentralization brings similar advantages as those provided by Full Decentralization when it is compared with full centralization. Recall that with partial decentralization there is another trade-off between a vertical externality and specialization. The central government can free-ride on the local governments' efforts (vertical externality) but the task assigned to the local governments is more accountable (specialization). When the shock variability is smaller than the thresholds defined in Conditions $1. - 3.$, the gains of specialization outweighs the losses of vertical free-riding, such that total effort is larger with Partial Decentralization than with Full Centralization. When the local governments are efficient to generate welfare (Case 3), partial decentralization always dominates full centralization if the outside option is low enough.

3.2.3 Partial Decentralization Versus Full Decentralization

Finally, the comparison of full decentralization with partial decentralization is summarized in the following Lemma (Proof in the Appendix),

Lemma 4 *With no favoritism, full decentralization provides larger social welfare than partial decentralization under conditions (1.-3.) defined in Lemma 2. Otherwise, partial decentralization provides larger social welfare.*

This Lemma shows that when Partial Decentralization is compared with Full Decentralization, it resembles to Full Centralization. The argument behind is the same as in the previous comparison but operates in the other way around. When the shocks' variability is larger than the thresholds defined in Conditions $1. - 3.$, the local governments' effort is low. With full decentralization it means that the welfare provided is low. With Partial Decentralization it means that the vertical free-riding is small and that the central government must provide larger effort to be re-elected. Therefore, Partial decentralization generates larger welfare. As in the first comparison, when the benefit of re-election of local governments are relatively large in comparison with the horizontal free-riding (Case 3), Partial Decentralization generates more welfare than Full Decentralization only if T is large enough.

Proposition 1 summarizes the institutional comparison with no favoritism (Proof in the Appendix),

Proposition 1 *With no favoritism,*

1. *If* $1+\varepsilon = 2p$*, full decentralization provides the largest welfare when* $Y < \underline{Y}_1$*. Otherwise, full centralization does.*

2. *If* $1+\varepsilon > 2p$*, full decentralization provides the largest welfare when* $Y < \underline{Y}_2$*. Otherwise, full centralization does.*

3. *If* $1+\varepsilon < 2p$ *and* $T^2 > \frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}$*, full decentralization provides the largest welfare when* $Y < \underline{Y}_3$ *or* $Y > \overline{Y}_3$*. Otherwise, full centralization does.*

4. *If* $1+\varepsilon < 2p$ *and* $T^2 < \frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}$*, full decentralization always provides the largest welfare.*

where $\underline{Y}_1$*,* $\underline{Y}_2$*,* $\underline{Y}_3$ *and* $\overline{Y}_3$ *as defined in Lemma 2.*

Partial decentralization never provides larger welfare than the other settings.

With no favoritism, partial decentralization is always dominated by one of the other two systems. In the more favorable case to local governments (when $1+\varepsilon < 2p$) full decentralization always provides the largest welfare unless the outside option T is larger than the threshold $\sqrt{\frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}}$ and, thus, the central government has more incentive to exert effort.

4 The Case of Favoritism

Since the central government is re-elected when a single region has a realized welfare larger than T, favoring one region may bring a larger utility to the national politician. Favoritism makes citizens to face a risk, the risk of the national politician to exert less effort in their region. Under my assumptions, regions are completely identical to politicians and, thus, have equal chances to be favored. Besides, citizens are risk neutral (See Equation (1)).

Since I am interested to investigate under which conditions Partial Decentralization reports a larger welfare than the other institutional settings, the following analysis sticks to the conditions established in Proposition 1. From there, if the shock variability Y is smaller than the low-bar thresholds $\underline{Y}_1$, $\underline{Y}_2$ and $\underline{Y}_3$, partial decentralization will be dominated by full decentralization. This section focuses on what happens above those thresholds. I am interested to look at how favoritism affects the social welfare provided by Full Centralization and under which conditions it makes Partial Decentralization to provide a larger social welfare. In the rest of the section I define two types of favoritism and show how favoritism affects the welfare output of Full Centralization.

4.1 Weak Favoritism

Let me define weak favoritism as follows

Definition 1 (Weak favoritism) *The central government weakly favors region i if maximizes the program $Wpr\left(S^1\right) - C(a,b)$ (Equation (6)) over a_i and b_i and exerts no effort in the other region ($a_j = b_j = 0$).*

As in the previous section, the central government maximizes the probability of being re-elected by at least one region. But now, the central government restricts the space of election variables to the tasks or policies in one region. It exerts no effort in the other region. Both things together mean that the central government *cares* about the welfare in the non favored region only through the externalities it receives. The equilibrium efforts of the central government under Full Centralization with weak favoritism are

$$a_1^{Cwf} = b_1^{Cwf} = \frac{(1+\varepsilon)^2\,(2T+Y)W}{2((1+\varepsilon)^2 Y^2 + 4\varepsilon W)} \tag{17}$$

Comparing the equilibrium efforts with no favoritism and weak favoritism (equations (7) and (17)), it can be seen that the effort exerted in the favored region is larger than the effort exerted in each region with no favoritism.

With weak favoritism the welfare in the favored region and the non favored region are, respectively, equal to[6]

$$U_1^{Cwf} = \frac{(1+\varepsilon)\,(2T+Y)W}{(1+\varepsilon)^2\,Y^2 + 4\varepsilon W} \qquad and \qquad U_2^{Cwf} = \frac{\varepsilon\,(1+\varepsilon)\,(2T+Y)W}{(1+\varepsilon)^2\,Y^2 + 4\varepsilon W}. \tag{18}$$

The expected welfare of the non favored region is equal to the welfare of the favored region times the coefficient of externality, ε. The social welfare with weak favoritism is

$$\frac{1}{2}U_1^{Cwf} + \frac{1}{2}U_2^{Cwf} = \frac{(1+\varepsilon)^2(2T+Y)W}{2((1+\varepsilon)^2 Y^2 + 4\varepsilon W)} \tag{19}$$

The following Lemma shows the condition in which both the national politician chooses to weakly favor one region and the social welfare from weak favoritism is larger than the social welfare from no favoritism (Proof in the Appendix)

Lemma 5 *Let $\widetilde{Y} \equiv \sqrt{\frac{2(1-\varepsilon)^2 W}{(1+\varepsilon)^2}}$. Weak favoritism is better than no favoritism for both the central government's utility and the social welfare if $Y < \widetilde{Y}$.*

Lemma 5 shows that when citizens are risk neutral, weak favoritism is good for both the national politician and social welfare under the same conditions. It shows that under Full Centralization, both the central government weakly favors a region and the

[6] With no loss of generality, the favored region is Region 1 and the non favored region is region 2. The symmetric case gives the same results in the analysis that follows.

resulting social welfare is larger when the economy is relatively stable, i.e. when the shocks' variance is relatively small. Notice that the central government will take the risk to (weakly) favor a region in an economy more unstable when the benefit of re-election is larger (the threshold $\widetilde{Y}$ increases with the benefit of re-election). Notice as well that weak favoritism decreases with the size of externality, ε.

As I said above, I investigate whether favoritism makes room for Partial Decentralization to provide the largest welfare. Proposition 1 shows the conditions in which Full Centralization provides the largest welfare with no favoritism. How does the welfare provided by Full Centralization is affected by weak favoritism? Does it make room for Partial Decentralization to provide the largest welfare? Formally, this is possible only if $\widetilde{Y}$ is larger than the thresholds $\underline{Y}_1$, $\underline{Y}_2$ and $\underline{Y}_3$ in Proposition 1. The following Proposition shows the conditions in which this is so (Proof in the Appendix),

Proposition 2 *If the central government weakly favors one region, full centralization provides the largest welfare if both the shock variability and the exogenous outside option are large enough, as follows*

1. *if* $1+\varepsilon = 2p$, $Y > \underline{Y}_1$ *and* $T > \frac{W}{\widetilde{Y}}$.
2. *if* $1+\varepsilon > 2p$, $Y > \underline{Y}_2$ *and* $T > \frac{4pW-(1+\varepsilon-2p)\widetilde{Y}^2}{2(1+\varepsilon)\widetilde{Y}}$.
3. *if* $1+\varepsilon < 2p$, $Y > \underline{Y}_3$ *and* $T > \frac{4pW+(2p-(1+\varepsilon))\widetilde{Y}^2}{2(1+\varepsilon)\widetilde{Y}}$.

Partial decentralization never provides the largest welfare.

Full centralization with weak favoritism provides the largest welfare if the outside option is above certain threshold. If T is interpreted as an exogenous minimum level of welfare required by citizens, this Proposition says that the central government weakly favors a region when citizens are very demanding. Regarding Case 3 ($1+\varepsilon < 2p$), Proposition 1 shows that T must be larger than $\sqrt{\frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}}$ to make the central government to provide the largest welfare. Proposition 2 shows that if T is too large, the central government weakly favors a region.

Besides, weak favoritism does not make room for Partial Decentralization to provide the largest welfare. With weak favoritism the central government is still accountable to both regions in the sense that the central government maximizes the probability to be reelected by at least one region. However, the central government cuts down the cost by exerting effort in a single region and affects the other region's welfare through externalities. In Section 4.3 I make additional comments on how this result is affected by citizens' risk aversion.

4.2 Strong Favoritism

Let me now consider the second type of favoritism. Strong favoritism is defined as follows

Definition 2 (Strong favoritism) *The central government strongly favors region i if maximizes the program $Wpr(s_i) - C(a,b)$ over a_1, b_1, a_2 and b_2.*

Strong favoritism is characterized by the fact that the central government maximizes the probability to be re-elected by the favored region. Formally, the probability function has changed. The central government only cares about the favored region's welfare. It exerts effort in the non favored region to further increase the welfare in the favored region. With strong favoritism the equilibrium efforts in the favored and non favored regions tasks are respectively,

$$a_1^{Csf} = b_1^{Csf} = \frac{W}{(1+\varepsilon)Y} \quad \text{and} \quad a_2^{Csf} = b_2^{Csf} = \frac{\varepsilon W}{(1+\varepsilon)Y} \tag{20}$$

which bring a social welfare equal to

$$\frac{1}{2}U_1^{Csf} + \frac{1}{2}U_2^{Csf} = \frac{W}{Y} \tag{21}$$

With strong favoritism, the central government behaves similarly to local governments. It coordinates policies across regions to maximize re-election by the favored one. To show whether Partial Decentralization provides the largest welfare when the central government strongly favors a region, I will first present two Lemmas. In the first one I show the conditions in which the national politician prefers strong favoritism to no favoritism. In the second one, I show the conditions in which Partial decentralization provides larger welfare than Full centralization with strong favoritism.

Lemma 6 *Let $\widetilde{Y} \equiv \sqrt{\frac{2(1-\varepsilon)^2 W}{(1+\varepsilon)^2}}$, $f(Y) \equiv \frac{\sqrt{(1+\varepsilon)^2 Y^2((1+\varepsilon)^2 Y^4 + 8\varepsilon W Y^2 - 4(1-\varepsilon)^2 W^2)}}{2(1+\varepsilon)^2 Y^2}$, $\underline{T}(Y) \equiv \frac{W}{Y} - f(Y)$ and $\overline{T}(Y) \equiv \frac{W}{Y} + f(Y)$.*

The central government is better off strongly favoring one region than not favoring anyone if

1. *$Y \leq \widetilde{Y}$*
2. *$Y > \widetilde{Y}$ and $T \notin (\underline{T}(Y), \overline{T}(Y))$*

Comparing Lemma 6 with Lemma 5, strong favoritism allows the central government to favor a region in more unstable economies, that is, in economies in which the shock variability is larger than $\widetilde{Y}$. This is so, if the outside option T is either large or small enough. The thresholds $\underline{T}(Y)$ and $\overline{T}(Y)$ have two parts. From the proof of Lemma 6 in the Appendix, if $Y = \widetilde{Y}$, $\underline{T}(Y) = \overline{T}(Y)$ because $f(\widetilde{Y}) = 0$. For $Y > \widetilde{Y}$, $\overline{T}(Y)$ increases with Y and $\underline{T}(Y)$ decreases with Y since the derivative of $f(Y)$ with respect to Y is positive. The larger the shock variance is, the harder strongly favor a region is. Let me present the second Lemma (Proof in the Appendix),

Lemma 7 *Let* $\underline{Y}_2^P \equiv \frac{(1+\varepsilon)T-\sqrt{(1+\varepsilon)^2T^2-2(1+\varepsilon-2p)(1+\varepsilon)W}}{(1+\varepsilon-2p)}$,
$\overline{Y}_2^P \equiv \frac{(1+\varepsilon)T+\sqrt{(1+\varepsilon)^2T^2-2(1+\varepsilon-2p)(1+\varepsilon)W}}{(1+\varepsilon-2p)}$,
$\underline{Y}_3^P \equiv \frac{-(1+\varepsilon)T+\sqrt{(1+\varepsilon)^2T^2+2(2p-(1+\varepsilon))(1+\varepsilon)W}}{2p-(1+\varepsilon)}$ *and* $\underline{T}^P \equiv \sqrt{\frac{2(1+\varepsilon-2p)W}{(1+\varepsilon)}}$.

Partial decentralization provides larger welfare than full centralization with strong favoritism when the shock variability is relatively large, as follows,

1. *If* $2p = 1+\varepsilon$, *when* $Y > \frac{W}{T}$.
2. *If* $1+\varepsilon > 2p$, *when* $Y \in \left(\underline{Y}_2^P, \overline{Y}_2^P\right)$ *and* $T > \underline{T}^P$.
3. *If* $1+\varepsilon < 2p$, *when* $Y > \underline{Y}_3^P$.

Partial Decentralization provides larger welfare than Full Centralization with strong favoritism when the shock variance is large enough. To understand why, recall that with strong favoritism the central government behaves like a local government. It is only accountable to the favored region. Therefore, like the local government would do, the central government exerts a big effort when the economy is very stable. But equilibrium efforts decrease rapidly when the shock variance increases. Above the thresholds in Lemma 7, partial decentralization does better. Governments free-ride each other, but the losses in welfare due to free-riding are smaller than the losses in welfare due to strong favoritism.

To show that Partial Decentralization provides the largest welfare when the central government strongly favors one region, I restrict attention to the interval in which Y is larger than $\widetilde{Y}$. It is sufficient to show four conditions to hold at the same time. First, that under Full Centralization, the central government prefers strong favoritism over both no favoritism and weak favoritism. Second, that the social welfare provided with strong favoritism is lower than the social welfare provided with no favoritism. Third, that the social welfare provided with Partial Decentralization is larger than the one provided with Full Centralization. And fourth, that all this happens when Full Centralization with no favoritism provides the largest welfare according to conditions in Proposition 1. With these conditions, full centralization would provide the largest welfare when there is no favoritism. However, the central government maximizes chances of re-election strongly favoring one region. Social welfare is smaller with strong favoritism than with partial decentralization. The following Proposition shows that Partial Decentralization provides the largest welfare when both the shock variance and the outside option are large enough (Proof in the Appendix),

Proposition 3 *Partial decentralization provides the largest welfare when the central government chooses to strongly favor one region under the following conditions*

1. $1+\varepsilon = 2p$. *When* $Y > max\{\underline{Y}_1, \widetilde{Y}\}$ *and* $T > \overline{T}(Y)$.
2. $1+\varepsilon > 2p$. *When* $Y \in (\widetilde{Y}, \overline{Y}_2^P)$ *and* $T > max\left\{\underline{T}^P, \overline{T}(Y), \widehat{T}\right\}$.

3. $1+\varepsilon < 2p$. *when* $Y > max\{\underline{Y}_3, \widetilde{Y}\}$ *and* $T > \overline{T}(Y)$.

where $\widehat{T} \equiv \frac{(1+\varepsilon-2p)}{2(1+\varepsilon)}\widetilde{Y} + \frac{W}{\widetilde{Y}}$. *The other thresholds are defined in Lemmas 2, 6 and 7.*

Since T is very demanding, it is too costly for the central government to exert effort in both regions. Thus, it chooses to strongly favor one region. Partial decentralization decreases the incentive for favoritism, cutting down the cost the central government must afford to exert effort in both regions. Free-riding on the local governments is the cost to make the central government to favor no region.

Given that $Y > \widetilde{Y}$, with Full Centralization the central government strongly favors one region if T is either high or low enough according to Lemma 6. When $T < \underline{T}(Y)$, the social welfare provided with strong favoritism is larger than the one provided with no favoritism. But, when $T > \overline{T}(Y)$, the social welfare is smaller with strong favoritism than with no favoritism. This explains why Partial Decentralization does better than strong favoritism only if $T > \overline{T}(Y)$.

Besides, notice that Case 2 ($1+\varepsilon > 2p$) is more demanding for Partial decentralization than Case 3 ($1+\varepsilon < 2p$). In Case 2, Y must belong to a bounded interval and T must be larger than the maximum of three thresholds. None dominates the others. Indeed, $\overline{T}(Y)$ is increasing with Y and $\overline{T}(\widetilde{Y}) = \frac{W}{\widetilde{Y}}$. On the other hand, $\underline{T}^P$ and $\widehat{T}$ do not vary with Y but it is easy to see that when Y is low enough $\widehat{T} > \overline{T}(Y)$.

Case 2 is more demanding than the other cases because the *rent scale* effect has an offsetting effect on the *reduced pivot probability* effect. To see this, assume the externality coefficient, ε, is given. On the one hand, if p is equal to one, the benefit from re-election is equal for both tiers of government and the rent scale effect is zero. From my assumptions, when $p = 1$, $1+\varepsilon < 2p$. Decreasing p decreases the benefit from re-election for local governments vis-a-vis the central government. On the other hand, with Full Centralization the citizens in either region have a lower probability to determine the outcome of the election. This is the *reduced pivot probability* effect. It allows the central government to favor a region. Proposition 3 shows that the rent scale effect does not preclude the *reduced pivot probability* effect. However, when the rent scale effect is large enough ($1+\varepsilon > 2p$), the *reduced pivot probability* effect is more limited, reducing the case for Partial Decentralization.

The results of this Section for the case in which $1+\varepsilon < 2p$ are depicted in Figure 1. It shows which institutional setting provides the largest welfare in the space of the shock variability Y (x-axis) against the outside option T (y-axis). I have assumed that $\widetilde{Y}$ is larger than the low-bar thresholds in Proposition 1. When the economy is very stable ($Y < \underline{Y}_3$), Full Decentralization provides the largest welfare (Proposition 1). When the economy is relatively stable ($\underline{Y}_2 < Y < \widetilde{Y}$)and T is low enough, Full Centralization with no favoritism is the best for welfare. When T is large enough, the central government has the incentive to weakly favor one region and weak favoritism does better for welfare (Proposition 2). When the economy is more unstable ($Y > \widetilde{Y}$), the central government may have the incentive to strongly favor one region. This may happens if T is too high or too low. When T is too high, Partial Decentralization provides the largest welfare.

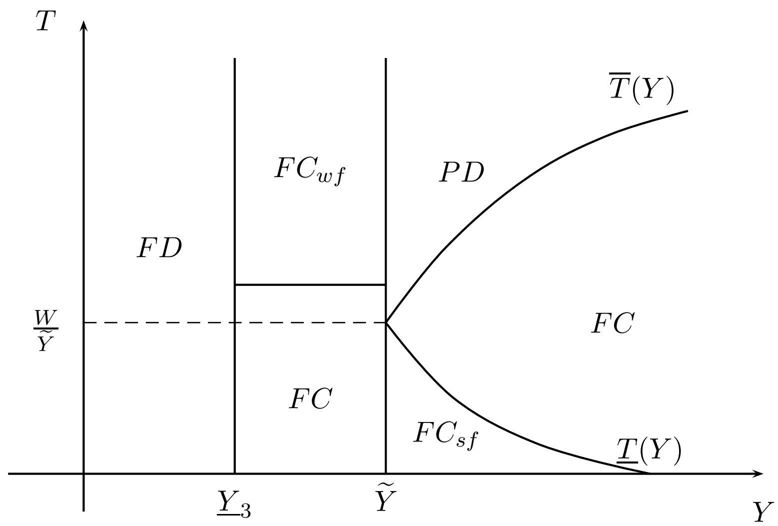

Figure 1: Separable Tasks, Case 3

4.3 Risk Aversion

In the previous analysis I have assumed that citizens are risk neutral. They are indifferent between the lottery of being favored or not and the average welfare of both states. If citizens are risk averse to favoritism, the social welfare with both types of favoritism will decrease and the institutional setting that provides the largest welfare may change.

To illustrate the point assume that citizens are infinitely risk averse. Hence, the social welfare is given by the lowest welfare provided by either region at the equilibrium. With full decentralization, partial decentralization and full centralization with no favoritism, the social welfare is the same as before. With weak favoritism the social welfare becomes equal to the social welfare in the non favored region, $\frac{\varepsilon(1+\varepsilon)(2T+Y)W}{(1+\varepsilon)^2Y^2+4\varepsilon W}$. The first part of Lemma 5 still holds, i.e., the central government is better off weakly favoring one region if $Y < \widetilde{Y}$. But the Social Welfare is no longer larger with weak favoritism than with no favoritism. Indeed, social welfare is larger with no favoritism (equation 14) than with weak favoritism for all Y. Under conditions in Proposition 2, Partial Decentralization would provide the largest welfare.

With strong favoritism the social welfare becomes equal to $\frac{\varepsilon W}{(1+\varepsilon)Y}$. The central government is still better off with strong favoritism according to conditions in Lemma 6. However, strong favoritism does not provide the largest welfare when $Y > \widetilde{Y}$ and $T < \underline{T}(Y)$. This increases the case for Partial decentralization to provide the largest welfare. Although I have considered an extreme case, it illustrates very well the idea that risk aversion to favoritism increases room for the case of Partial Decentralization.

5 Non Separable Tasks

In this section I turn to the case in which $k^L = k^C = k^{PD} = k$. I want to know how the results of the previous section are affected by task substitution. When tasks are substitutes (i.e. $k > 0$), tasks are more costly for politicians. An additional unit of effort in one task increases the marginal cost of effort in the other. On the contrary, when tasks are complements (i.e. $k < 0$), an additional unit of effort in one task decreases the marginal cost of effort in the other.

5.1 Full Decentralization and Full Centralization

The local governments choose efforts to maximize equation (8) with the cost function equal to $C_i(a_i, b_i) = \frac{a_i^2}{2} + \frac{b_i^2}{2} + ka_ib_i$. The equilibrium efforts are

$$a_i^L = b_i^L = \frac{pW}{(1+k)(1+\varepsilon)Y}. \tag{22}$$

The central government faces a cost function equal to $C(a,b) = \sum_{i=1}^{2}\left[\frac{a_i^2}{2} + \frac{b_i^2}{2} + ka_ib_i\right]$. With no favoritism the equilibrium efforts are

$$a_i^C = b_i^C = \frac{(2T+Y)W}{2((1+k)Y^2 + 2W)}, \text{ for } i = 1, 2. \tag{23}$$

With weak favoritism the equilibrium efforts in the favored region become $a_1^{Cwf} = b_1^{Cwf} = \frac{(1+\varepsilon)^2(2T+Y)W}{2((1+k)(1+\varepsilon)^2Y^2+4\varepsilon W)}$ and with strong favoritism the equilibrium efforts in the favored and non favored regions are $a_1^{Csf} = b_1^{Csf} = \frac{W}{(1+k)(1+\varepsilon)Y}$ and $a_2^{Csf} = b_2^{Csf} = \frac{\varepsilon W}{(1+k)(1+\varepsilon)Y}$, respectively.

Proposition 4 (Proof in the Appendix) *With full decentralization and full centralization, increasing (decreasing) task substitution always decreases (increases) the equilibrium effort of politicians.*

Since tasks are more costly for politicians when tasks are substitutes, the politicians reduce their equilibrium efforts compared to those with no substitution. In contrast, when tasks are complements, they are less costly, and politicians increase their equilibrium efforts compared to those with no substitution. With Full Decentralization, local governments reduce (increase) free-riding when the tasks are substitutes (complements). With both types of favoritism, the central government uses externalities to further reduce effort when tasks are substitutes and to further increase effort when tasks are complements.

5.2 Partial Decentralization

With Partial Decentralization the local government of region i faces a cost function equal to $C_i(a_i) = \frac{a_i^2}{2} + ka_ib_i$ and the Central Government has a cost function equal to $C(b) = \sum_{i=1}^{2}\left[\frac{b_i^2}{2} + ka_ib_i\right]$.

The optimal behavior of the local government from region i is given by the following reaction function,

$$a_i = \frac{pW}{(1+\varepsilon)Y} - kb_i \tag{24}$$

Tasks as strategic substitutes if $k > 0$ and strategic complements if $k < 0$. When tasks are substitutes, local governments free-ride on the central government's effort, reducing effort in k units for each unit of the central government's effort. When tasks are complements, local governments increase effort in k units for each unit of central government's effort.

Taking the symmetric solution ($a_1 = a_2$, $b_1 = b_2$), the optimal behavior of the central government is given by the following reaction function,

$$b_i = \frac{W(T + \frac{Y}{2})}{Y^2 + W} - \left[\frac{k + \frac{W}{Y^2}}{1 + \frac{W}{Y^2}}\right] a_i \tag{25}$$

Tasks are strategic substitutes for the central government even for some negative values of k. Tasks are strategic complements for the central government only when $k < -\frac{W}{Y^2}$. The equilibrium efforts are

$$a_i^{PD} = \frac{W(2p(W+Y^2) - (1+\varepsilon)kY(2T+Y))}{2(1+\varepsilon)(1-k)Y(W+(1+k)Y^2)} \tag{26}$$

and

$$b_i^{PD} = \frac{W((1+\varepsilon)Y(2T+Y) - 2p(W+kY^2))}{2(1+\varepsilon)(1-k)Y(W+(1+k)Y^2)} \tag{27}$$

Equations (26) and (27) show that the task substitution coefficient has two contrary effects on efforts. On the one hand, increasing task substitution reduces the equilibrium efforts because it makes efforts more costly at the margin. On the other hand, increasing task substitution increases the equilibrium efforts because each government increases efforts to compensate the reduction of the other government's effort.

Notice that the equilibrium effort from the Central Government will be negative if $T < \frac{pW}{(1+\varepsilon)Y} - \frac{2pk-(1+\varepsilon)}{2(1+\varepsilon)}Y$. This is not possible under the conditions on T in Proposition 3. Therefore, when the partial decentralization provides the largest welfare, the central government provides a positive effort.

5.3 Institutional Comparisons

Task substitution affects the social welfare provided by the different institutional settings. Comparing full decentralization with full centralization (Proof in the Appendix),

Proposition 5 *Increasing (decreasing) task substitution increases (decreases) the case for full centralization with no favoritism vis-a-vis full decentralization.*

When the central government favors no region, full centralization is better than full decentralization to deal with task substitution. This is so because the central government can coordinate policies across regions in order to reduce effort. Indeed, when the central government weakly favors one region, it is not clear any more if increasing task substitution increases weak favoritism. Recall that weak favoritism provides the largest welfare in the intervals defined between the low-bar thresholds ($\underline{Y}_1$, $\underline{Y}_2$, $\underline{Y}_3$) and $\widetilde{Y}$. The low-bar thresholds are decreasing in k (see Proof of Proposition 5). On the other hand, $\widetilde{Y}$ becomes equal to $\sqrt{\frac{2(1-\varepsilon)^2 W}{(1+\varepsilon)^2(1+k)}}$. Both thresholds are thus decreasing in k. It is not clear what is the net effect on the interval of shock variability. Increasing tasks substitution makes full centralization to be better than full decentralization in providing welfare in more stable economies. However, it also increases the cost of not coordinating externalities across regions, that is, the cost of weakly favoring a region.

The following Proposition shows the effect of task substitution on partial decentralization (Proof in the Appendix),

Proposition 6 *Increasing task substitution makes partial decentralization to provide the largest welfare in more stable economies. It also reduces the constraint on the exogenous outside option in those economies.*

The Proposition is depicted in Figure 2. The central government becomes able to strongly favor one region in more stable economies. This is so because it uses the externalities the favored region is recipient of to reduce the extra-cost brought about by task substitution. As before partial decentralization does better than strong favoritism according to conditions in Proposition 3. However, only is possible to say something with respect to the more stable economies of the interval (i.e. for Y near $\widetilde{Y}$) due to the constraints on T. Indeed, $\overline{T}(Y)$ is increasing in Y, so that increasing task substitution makes the minimum T more demanding for Y high enough.

6 Conclusions

This paper addresses the issue of the appropriate level of decentralization of a multi-task government. In the setting developed above, citizens provide incentives to politicians through re-election. The basic trade-off is between coordination of externalities across regions (which favors assigning tasks to the central government) and accountability (which favors assigning tasks to the local governments). In such a framework, I

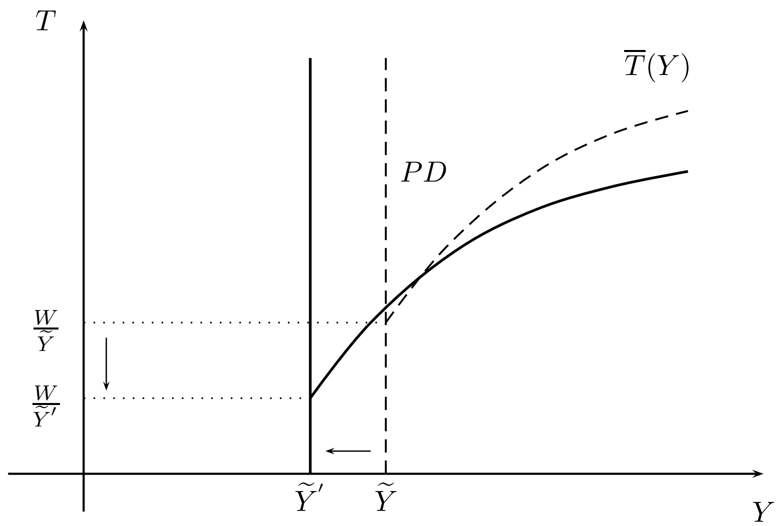

Figure 2: Partial Decentralization and increasing task substitution

compared three forms of government: full centralization, full decentralization and partial decentralization. I fully characterized which form of government provides the largest welfare depending on the stability of the economy, given by the variance of regional shocks, and the exogenous outside option that citizens require to re-elect politicians.

The paper shows that full decentralization provides the largest welfare in more stable economies and that full centralization does so in more unstable economies. Besides, the central government favors one region while being accountable to both regions (weak favoritism) if the economy is relatively stable and the exogenous outside option is high. The central government favors one region while being accountable to a single region (strong favoritism) if the economy is relatively unstable and the exogenous outside option is either high or low enough.

Moreover, I find that partial decentralization provides the largest social welfare when the central government strongly favors a region. Partial decentralization reduces the central government' incentives to shirk while taking advantage from the larger accountability of local governments. Central government's free-riding on local governments' efforts is the cost of precluding favoritism. The result extents to weak favoritism, at least in part, when citizens are risk averse to favoritism.

The paper also sheds light on how task substitution affects the provision of social welfare. The central government is better than local governments to deal with task substitution. More precisely, increasing task substitution makes the case for full centralization in more stable economies. In addition, increasing tasks substitution makes likely for the central government to strongly favor one region in economies relatively more stable, which increases the case for partial decentralization.

These results seem to be in line with the empirical findings in Barankay and Lockwood (2007). Using Swiss data, the authors show that expenditure decentralization increases educational attaintment. Indeed, expenditure decentralization is strongly associated to the assignment of both teachers appointment and teachers incentive pay to local governments. However, it is not associated to the local power on determining teacher salaries nor structural school organization. It appears that the assignment of the different tasks of educational policy is related to the productive efficiency of government.

Some extensions to the model would cast additional results. Citizens may estimate the performance of politicians and take into account this estimation in order to re-elect or not the incumbent. A way to do so is making T endogenous. In addition, other sources of favoritism are noteworthy. Correlation between shocks seems to be important. If shocks are positively correlated both central and local governments are more likely to reduce efforts, affecting the probability of favoring one region. However, in that case citizens may use yardstick competition to decide whether to re-elect the incumbent. These and other interesting extensions are left for future research.

References

Alesina, A. and G. Tabellini (2004). Bureaucrats or politicians? CESifo Working paper 1130.

Barankay, I. and B. Lockwood (2007). Decentralization and the productive efficiency of government: Evidence from swiss cantons. *Journal of Public Economics 91*, 1197–1218.

Barro, R. (1973). The control of politicians: an economic model. *Public Choice 14*, 19–42.

Besley, T. (2005). Principled agents? the political economy of good government. Book manuscript.

Besley, T. and S. Coate (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics 87*, 2611 – 37.

Bordignon, M., L. Colombo, and U. Galmarini (2003). Fiscal federalism and endogenous lobbies' formation. CESifo Working paper No. 1017.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*, 5–26.

Hindriks, J. and B. Lockwood (2004). Centralization and political accountability. Typescript.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization 7*, 24 – 52.

Lockwood, B. (2002). Distributive politics and the costs of centralization. *Review of Economic Studies 69*(2), 313 – 338.

OECD (1998). Education at a glance: Oecd indicators. *Centre for Educational Research and Innovation*.

Padro i Miquel, G. (2004). Multitask political agency. Working paper.

Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy.* MIT press.

Seabright, P. (1996). Accountability and decentralisation in government: an incomplete contracts model. *European Economic Review 40*(1), 61–89.

Appendix

Proof of Lemma 1 Deriving equilibrium effort in equation (7) with respect to W, $\frac{\partial a_i^C}{\partial W} = \frac{(2T+Y)Y^2}{2(Y^2+2W)^2} > 0$. Deriving equation (7) with respect to Y, $\frac{\partial a_i^C}{\partial Y} = \frac{W\left(2W+Y^2\right)-2YW(2T+Y)}{2(Y^2+2W)^2}$.

This derivative is positive if $W\left(2W+Y^2\right) > 2YW(2T+Y)$ and negative if the inequality goes in the other sense. Rearranging terms the derivative is positive if $Y^2 + 4TY - 2W < 0$. The single positive root of this quadratic expression is $Y = -2T + \sqrt{4T^2+2W}$. If $Y < -2T + \sqrt{4T^2+2W}$ then $\frac{\partial a_i^C}{\partial Y} > 0$. If $Y > -2T + \sqrt{4T^2+2W}$ then $\frac{\partial a_i^C}{\partial Y} < 0$. ■

Proof of Lemma 2 Taking equations (14) and (15), Full decentralization provides larger welfare if $\frac{2pW}{(1+\varepsilon)Y} > \frac{W(2T+Y)}{Y^2+2W}$. Manipulating this inequality, it holds when

$$(1+\varepsilon-2p)Y^2 + 2T(1+\varepsilon)Y - 4pW < 0 \tag{28}$$

If $1+\varepsilon = 2p$, then inequality holds if $Y < \frac{2pW}{T(1+\varepsilon)}$.

If $1+\varepsilon \neq 2p$ the real roots of equation (28) provide the conditions. They are given by the function $\frac{-T(1+\varepsilon)\pm\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{1+\varepsilon-2p}$

If $1+\varepsilon > 2p$, one of the roots is negative. Inequality in Equation (28) holds if $Y < \frac{-T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{1+\varepsilon-2p}$.

If $1+\varepsilon < 2p$, both roots are positive. Inequality in Equation (28) holds if $Y < \frac{T(1+\varepsilon)-\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ or $Y > \frac{T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$. ■

Proof of Lemma 3 This proof follows the same steps of the proof of Lemma 2. Taking equations (14) and (16), partial decentralization provides larger welfare if $\frac{W(2(1+\varepsilon)T+(1+\varepsilon+2p)Y)}{2(1+\varepsilon)(Y^2+W)} > \frac{W(2T+Y)}{Y^2+2W}$. Manipulating this inequality, it holds when

$$Y[(1+\varepsilon-2p)Y^2 + 2T(1+\varepsilon)Y - 4pW] < 0. \tag{29}$$

The expression in squared brackets is equal to Equation (28). The rest of the proof is equal to proof of Lemma 2. ■

Proof of Lemma 4 It follows the same steps of proofs of Lemmas 2 and 3. Taking equations (15) and equation (16), welfare with full decentralization is larger than welfare with partial decentralization if $\frac{2pW}{(1+\varepsilon)Y} > \frac{W(2(1+\varepsilon)T+(1+\varepsilon+2p)Y)}{2(1+\varepsilon)(Y^2+W)}$. Manipulating this inequality, it holds when inequality in Equation (28) holds. The rest of the proof is equal to that of Lemma 2. ■

Proof of Proposition 1 The proof readily follows from the previous Lemmas.

If $1+\varepsilon = 2p$, full decentralization provides larger welfare than the other regimes when $Y < \frac{2pW}{T(1+\varepsilon)}$ (Condition 1 in Lemmas 2 and 4). When $Y < \frac{2pW}{T(1+\varepsilon)}$, full centralization provides larger welfare than the other regimes (Condition 1 in Lemmas 2 and 3).

If $1+\varepsilon > 2p$, full decentralization provides larger welfare than the other regimes when $Y < \frac{-T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{1+\varepsilon-2p}$ (Condition 2 in Lemmas 2 and 4). When $Y > \frac{-T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{1+\varepsilon-2p}$, full centralization provides larger welfare than the other regimes (Condition 2 in Lemmas 2 and 3).

If $1+\varepsilon < 2p$, the quadratic form in equation (28) has real roots $Y = \frac{T(1+\varepsilon)\pm\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$. These roots do not exist if $(1+\varepsilon)^2T^2 < 4pW(2p-(1+\varepsilon))$. Since $1+\varepsilon < 2p$ equation (28) can be rewritten as

$$(2p-(1+\varepsilon))Y^2 - 2T(1+\varepsilon)Y + 4pW > 0 \tag{30}$$

This quadratic form is always positive if $(1+\varepsilon)^2T^2 < 4pW(2p-(1+\varepsilon))$. Therefore, if $1+\varepsilon < 2p$ and $(1+\varepsilon)^2T^2 < 4pW(2p-(1+\varepsilon))$ hold, full decentralization always provides the largest welfare.

If $(1+\varepsilon)^2T^2 > 4pW(2p-(1+\varepsilon))$, the real roots exist and full decentralization generates larger welfare than the other settings when $Y < \frac{T(1+\varepsilon)-\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ or $Y > \frac{T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ (Condition 3 in Lemmas 2 and 4). Full centralization generates larger welfare than the other settings when $\frac{T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ $< Y < \frac{T(1+\varepsilon)-\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$ (Condition 3 in Lemmas 2 and 3).

Finally, from Lemmas 3 and 4, if all the previous conditions on Y are verified with equality, welfare under partial decentralization is equal to welfare with the other settings. As I have shown, otherwise, it is always lower. ■

Proof of Lemma 5 To prove the first part it is enough to compare the indirect utility of the central government with no favoritism and weak favoritism. Let V^C and V^{Cwf} the politician's indirect utility with no favoritism and weak favoritism, respectively. I need to prove that if $Y^2 < \frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2}$ then $V^{Cwf} > V^C$.

With no favoritism the problem of the central government is to maximize

$$\begin{aligned} Wpr\left(S^1\right) - C(a,b) = \\ & W\left[\frac{Y^2 - \left(\frac{Y}{2}+T-U_1(x)\right)\left(\frac{Y}{2}+T-U_2(x)\right)}{Y^2}\right] \\ & -\sum_{i=1}^{2}\left[\frac{a_i^2}{2}+\frac{b_i^2}{2}\right] \end{aligned} \tag{31}$$

over a_1, b_1, a_2 and b_2.

With weak favoritism the problem of the central government is to maximize Equation (31) over a_i, b_i subject to $a_j = b_j = 0$.

Replacing equilibrium efforts from Equation (7) in Equation (31), I obtain, after some simplifications,

$$V^C = \frac{W(-4T^2 + 8W - 4TY + 3Y^2)}{4(2W + Y^2)} \tag{32}$$

Similarly, replacing equilibrium efforts from Equation (17) in Equation (31), I obtain,

$$\begin{aligned} V^{Cwf} =& \frac{W}{4Y^2[(1+\varepsilon)^2Y^2 + 4\varepsilon W]} \times \\ & [(-4T^2 - 4TY + 3Y^2)((1+\varepsilon)^2Y^2 - (1-\varepsilon)^2W) \\ & + 16\varepsilon WY^2 + 4(1-\varepsilon)^2WY^2] \end{aligned} \tag{33}$$

Rearranging terms and after some simplifications $V^{Cwf} > V^C$ if $[2(1-\varepsilon)^2W - (1+\varepsilon)^2Y^2][4T^2 + 4TY + Y^2] > 0$.

The second term in squared brackets is always positive. If $Y^2 < \frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2}$, the first term in squared brackets is positive and, therefore, $V^{Cwf} > V^C$.

To prove the second part, let us compare the welfare. With no favoritism, the central government provides welfare up to

$$\frac{1}{2}U_1^C + \frac{1}{2}U_2^C = \frac{W\,(2T + Y)}{Y^2 + 2W} \tag{34}$$

With weak favoritism, the central government provides

$$\frac{1}{2}U_1^{Cwf} + \frac{1}{2}U_2^{Cwf} = \frac{(1+\varepsilon)^2(2T + Y)W}{2((1+\varepsilon)^2Y^2 + 4\varepsilon W)} \tag{35}$$

Welfare with weak favoritism is larger than welfare with no favoritism if $(1+\varepsilon)^2(Y^2 + 2W) > 2((1+\varepsilon)^2Y^2 + 4\varepsilon W)$. The latter inequality holds if $Y^2 < \frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2}$. ■

Proof of Proposition 2 To prove this Proposition, I need to show three conditions to hold at the same time: first, that the central government is better off favoring a region; second, that it provides larger social welfare favoring a region than not doing so; and third, that it provides larger social welfare than the other institutional settings. Lemma 5 shows that if $Y^2 < \widetilde{Y}^2 \equiv \frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2}$, the first two conditions above hold.

Let us proceed with the last condition. From Proposition 1 and Lemma 2, full centralization provides the largest welfare (with no favoritism), as follows,

1. If $1 + \varepsilon = 2p$, when $Y > \underline{Y}_1$.
2. If $1 + \varepsilon > 2p$, when $Y > \underline{Y}_2$.

3. If $1+\varepsilon < 2p$ and $T^2 > \frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}$, when $\underline{Y}_3 < Y < \overline{Y}_3$.

I need to show which conditions make the threshold $\widetilde{Y}$ to belong to these intervals.

- Case 1: $1+\varepsilon = 2p$. The central government with weak favoritism provides the largest welfare if $\frac{2pW}{T(1+\varepsilon)} < \widetilde{Y}$. It holds if $T > \frac{W}{\widetilde{Y}}$.
- Case 2: $1+\varepsilon > 2p$. The central government with weak favoritism provides the largest welfare if $\frac{-T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2+4pW(1+\varepsilon-2p)}}{(1+\varepsilon-2p)} < \widetilde{Y}$.

  Rearranging gives $\sqrt{T^2\left(1+\varepsilon\right)^2+4pW\left(1+\varepsilon-2p\right)} < \left(1+\varepsilon-2p\right)\widetilde{Y}+T\left(1+\varepsilon\right)$. This inequality holds if $T > \frac{4pW-(1+\varepsilon-2p)\widetilde{Y}^2}{2(1+\varepsilon)\widetilde{Y}}$.

- Case 3: $1+\varepsilon < 2p$ and $T^2 > \frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}$. Let $\widetilde{T} \equiv \sqrt{\frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}}$. The central government with weak favoritism provides the largest welfare if $\frac{T(1+\varepsilon)-\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)} < \widetilde{Y}$ and $\widetilde{Y} < \frac{T(1+\varepsilon)+\sqrt{T^2(1+\varepsilon)^2-4pW(2p-(1+\varepsilon))}}{2p-(1+\varepsilon)}$.

  The first inequality holds if

  $$\left(1+\varepsilon\right)T-\left(2p-\left(1+\varepsilon\right)\right)\widetilde{Y} < \sqrt{\left(1+\varepsilon\right)^2T^2-4pW\left(2p-\left(1+\varepsilon\right)\right)}$$

  Taking squares, simplifying and isolating T, it boils down to $T > \frac{(2p-(1+\varepsilon))\widetilde{Y}^2+4pW}{2(1+\varepsilon)\widetilde{Y}} \equiv \underline{T}$. Following a similar procedure the second inequality holds if $T > \frac{(2p-(1+\varepsilon))\widetilde{Y}^2-4pW}{2(1+\varepsilon)\widetilde{Y}} \equiv \underline{T}'$. Since $\underline{T} > \underline{T}'$, both conditions hold when $T > \underline{T}$.

  I must show that $\underline{T} > \widetilde{T}$, i.e., $\frac{(2p-(1+\varepsilon))\widetilde{Y}^2+4pW}{2(1+\varepsilon)\widetilde{Y}} > \sqrt{\frac{4pW(2p-(1+\varepsilon))}{(1+\varepsilon)^2}}$. Taking squares, simplifying and rearranging terms the inequality becomes

  $$\left(1+\varepsilon\right)^4p^2+4\left(1-\varepsilon\right)^4\left(2p-\left(1+\varepsilon\right)\right)^2+2\left(1+\varepsilon\right)^2\left(1-\varepsilon\right)^2p\left[2p-\left(1+\varepsilon\right)\right] > 0$$

  Since $2p > 1+\varepsilon$ the expression is positive and $\underline{T} > \widetilde{T}$.

■

Proof of Lemma 6 Let $\widetilde{Y} \equiv \sqrt{\frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2}}$, $f(Y) \equiv \frac{\sqrt{(1+\varepsilon)^2Y^2((1+\varepsilon)^2Y^4+8\varepsilon WY^2-4(1-\varepsilon)^2W^2)}}{2(1+\varepsilon)^2Y^2}$, $\underline{T}(Y) \equiv \frac{W}{Y}-f(Y)$ and $\overline{T}(Y) \equiv \frac{W}{Y}+f(Y)$. The politician's utility with strong favoritism is

$V^{Csf} = \frac{W\left(2\left(1+\varepsilon^2\right)W-(1+\varepsilon)^2(2T-Y)Y\right)}{2(1+\varepsilon)^2Y^2}$. With no favoritism is $V^C = \frac{W\left(8W+3Y^2-4T^2-4TY\right)}{4(2W+Y^2)}$.

Utility from strong favoritism is larger than utility from no favoritism if

$4(1+\varepsilon)^2 Y^2T^2 - 8(1+\varepsilon)^2 WYT + 8\left(1+\varepsilon^2\right)W^2 - 8\varepsilon WY^2 - (1+\varepsilon)^2 Y^4 > 0$. This expression is a quadratic form of T, whose real roots are given by $\underline{T}(Y), \overline{T}(Y) = \frac{W}{Y} \pm f(Y)$.

Examining $f(Y)$ there are three cases. If $(1+\varepsilon)^2 Y^4 + 8\varepsilon WY^2 - 4(1-\varepsilon)^2 W^2 < 0$, there are no real roots and strong favoritism is better than no favoritism for all T.

If $(1+\varepsilon)^2 Y^4 + 8\varepsilon WY^2 - 4(1-\varepsilon)^2 W^2 = 0$, $\underline{T}(Y) = \overline{T}(Y) = \frac{W}{Y}$. Therefore, strong favoritism is at least a good as no favoritism to the politician.

If $(1+\varepsilon)^2 Y^4 + 8\varepsilon WY^2 - 4(1-\varepsilon)^2 W^2 > 0$, there are two different real roots.

Notice that $(1+\varepsilon)^2 Y^4 + 8\varepsilon WY^2 - 4(1-\varepsilon)^2 W^2 = 0$ if $Y^2 =$ $\frac{-8\varepsilon W + \sqrt{64\varepsilon^2 W^2 + 16(1+\varepsilon)^2(1-\varepsilon)^2 W^2}}{2(1+\varepsilon)^2} = \frac{-4\varepsilon W + 2W\sqrt{(1+\varepsilon^2)^2}}{(1+\varepsilon)^2} = \frac{2(1-\varepsilon)^2 W}{(1-\varepsilon)^2}$. The last expression is the square of $\widetilde{Y}$.

Therefore, reexpressing in terms of $\widetilde{Y}$: if $Y < \widetilde{Y}$, $V^{Csf} > V^C$; if $Y = \widetilde{Y}$, $V^{Csf} \geq V^C$ (with $V^{Csf} = V^C$ when $T = \frac{W}{Y}$); and if $Y > \widetilde{Y}$,$V^{Csf} > V^C$ when $T \notin \left(\underline{T}(Y), \overline{T}(Y)\right)$.
■

Proof of Lemma 7 Social welfare with partial decentralization is $\frac{1}{2}\ U_1^{PD} + \frac{1}{2}\ U_2^{PD} = \frac{W(2(1+\varepsilon)T + 2(1+\varepsilon)Y)}{2(1+\varepsilon)(W+Y^2)}$ (Eq. (16)) and with strong favoritism is $\frac{1}{2}U_1^{Csf} + \frac{1}{2}U_2^{Csf} = \frac{W}{Y}$.

Considering the cases of Lemma2,

- $2p = 1+\varepsilon$. Welfare with partial decentralization is larger than welfare with strong favoritism if $Y > \frac{W}{T}$.
- $1+\varepsilon > 2p$. Welfare with partial decentralization is larger than welfare with strong favoritism if $(1+\varepsilon-2p)Y^2 - 2(1+\varepsilon)TY + 2(1+\varepsilon)W < 0$. This expression has real roots equal to $\underline{Y}_2^P, \overline{Y}_2^P \equiv \frac{2(1+\varepsilon)T \pm \sqrt{4(1+\varepsilon)^2 T^2 - 8(1+\varepsilon-2p)(1+\varepsilon)W}}{2(1+\varepsilon-2p)}$. Therefore, if $T > \sqrt{\frac{2(1+\varepsilon-2p)W}{1+\varepsilon}} \equiv \underline{T}^P$ and $Y \in \left(\underline{Y}_2^P, \overline{Y}_2^P\right)$ partial decentralization is better to social welfare.
- $(1+\varepsilon) < 2p$. Welfare with partial decentralization is larger if $(2p-(1+\varepsilon))Y^2 + 2(1+\varepsilon)TY - 2(1+\varepsilon)W > 0$. The single positive real roots of this expression is $\underline{Y}_3^P \equiv \frac{-(1+\varepsilon)T + \sqrt{(1+\varepsilon)^2 T^2 + 2(2p-(1+\varepsilon))(1+\varepsilon)W}}{2p-(1+\varepsilon)}$. If $Y > \underline{Y}_3^P$ the equation is positive.

■

Proof of Proposition 3

Partial decentralization provides the largest welfare if,

1. The central government is better off strongly favoring one region, $V^{Csf} > V^C$ and $V^{Csf} > V^{Cwf}$.

2. Welfare provided with strong favoritism is smaller than welfare provided with no favoritism, $\frac{1}{2}U_1^{Csf} + \frac{1}{2}U_2^{Csf} < \frac{1}{2}U_1^{C} + \frac{1}{2}U_2^{C}$.

3. Welfare provided with partial decentralization is even better, $\frac{1}{2}U_1^{PD} + \frac{1}{2}U_2^{PD} > \frac{1}{2}U_1^{Csf} + \frac{1}{2}U_2^{Csf}$.

4. Full decentralization provides less welfare than full centralization (Lemma 2).

I restrict attention to the case in which $Y > \widetilde{Y}$. Therefore if $V^{Csf} > V^{C}$, $V^{Csf} > V^{Cwf}$. Condition 2 holds if $\frac{W}{Y} < \frac{W(2T+Y)}{2W+Y^2}$. Rearranging and simplifying it holds when

$$Y > \frac{W}{T} \tag{36}$$

Considering the cases of Lemma2,

- $1+\varepsilon = 2p$. From Lemma 7, conditions 2 and 3 hold when inequality in Equation (36) is verified. From Lemma 6, if $Y > \frac{W}{T}$ and $Y > \widetilde{Y}$ conditions 2 and 3 hold if $T > \overline{T}(Y)$. Since $\underline{Y}_1$ may be smaller or larger than $\widetilde{Y}$, conditions $1-4$ hold when $Y > max\{\underline{Y}_1, \widetilde{Y}\}$.

- $1+\varepsilon > 2p$. From Lemma 7 condition 3 holds if $Y \in \left(\underline{Y}_2^P, \overline{Y}_2^P\right)$. Since $Y > \widetilde{Y}$, it holds when $\widetilde{Y} \in \left(\underline{Y}_2^P, \overline{Y}_2^P\right)$. A sufficient condition for both $\widetilde{Y} > \underline{Y}_2^P$ and $\widetilde{Y} < \overline{Y}_2^P$ to hold is
$T > \frac{(1+\varepsilon-2p)\widetilde{Y}^2+2(1+\varepsilon)W}{2(1+\varepsilon)\widetilde{Y}} = \frac{(1+\varepsilon-2p)}{2(1+\varepsilon)}\widetilde{Y} + \frac{W}{\widetilde{Y}} \equiv \widehat{T}$. Since $1+\varepsilon > 2p$, this inequality implies equation (36) when Y is not large. Since T must be larger than $\frac{W}{Y}$, T cannot be smaller than $\underline{T}(Y) = \frac{W}{Y} - f(Y)$ and, therefore, T must be larger than $\overline{T}(Y) = \frac{W}{Y} + f(Y)$. Finally, $\widehat{T}$ and $\underline{T}^P$ are larger than $\overline{T}(Y)$ for Y near to $\widetilde{Y}$. The opposite is true for larger Y. Therefore T must be larger than $max\left\{\overline{T}(Y), \underline{T}^P, \widehat{T}\right\}$.

- $(1+\varepsilon) < 2p$. From Lemma7, condition 3 holds if $Y > \underline{Y}_3^P$. Since $Y > \widetilde{Y}$ then either $\widetilde{Y} > \underline{Y}_3^P$ or $\underline{Y}_3^P > \widetilde{Y}$. Taking the definitions in Lemmas 5 and 7, $\underline{Y}_3^P < \widetilde{Y}$ if $T > \frac{2(1+\varepsilon)W-(2p-(1+\varepsilon))\widetilde{Y}^2}{2(1+\varepsilon)\widetilde{Y}}$ and $\underline{Y}_3^P > \widetilde{Y}$ if $T < \frac{2(1+\varepsilon)W-(2p-(1+\varepsilon))\widetilde{Y}^2}{2(1+\varepsilon)\widetilde{Y}}$. Since $1+\varepsilon < 2p$, $\underline{Y}_3^P > \widetilde{Y}$ implies that $T < \frac{W}{\widetilde{Y}} - \frac{(2-(1+\varepsilon))}{2(1+\varepsilon)}\widetilde{Y}$, which contradicts $Y > \widetilde{Y}$ and $Y > \frac{W}{T}$. Therefore $\widetilde{Y}$ must be larger than $\underline{Y}_3^P$. Indeed, since $2p > 1+\varepsilon$, $T > \frac{W}{\widetilde{Y}}$ implies $T > \frac{W}{\widetilde{Y}} - \frac{(2p-(1+\varepsilon))}{2(1+\varepsilon)}\widetilde{Y}$. From Lemma6, $T \notin \left(\underline{T}(Y), \overline{T}(Y)\right)$. From equation (36), T must be larger than $\frac{W}{Y}$. Therefore, partial decentralization is better if $T > \overline{T}(Y)$.

■

Proof of Proposition 4

The equilibrium efforts of local governments with full decentralization are $a_i^L = b_i^L = \frac{pW}{(1+k)(1+\varepsilon)Y}$. Deriving with respect to k, $\frac{\partial a_i^L}{\partial k} = \frac{\partial b_i^L}{\partial k} = \frac{-pW}{(1+\varepsilon)Y(1+k)^2} < 0$.

The equilibrium efforts of the central government with full centralization are: With no favoritism, $a_i^C = b_i^C = \frac{W(2T+Y)}{2(2W+(1+k)Y^2)}$. Deriving with respect to k, $\frac{\partial a_i^C}{\partial k} = \frac{\partial b_i^C}{\partial k} = \frac{-W(2T+Y)Y^2}{2(2W+(1+k)Y^2)^2} < 0$. With weak favoritism, $a_1^{Cwf} = b_1^{Cwf} = \frac{(1+\varepsilon)^2 W(2T+Y)}{2\left((1+k)(1+\varepsilon)^2 Y^2+4\varepsilon W\right)^2}$. Deriving with respect to k, $\frac{\partial a_1^{Cwf}}{\partial k} = \frac{\partial b_1^{Cwf}}{\partial k} = \frac{(1+\varepsilon)^4 W(2T+Y)Y^2}{2\left((1+k)(1+\varepsilon)^2 Y^2+4\varepsilon W\right)^2} < 0$. With strong favoritism, $a_1^{Csf} = b_1^{Csf} = \frac{W}{Y(1+\varepsilon)(1+k)}$, $a_2^{Csf} = b_2^{Csf} = \frac{\varepsilon W}{Y(1+\varepsilon)(1+k)}$. Deriving with respect to k,$\frac{\partial a_1^{Csf}}{\partial k} = \frac{\partial b_1^{Csf}}{\partial k} = \frac{-W}{Y(1+\varepsilon)(1+k)^2} < 0$; $\frac{\partial a_2^{Csf}}{\partial k} = \frac{\partial b_2^{Csf}}{\partial k} = \frac{-\varepsilon W}{Y(1+\varepsilon)(1+k)^2} < 0$. ■

Proof of Proposition 5

From proposition (1), the case for full centralization with no favoritism increases if $\underline{Y}_1, \underline{Y}_2, \underline{Y}_3$ decrease and $\overline{Y}_3$ increases. When $k^L = k^L = k$, welfare becomes $\frac{2pW}{(1+\varepsilon)(1+k)Y}$ with full decentralization and $\frac{W(2T+Y)}{2W+(1+k)Y^2}$ with full centralization.

Full decentralization provides larger welfare than full centralization if
$(1+k)(1+\varepsilon-2p)Y^2 + 2(1+\varepsilon)(1+k)TY - 4pW < 0$.

1. If $1+\varepsilon = 2p$, $\underline{Y}_1 \equiv \frac{W}{(1+k)T}$. Deriving with respect to k, $\frac{\partial \underline{Y}_1}{\partial k} = \frac{-W}{(1+k)^2 T} < 0$.

2. If $1+\varepsilon > 2p$, $\underline{Y}_2 \equiv \frac{T(1+\varepsilon) \pm \sqrt{(1+\varepsilon)^2 T^2 - \frac{4pW(1+\varepsilon-2p)}{1+k}}}{1+\varepsilon-2p}$. Let $C = (1+\varepsilon)^2 T^2 + \frac{4pW(1+\varepsilon-2p)}{1+k}$. Deriving with respect to k, $\frac{\partial \underline{Y}_2}{\partial k} = \frac{-2pWC^{-\frac{1}{2}}}{(1+k)^2} < 0$.

3. If $1+\varepsilon < 2p$, $\underline{Y}_{3,} \overline{Y}_3 = \frac{T(1+\varepsilon) \pm \sqrt{(1+\varepsilon)^2 T^2 - \frac{4pW(2p-(1+\varepsilon))}{1+k}}}{2p-(1+\varepsilon)}$. Deriving with respect to k, $\frac{\partial \underline{Y}_3}{\partial k} = \frac{-2pWC^{-\frac{1}{2}}}{(1+k)} < 0$ and $\frac{\partial \overline{Y}_3}{\partial k} = \frac{2pWC^{-\frac{1}{2}}}{(1+k)^2} > 0$.

■

Proof of Proposition 6

From Proposition (3), Partial Decentralization provides the largest welfare above some thresholds in the shock variability and the exogenous outside option. To prove the first part of this proposition I need to show that the threshold for the shock variability decreases when k increases. From proposition (3), the thresholds are: 1. $max\left\{\underline{Y}_1, \widetilde{Y}\right\}$ if $1+\varepsilon = 2p$; 2. $\widetilde{Y}$ if $1+\varepsilon > 2p$; 3. $max\left\{\underline{Y}_3, \widetilde{Y}\right\}$ if $1+\varepsilon < 2p$.

When $k \neq 0$, these thresholds become $\widetilde{Y} = \sqrt{\frac{2(1-\varepsilon)^2 W}{(1+\varepsilon)^2(1+k)}}$, $\underline{Y}_1 = \frac{W}{(1+k)T}$, and $\underline{Y}_3 =$

$\frac{T(1+\varepsilon)-\sqrt{(1+\varepsilon)^2T^2-\frac{4pW(2p-(1+\varepsilon))}{1+k}}}{2p-(1+\varepsilon)}$. Their derivatives with respect to k are $\frac{\partial \widetilde{Y}}{\partial k}=-\frac{\widetilde{Y}}{2(1+k)}<0$; $\frac{\partial \underline{Y}_1}{\partial k}=-\frac{W}{T(1+k)^2}<0$ and $\frac{\partial \underline{Y}_3}{\partial k}=-\frac{2C^{-\frac{1}{2}}pW}{(1+k)^2}<0$, where $C\equiv(1+\varepsilon)^2T^2-\frac{4pW(2p-(1+\varepsilon))}{1+k}$.

Besides, notice that if $1+\varepsilon>2p$, Y must belong to $\left(\widetilde{Y},\overline{Y}_2^P\right)$. When $k\neq 0$, $\overline{Y}_2^P=\frac{(1+\varepsilon)T+\sqrt{(1+\varepsilon)^2T^2-\frac{2(1+\varepsilon-2p)(1+\varepsilon)W}{1+k}}}{1+\varepsilon-2p}$ and $\frac{\partial \overline{Y}_2^P}{\partial k}=\frac{D^{-\frac{1}{2}}(1+\varepsilon)W}{(1+k)^2}>0$, where $D\equiv(1+\varepsilon)^2T^2-\frac{2(1+\varepsilon-2p)(1+\varepsilon)W}{1+k}$. Therefore, increasing k enlarges the interval $\left(\widetilde{Y},\overline{Y}_2^P\right)$.

To prove the second part of the proposition I need to show that the threshold for the outside option decreases when both k increases and $Y>\widetilde{Y}$ is small enough. From Proposition (3), the minimum T required is $\overline{T}(Y)$ if $(1+\varepsilon)\leq 2p$ and $max\left\{\underline{T}^P,\overline{T}(Y),\widehat{T}\right\}$ if $1+\varepsilon>2p$. When $k\neq 0$ these thresholds become $\overline{T}(Y)=\frac{W}{Y(1+k)^2}+f(Y)$, where $f(Y)=\frac{\sqrt{(1+\varepsilon)^2Y^2\left((1+\varepsilon)^2Y^2+\frac{8\varepsilon WY^2}{1+k}-\frac{4(1-\varepsilon)^2W^2}{(1+k)^2}\right)}}{2(1+\varepsilon)^2Y^2}$; $\underline{T}^p=\sqrt{\frac{2(1+\varepsilon-2p)W}{(1+\varepsilon)(1+k)}}$ and $\widehat{T}=\frac{(1+\varepsilon-2p)}{2(1+\varepsilon)}\widetilde{Y}+\frac{W}{(1+k)\widetilde{Y}}$, where $\widetilde{Y}=\sqrt{\frac{2(1-\varepsilon)^2W}{(1+\varepsilon)^2(1+k)}}$.

Deriving $\overline{T}$ with respect to k, $\frac{\partial\overline{T}(Y)}{\partial k}=-\frac{W}{Y(1+k)^2}+4E^{-\frac{1}{2}}\left[\frac{(1-\varepsilon)^2W^2}{(1+k)}-\frac{\varepsilon WY^2}{(1+k)^2}\right]$, where $E\equiv(1+\varepsilon)^2Y^2\left((1+\varepsilon)^2Y^2+\frac{8\varepsilon WY^2}{1+k}-\frac{4(1-\varepsilon)^2W^2}{(1+k)^2}\right)$. Notice that if $Y=\widetilde{Y}$ then $f\left(\widetilde{Y}\right)=0$ and $\frac{\partial\overline{T}(\widetilde{Y})}{\partial k}=-\frac{W}{\widetilde{Y}(1+k)^2}<0$. Besides, notice that $\frac{\partial f(Y)}{\partial k}>0$ if $Y<Y^*\equiv\sqrt{\frac{(1-\varepsilon)W}{\varepsilon(1+k)}}$; and $\frac{\partial f(Y)}{\partial k}<0$ if $Y>Y^*$. Since $\frac{\partial\overline{T}(Y)}{\partial k}$ is continuous, it is negative for Y near to $\widetilde{Y}$ and becomes positive at some level of Y. Therefore if $1+\varepsilon\leq 2p$, an increase in k will reduce $\overline{T}(Y)$ for $Y>\widetilde{Y}$ small enough.

Finally, the derivatives of $\underline{T}^P$ and $\widehat{T}$ with respect to k are $\frac{\partial\underline{T}^P}{\partial k}=-\frac{(1+\varepsilon-2p)W}{(1+\varepsilon)(1+k)^2}F^{-\frac{1}{2}}<0$, where $F\equiv\frac{2(1+\varepsilon-2p)W}{(1+\varepsilon)(1+k)}$; and $\frac{\partial\widehat{T}}{\partial k}=\frac{1+\varepsilon-2p}{2(1+\varepsilon)}\frac{\partial\widetilde{Y}}{\partial k}-\frac{1}{2(1+k)^2\widetilde{Y}}$. Since $\frac{\partial\widetilde{Y}}{\partial k}<0$ then $\frac{\partial\widehat{T}}{\partial k}<0$. ■