Chapter 10

Redistribution, Decentralization and Constitutional Rules[1]

Federico Etro and Piero Giarda

Introduction

Policymaking at all levels faces a crucial trade-off between transparency of rules and efficiency. At the constitutional level, such an issue often emerges as a trade-off between a close control on the decision-makers and their autonomy.

An extreme degree of transparency is consistent with the possibility to check and balance the use and the abuse of political power. This balance is performed not only by the citizens-voters but also by the mass media and the opinion makers in general. Rules transparency is higher when the relationship between policymakers and their electoral basis is closer. Hence, there is naturally more transparency at the level of local governments than at the level of the central government. For this reason, federal states in which crucial powers are delegated from the central level to the local level are typically characterized by an high degree of transparency.

An extreme degree of independence and insulation of the politicians is a precondition to implement efficient policies and certain reforms which may be undermined by status quo biases. For instance, lobbies or unions representing a minority in the population may have strong tools to avoid reforms which make a majority of the people better off but make them worse off, however a high degree of insulation of a government can overcome this problem. Obviously, the cost of such a system is the excessive concentration of power at the central level. Paradoxically, a dictator has the easiest access to efficient reforms, but history is full of examples in which extensive power is abused.

This trade-off is at the basis of any problem of constitutional design. The horizontal allocation of executive, legislative and judiciary power and their separation is the classic example, but also the vertical allocation of power between local and central governments is a crucial example. The literature on fiscal federalism, for instance, has emphasized the trade-off emerging in problems of public good provision: uniform provision decided at the central level allows to internalise the spillovers between local districts, but it cannot adapt local provision to the local preferences and so it is optimal only when heterogeneity is low (Oates 1972; 1999).

1 We are grateful to Albert Breton, seminar participants, and the discussant for precious suggestions.

Other more specific constitutional rules are clearly based on the trade-off previously suggested: for instance, the period of time between one election and another at any level of government is an example: longer terms insulate more the policymakers and allow them to implement long term reform processes even though they also reduce the accountability of the politicians. A parliamentary privilege has a similar and even more delicate role. The choice of the degree of insulation of central bankers is a more important example from an economic point of view: more independent and conservative central bankers avoid the well known inflationary bias associated with the time-inconsistency problem of monetary policy. Moreover, a certain lack of transparency (in the sense of not revealing all information to the public), allows a certain flexibility in dealing with shocks to the economy without losing anti-inflationary reputation. Finally, simple tax rules are typically adopted for their simplicity even if more complex mechanisms may be optimal, while uniform tariff rates are often adopted to limit the influence of lobbying activity even though differentiated tariffs may be optimal.

From a normative point of view, this trade-off implies the existence of an optimal degree of transparency, which depends on the context to which it applies. In the real world such an optimum is difficult to realize and it is not even clear that there are the right incentives to realize it. The political economy literature has recently emphasized the trade-off between rules transparency and efficient reforms. In particular, Bordignon and Minelli (2001) have endogenized the choice between simple rules which allow to control the politicians in their application and complex rules which are more efficient but also more difficult to be monitored by the citizens and easy to be manipulated by the politicians. More recently, Aghion, Alesina, and Trebbi (2004) have developed a model where reforms are easier to implement when the government is more insulated from the voters and the choice of the degree of insulation is endogenized at a constitutional level. In both models, the trade-off between transparency and efficiency is clear, and more transparency is associated with a greater difficulty to implement efficient reforms.

This chapter belongs to this tradition and deals with the trade-off between transparency and efficiency in the constitutional design of a federal organization. In this context, transparency is synonymous with decentralization since delegation of power at the local level allows policies to be clearly adapted to local political preferences, and efficiency is synonymous with efficient allocation of resources between poor and rich regions (efficient redistribution) and between local public goods and private goods (efficient public goods provision). In our model of fiscal federalism, the trade-off between transparency and efficiency boils down to a *trade-off between decentralization and redistribution*.

We compare a centralized and a decentralized system in a two regions model with income inequality and local public good provision. The former achieves optimal income redistribution with an inefficient allocation of local public goods across the regions, while under decentralization there is limited redistribution and a tendency to generate overprovision of local public goods. Decentralization turns out to be preferred to centralization when inequality is high. Two other crucial results will also suggest a rationale for the trade-off between decentralization and redistribution: first, a transition from centralization to decentralization is supported by both regions

if and only if the central government accepts incomplete redistribution; second, a decentralized system with a House of Regions, where the redistributive policy is decided by bargaining between regions, induces efficient policies but attains only partial redistribution.

The chapter is organized as follows. First, it provides an empirical discussion on the trade-off between decentralization and redistribution. It then presents the model. Following this, it studies the constitutional choice between a centralized and a decentralized system and provides some theoretical rationales for the existence of a trade-off between redistribution and decentralization. Afterwards, it discusses the relationship between transparency in politics and redistributive policies in a wider perspective and then concludes.

The trade-off between decentralization and redistribution

The distribution of powers and functions among different levels of government varies greatly among countries, ranging from the high centralization with complete redistribution and total lack of transparency of the old URSS and other communist countries to the high decentralization with low redistributive programs of the United States and other Anglo-Saxons countries. European continental countries are somewhere in the middle. The distribution of powers is also changing over time: while after World War II there has been a long tendency toward increasing centralization and social spending with the rise of the welfare state in most western countries, the last two decades have witnessed a discernible pattern of increasing decentralization of spending and taxing powers − which in certain cases is even going toward the last line – break up of nations in separate entities[2] − and a stop to the rise of redistributive efforts or even a decrease in them.

A widespread view claims that there is indeed a trade-off between redistribution and decentralization: a centralized state can implement a wider and more efficient redistribution of incomes among individuals and among regions than a decentralized state. In Table 1 we report some preliminary decentralization indexes for a number of OECD countries (built on the basis of the Fiscal Decentralization Indicators of the World Bank, based on Government Finance Statistics (GFS) data from the IMF)[3] ordered by social spending/GDP, which can be seen as an index of redistribution.

2 See Alesina (2003) for a related discussion.

3 We included all the OECD countries except for the smaller ones (Luxembourg, Cyprus, Iceland, Portugal, Greece, Israel and New Zealand) since the relevance (and even the same concept) of decentralization vanishes (or at least changes) for small countries. Developing countries were not considered because, despite they may also show a trade-off between decentralization and redistribution, they are characterized by less democratic and more centralized regimes, and hence they are not directly comparable with the developed countries. Ex Communist Countries were not considered since their process of decentralization was mainly developed through the break up of nations so that most of the subnational expenditures in the ex URSS, the ex Yugoslavia and the ex Czechoslovak became central expenditures in the new born countries. It would be interesting to verify that the trade-off between redistribution and decentralization holds for other homogeneous groups of countries.

Table 10.1 Decentralization Indexes

Countries ranked by increasing Social Spending/ GDP	Subnational Public Expenditures (percent of total − averages over 1972-2000)	Subnational Tax Revenue (percent of total revenues and grants − averages over 1972-2000	Vertical Imbalance (in percentages)
USA	44.6	47.6	32.6
Australia	41.2	33.1	44.4
Japan	43.5	99.0	NA
Canada	57.4	56.0	26.4
Switzerland	51.9	52.6	22.8
Finland	38.1	49.3	31.5
UK	25.4	23.8	53.7
Ireland	25.1	8.8	66.8
Norway	34.6	45.6	32.3
Spain	20.0	41.5	40.4
Italy	21.2	19.1	65.4
Denmark	45.4	44.7	46.1
Germany	42.1	54.5	22.1
Austria	30.8	50.5	28.5
France	18.2	47.9	36.4
Sweden	37.7	62.6	22.8
Netherlands	25.0	7.1	75.1
Belgium	11.8	34.2	54.9

We adopt as a main Decentralization Index the *share of subnational (local, state and provincial) public expenditure on total public expenditure*. The main implication we want to draw from the table is that a preliminary look at these cross country data is in favour of a trade-off between decentralization and redistribution.

Notice that the variability of the Decentralization Index is very small in the OECD countries. Figure 10.1 illustrates its pattern for OECD countries and some other groups of countries in the period 1972-1997. Despite the lack of comprehensive data makes it difficult to evaluate these patterns, it emerges as a stable behaviour of the fraction of local public spending with at most a recent moderate upward trend.

Some problems with our index of decentralization should be emphasized. Defence and interest payments are rarely decentralized, and they can potentially distort measures of decentralization. For instance, the United States, despite being a much larger country, has a lower subnational share of expenditures than Switzerland. However, when defence and interest expenses are excluded from the subnational-to-total ratio, the United States has a higher subnational share of expenditures than Switzerland. While various expenditure patterns can be assessed by the GFS data, less can be said about expenditure autonomy. Expenditures that are mandated by the central government appear as subnational expenditures, even though subnational

governments may have no autonomy in these spending decisions. Below, we also provide data on the *share of subnational (local, state and provincial) tax revenues on total tax revenue*, which is however a worst measure of fiscal decentralization. Indeed, a typical question asked in a decentralization context is how much autonomy do subnational governments have in raising revenue; for instance how much is collected through shared taxes versus piggybacked taxes versus locally determined taxes? Shared taxes appear as subnational revenue, although the subnational government has no autonomy in determining the revenue base or rate, since the GFS reports revenues based on which level of government ultimately receives the revenues. More qualitative information on legal and constitutional rules about fiscal decentralization is needed to build a better index based on all the available information. Finally, we provide an index of *vertical imbalance*, or the degree to which subnational governments rely on central government revenues to support their expenditures. This can be measured by intergovernmental transfers as a share of subnational expenditures. This measure does not distinguish what proportion of transfers is conditional versus general purpose, and the GFS data do not provide this information. Hence also this is not a good measure of fiscal decentralization. Finally we show in Figure 10.2 data on the Decentralization Index and the average *share of subnational (local, state and provincial) public expenditure on GDP* in the period 1972-2000 for the selected group of OECD countries. There is a high correlation between the two data sets (0.79) while the correlation between the Decentralization Index and the average *share of total public expenditure on GDP* is negative (–0.25).

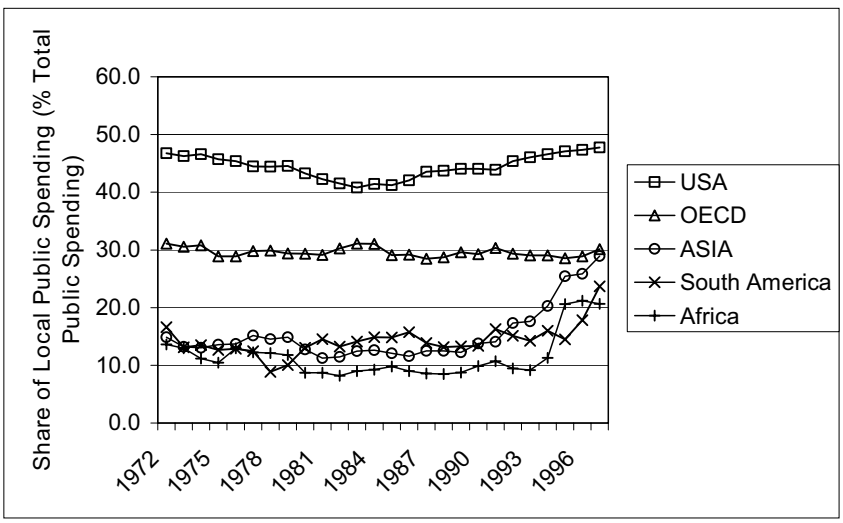

Figure 10.1 Patterns of Decentralization

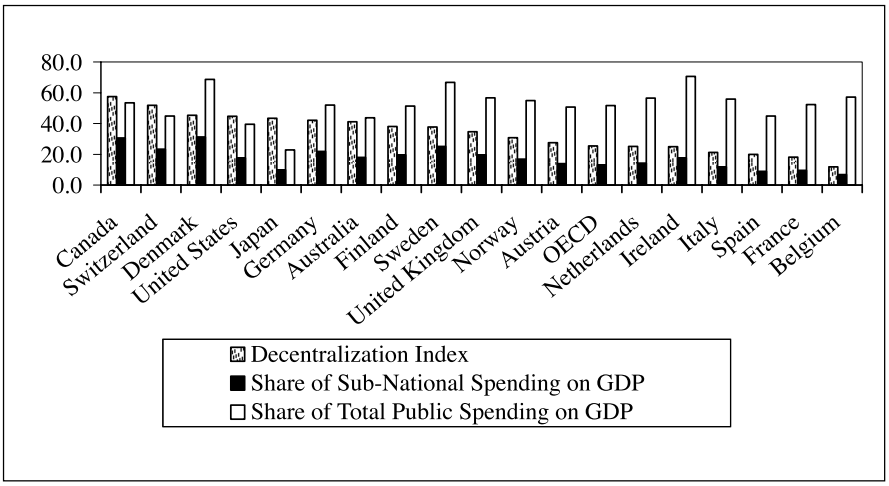

Figure 10.2 OECD Countries

Table 10.1 and Figure 10.1 suggest that public spending, as a proxy for redistribution spending, and decentralization are negatively correlated. Despite such a correlation may be spurious, since countries with high decentralization are also more liberal countries with stronger preferences for right wing policies and, hence, less redistribution, and also other factors (like social heterogeneity, national wealth, growth, labour market conditions affecting the distribution of earnings, size, openness) may affect both indexes, it is important to understand if and why such a correlation exists. A more accurate empirical research should try to disentangle the correct relationship between decentralization and redistribution controlling for other political, sociological and economic factors.

In the rest of the chapter we turn to the theory. It is not clear at all from a theoretical point of view why there should be a trade-off between redistribution and decentralization. A number of economic, sociological, cultural and political factors affect both variables and in many cases this happens in different directions. Some may well act against our suggested hypothesis. For instance, higher (regional) inequality may be associated with a stronger tendency toward decentralization but also toward more income redistribution for well known political economic reasons (Meltzer and Richard 1981): if this is true, the pure negative correlation between decentralization and redistribution may be even stronger than what is suggested by our previous data. In other words, we need a more accurate theoretical investigation of the relation between redistribution and decentralization.

A model of interregional redistribution

We will adopt a simple model of fiscal federalism between two regions. Since we want to relate redistribution and decentralization, we assume that the two regions

are heterogeneous in income and have the same preferences.[4] In particular, let us consider a country composed of two regions, of equal population normalized to one, with different incomes, w for the poorer region B and kw with, $k \geq 1$, for the richer region A. Hence, the higher is the parameter *k*, the higher is inequality between the regions. Utility in region *i=A,B* is given by:

$$U_i = u_i(c_i) + H_i(g_i)$$

where the sub-utilities of the two regions over net income (or private consumption) c_i and expenditure on a regional public good g_i are given by the same functions. For the moment we will not introduce spillovers from public good production in the other region.

For the sake of simplicity we will focus on the particular case in which $u_i(x) = H_i(x) = \log x$, so as to obtain closed form solutions. The public good is financed through a proportional income tax at rate t_i. We will always assume that the Constitution sets a social welfare function of the Bentham variety for the country as a all, that is an un-weighted sum of the regional utility functions: [5]

$$W = u_A(c_A) + H_B(g_A) + u_B(c_B) + H_B(g_B)$$

As a benchmark for our model, we establish what is the first best allocation of resources in this model. This could be implemented choosing the tax rates t_A and t_B and two local public expenditures g_A and g_B to maximize:

$$W = u_A\left[kw(1-t_A)\right] + u_B\left[w(1-t_B)\right] + H_A(g_A) + H_B(g_B)$$
$$s.v. kwt_A + wt_B = g_A + g_B$$

The first order conditions for the optimum centralized redistribution are:

$$u'_A(c_A) = u'_B(c_B) = H'_A(g_A) = H'_B(g_B)$$

which imply equalization of private consumption and public spending across the regions. In particular, under our functional form assumptions we obtain:

$$c_A^{FB} = c_B^{FB} = g_A^{FB} = g_B^{FB} = \frac{(k+1)w}{4}$$

and the utilities:

4 Income heterogeneity is also considered by Alesina, Angeloni and Etro (2005) and especially Bordignon, Manasse and Tabellini (2001) For models with heterogeneity in preferences see Alesina, Angeloni and Etro (2001; 2005) and Besley and Coate (2002). Persson and Tabellini (1996a; 1996b) are the classic references on risk sharing in fiscal federalist constitutions.

5 The earliest example of such a welfare function in the analysis of fiscal regional problems we are aware of is Pantaleoni (1891).

$$U_A^{FB} = U_B^{FB} = \log\left[\frac{(k+1)^2 w^2}{16}\right] \qquad W^{FB} = 2\log\left[\frac{(k+1)^2 w^2}{16}\right]$$

where *FB* clearly stands for first best.

If the two regions are completely separated in their decision making, region A would choose t_A to maximize $U_A = u_A[kw(1-t_A)] + H(kwt_A)$ and region B would choose t_B to maximize $U_B = u_B[w(1-t_B)] + H_B(wt_B)$, from which $t_A^I = t_B^I = 1/2$, which implies:

$$c_A^I = g_A^I = \frac{kw}{2} \qquad c_B^I = g_B^I = \frac{w}{2}$$

with utilities:

$$U_A^I = 2\log\left(\frac{kw}{2}\right) \quad U_B^I = 2\log\left(\frac{w}{2}\right) \quad W^I = 2\log\left(\frac{kw^2}{4}\right)$$

where *I* stands for independence (a possible bench-mark) situation with no inter-regional redistribution. Notice that $W^{FB} > W^I$ for any *k*, but also the independence outcome is Pareto-efficient, since any redistribution to the poorer region must hurt the richer one.

Let us now consider the two regions as members of a federal or unitary state country. The Constitution assigns the production of the public good to regions but the policy of taxes and transfers can follow a centralized system or a decentralized one.

Centralization

Under centralization, the Constitution requires a national and uniform income tax rate in all regions and no regional or central government deficit spending. Furthermore, it may entitle the central government: (i) to levy a lump sum tax on citizens in the rich regions to finance a negative tax program in favour of citizens in the poor regions *or* (ii) to set up a transfer programs by which rich regions are "forced" to transfer financial resources to poor regions. The parameter θ_1 represents the instrumental fiscal variable used to generate the rich citizen-to-poor citizen transfer; the parameter θ_2 represents a financial transfer from region A to region B required to ensure that regional spending on public goods is fully financed (by tax revenue and interregional transfer). We will assume that the Central Government cannot, at the same time, use both instruments: we assume that it can only redistribute income or transfer resources across regions. The first alternative being more interesting for our purposes we will focus on it, but comment also on the latter. In general such a constraint is well grounded on the limited availability of tax instruments of the government and it may be a short cut representation of its second best environment.

More formally, let us start considering the environment where the central government can choose the tax rate *t* and the lump-sum transfer θ_1 from rich to poor citizens. Local public expenditure is financed by each region through its own

tax revenue as the central government does not have the power to enforce financial transfers from rich to poor Regions; in other words $\theta_2 = 0$. The objective function (where, for simplicity, θ stands for θ_1), is:

$$W = u_A[kw(1-t)-\theta\ ]+u_B[w(1-t)+\theta\ ]+H_A\left(kwt\right)+H_B\left(wt\right)$$

The first order conditions (with respect to *t* and θ) are:

$$u'_A[kw(1-t)-\theta\ ]=u'_B[w(1-t)+\theta\ ]=\frac{kH'_A\left(kwt\right)+H'_B\left(wt\right)}{k+1}$$

from which $t^C = 1/2$ and $\theta^C = (k-1)w/4$, where *C* stands for centralization. The allocation of resources is:

$$c_A^C = c_B^C = \frac{(k+1)w}{4} \qquad g_A^C = \frac{kw}{2} \qquad g_B^C = \frac{w}{2}$$

The outcome implies the right allocation of resources between private and public consumption in the country, but while we obtain perfect equalization of private consumption at the first best level for both regions – or optimal redistribution in our set up − there is no equalization of public expenditure, which is the same as in the independent regions case. This inefficiency is due to the lack of enough redistributive instruments which is constraining the central government. The utilities of the two regions are:[6]

$$U_A^C = \log\left(\frac{(k+1)kw^2}{8}\right) \quad U_B^C = \log\left(\frac{(k+1)w^2}{8}\right) \quad W^C = 2\log\left(\frac{(k+1)kw^2}{8\sqrt{k}}\right)$$

The alternative environment would be one where the central government can only choose the uniform tax rate and the inter-regional transfer directed to finance public expenditure in the poor region. In this case the situation is symmetric to the previous one with perfect equalization of public expenditure (at the first best level), but not of private consumption (which comes out equal to the value computed in the independent regions case).

Summarizing, *a centralized setting in which the central government can only redistribute personal incomes (or, alternatively, enforce interregional transfers) generates equalization of incomes after taxes (alternatively, equalization of local public expenditure) but differentiation in public good provisions (alternatively, incomes after-taxes).*

6 Of course, there is no reason why region A, if given the choice, should have chosen to sign a founding contract for a federation or an unification with region B. Nor that it should be glad to stay in the contract. In any case, the model has nothing to say on the birth or death of federations.

Decentralization

Consider now the case where the Constitution, while maintaining the same social welfare function, assigns tax autonomy to regional governments and assumes that private consumption and regional public goods production be regulated by lump-sum taxes or transfers.

The allocation of fiscal powers to regional and federal governments poses some problems of timing and of budgetary procedures. It is possible to envisage at least two alternatives. Remember that budgetary decisions by regional and central governments are interdependent. Somehow, one of the two levels of governments has to make a first step. Consequently, in our model it is possible to envisage at least two alternatives. In the first alternative, the central government chooses the redistributive policy first and the regions choose their regional tax rates afterwards. In the second alternative, the regions move first and choose their tax rates and the central government steps in a second stage to decide on the redistributive policy. These two different time sequences can be taken to describe two different budgetary procedures.

We will focus on the case in which regions are first movers and the central government cannot commit to a redistributive policy, since this is more realistic (indeed it is simple for the central government to operate transfers between regions, while tax rates are typically set in advance and require some time to be changed). This no-commitment case can be described by the following sequence of decisions:

1. each region chooses its own tax rate t_A and t_B : a) taking the choice of the other region as given; b) knowing that the central government is bound by the Constitutional objectives as defined in 2).
2. the tax-transfer θ is chosen that maximizes the social welfare function.

Each region knows that the government maximizes:

$$W^D = u_A[kw(1-t_A)-\theta]+u_B[w(1-t_B)+\theta]+H_A\left[kwt_A\right]+H_B\left[wt_B\right]$$

where *D* stands for decentralization. For given tax rates, the first order conditions imply:

$$u'_A[kw(1-t_A)-\theta]=u'_B[w(1-t_B)+\theta]$$

which generates equalization of private consumption in the two regions through the following transfer function:

$$\theta^D(t_A,t_B)=\frac{kw(1-t_A)-w(1-t_B)}{2}$$

Each region takes this in consideration so as to maximize:

$$U_A=u_A\left[kw(1-t_A)-\theta^D(t_A,t_B)\right]+H_A\left(kwt_A\right)=u_A\left[\frac{kw(1-t_A)+w(1-t_B)}{2}\right]+H_A\left(kwt_A\right)$$

and

$$U_B = u_B\left[w(1-t_B)+\theta^D(t_A,t_B)\right]+H_B\left(wt_B\right)=u_B\left[\frac{kw(1-t_A)+w(1-t_B)}{2}\right]+H_B\left(wt_B\right)$$

Taking into consideration the optimality condition for the central government tax-transfer, the first order conditions of the two regions imply:

$$u'_A\left[\frac{kw(1-t_A)+w(1-t_B)}{2}\right]=2H'_A\left(kwt_A\right)=u'_B\left[\frac{kw(1-t_A)+w(1-t_B)}{2}\right]=2H'_B\left(wt_B\right)$$

Given our functional form assumptions, Regions choose the tax rates:

$$t_A^D=\frac{k+1}{3k} \quad t_B^D=\frac{k+1}{3}$$

which imply the transfer $\theta^D=(k-1)w/2$. Utilities are:

$$U_A^D=U_B^D=\log\left(\frac{(k+1)^2w^2}{18}\right) \quad W^D=2\log\left(\frac{(k+1)^2w^2}{18}\right)$$

The allocation of resources is inefficient since public expenditure is now higher than in the first best while private consumption is lower than in the first best:

$$c_A^D=c_B^D=\frac{(k+1)w}{6} \qquad g_A^D=g_B^D=\frac{(k+1)w}{3}$$

Notice that income redistribution is less than optimal and less than under centralization ($c_i^D<c_i^C=c_i^{FB}$ for *i=A,B*). The nature of this inefficiency is interesting. Each region is now taking in consideration the effect of its own tax rate on the redistributive policy adopted by the nation state government. The higher is the taxation of the poorer region, the higher will be the negative tax assigned to this region. Hence, the incentive to set a higher tax rate. The opposite happens to the richer region, but the overall result is over-production of regional public goods, which is in contrast to the one derived by Bordignon et al. (2001), who find a tendency toward under-provision of regional public goods. It is to be noted that their result would be obtained if our model were constructed on a the central government with the possibility to enforce interregional transfers − θ_2 − to finance public goods production but without the possibility to apply the tax-transfer θ_1 on individuals of Regions A and B. Such a model would generate perfect equalization of public expenditure and private consumption, but with under-provision of public goods. Utilities would remain U_A^D and U_B^D as previously defined and country utility would remain W^D.[7]

7 Indeed, the Bordignon, Manasse and Tabellini model assumes that the central government can only force such an interregional financial transfer that equalizes public expenditure in both regions, thus proportional to the difference in tax revenues. As a consequence, both the rich and the poor regions have small incentives to tax: higher local taxes

Summarizing, *a decentralized setting in which the central government only controls its tax parameters (or, alternatively, can only enforce interregional financial transfers) obtains perfect redistribution of income and equalization of local public expenditure but generates over-provision (alternatively, under-provision) of public goods. It is important to remember that this result rests on the institutional assumption that regional governments are assigned a prominent role in the making of the economy-wide budgetary policy.* In later pages this solution will be referred to as a case of constrained decentralization.

As Bordignon et al. (2001) have shown, when the Central Government can commit to the redistributive policy, the first best outcome can be achieved. Moreover, if it could choose both a transfer of income and an inter-regional financial transfer, even without such a commitment, the first best allocation of resources could also be achieved. Once again the lack of a full set of policy instruments is crucial for the emergence of inefficiencies.

Discussion

In the real world, redistribution is implemented with different programs, with the central government involved primarily in personal income redistribution via taxation of income at differentiated rates and with richer regions forced to participate in equalization plans directed to the financing of local public goods in poorer regions.[8] If the redistributive activity is more relevant, we have shown that:
under centralization optimal income redistribution is achieved but the allocation of local public goods across the regions is suboptimal, while under decentralization there are limited redistribution and a tendency to generate overprovision of local public goods.[9]

---

for the rich region are lost in part to increase transfers to the poor one and higher local taxes in the poor region are lost in a partial reduction of the transfer from the rich region. Both regions reduce taxation – compared to the optimal level obtained under centralization – and public expenditure is sub-optimal. Our model, instead, focuses on redistribution of personal income, in the sense that the richer region finances, with the yield of the central government tax, part of the private consumption of the poor region. Hence, the central government chooses a tax-transfer so as to equalize private consumption in both regions (that is to implement a perfect redistribution) and it chooses a transfer which is proportional to the difference in net income. As a consequence, the two regions have now different incentives. For the poor region there is an incentive to increase local tax rates since this increases the net income differential between the regions, and hence it increases redistribution from the rich region. This effect determines the increase in local taxation and hence the overprovision of public goods in equilibrium.

8 Our result seems to be relevant not only for the constitutional design of fiscal federalism in countries, but also for the organization of unions of countries. For instance, the European Union is exactly characterized by a group of countries with large autonomy on their policies of taxation and expenditure in public goods, while a net of transfers between countries is managed at the union level with redistributive purposes: our model emphasizes that in this environment countries may engage in excessive expenditures on local public goods.

9 Such a result is more plausible than the under provision result from an empirical point of view. Indeed, the correlation of our index of decentralization with the average share of subnational public expenditure in GDP is highly positive; 0.87 worldwide and 0.9 for the

Clearly the tendency would be limited in case of spillovers between regions due to a well known free-riding problem.

Centralization versus decentralization

In this section we will deal with the choice between centralization and decentralization and the relationship between inequality and such a choice. It is interesting to investigate under what conditions the transition a) from the independent regions setting to a centralized or a decentralized system and b) from a centralized to a decentralized system will be accepted. Since the function of the central government in the model is to implement income redistribution, it is clear that the degree of inequality will play a crucial role in answering the question.

We start comparing centralization with independence. On the one side, the rich country always prefers a no-central government setting to a centralized system. Indeed $U_A^C - U_A^I = \log[(k+1)/2k] < 0$ for any k since under centralization the rich region redistributes income toward poor region. On the other side, the poorer region, that is benefiting from redistribution, is always in favour of centralization, no matter how inefficient. In fact $U_B^C - U_B^I = \log[(k+1)/2] \geq 0$ for any k.

The most important comparison is between the nation-wide utilities under independency and under centralization. It shows that a centralized system is always preferred to independence:

$$W^C - W^I = \log\left(\frac{(k+1)^2}{4k}\right) \geq 0 \text{ for any } k$$

Though the allocation of resources may be not perfectly efficient under centralization, a move from independence to centralization makes the poor region better off enough to compensate the losses of rich region: *centralized redistribution always makes the country better off compared to independence.*

Now let us move to decentralization. Again, the rich region always prefers the no-central government situation to this decentralized system. Indeed $U_A^D - U_A^I = \log\left[2(k+1)^2/9k^2\right] < 0$ for any k. More interestingly, the poor region, that is benefiting from redistribution, is not always in favour of decentralization, since this implies a considerable inefficiency. More exactly we have $U_B^D - U_B^I = \log\left[2(k+1)^2/9\right] \geq 0$ if and only if $k \geq \left[\sqrt{18}/2 - 1\right]$, that is, the poorer region prefers the decentralized structure if and only if income inequality is great enough. The reason is that under decentralization, only when inequality is very high are the benefits from inefficient redistribution are higher than those from the efficient allocation of resources under independence.

OECD countries. In other words, more decentralized countries tend to spend a higher fraction of their income in local goods than centralized countries. Notice that in our model there is no difference between local public expenditure in pure public goods or private goods, hence it is correct to look at data on local public expenditure without distinguishing between provision of public and private goods.

The comparison of the nation-wide utilities under independence and decentralization shows that:

$$W^D - W^I = \log\left(\frac{4(k+1)^4}{81k^2}\right) \geq 0 \text{ if and only if } k \geq 2$$

Again, *only a great enough degree of income inequality justifies a transition from regional independence to the inefficient decentralized redistribution*.

Finally let us compare the model of centralization with the model of decentralization. From the point of view of the rich region, centralization is always better:

$$U_A^C - U_A^D = \log\left(\frac{9k}{4(k+1)}\right) > 0 \text{ for any } k$$

the reason being that centralization allows to limit redistribution. Instead, the poor region prefers centralization if and only if the degree of inequality is low enough:

$$U_B^C - U_B^D = \log\left(\frac{9}{4(k+1)}\right) > 0 \text{ if and only if } k \in [1,1.25]$$

Most important, centralization is the efficient choice if and only if the degree of inequality is low enough and decentralization is better when inequality is very high:

$$W^C - W^D = 2\log\left(\frac{9\sqrt{k}}{4(k+1)}\right) > 0 \text{ if and only if } k \in [1,(9+\sqrt{17}))^2/64]$$

Heterogeneity between the regions must be very high for the gains from decentralization to compensate the losses in efficiency associated with decentralized redistribution. Thus, there are some good reasons for a very heterogeneous country to be organized in a decentralized way, while the opposite holds for a relatively homogeneous country.

Notice that if we were going to introduce spillovers between local public goods (as in Alesina, Angeloni and Etro, 2001; 2005) a further tendency toward under provision of local public goods would emerge under decentralization. In such a model, decentralization would be socially preferred to centralization if and only if income inequality between the regions was great enough or if spillovers between local public goods were small enough.

In the rest of this section we will show other ways in which our simple model can rationalize the trade-off between redistribution and decentralization. First, we will show that a transition from a centralized system to an efficient decentralized one always implies giving up to some redistribution to be politically feasible. Later on, we will discuss a system of fiscal federalism with a House of Regions where an efficient outcome is realized with partial redistribution.

Transition from centralization to decentralization

Now we analyse an issue of transition. We consider a reform which moves the organization from a centralized setting to a decentralized one, as many countries have recently experienced or are experiencing. The question is: does a reform towards decentralization limit the extent to which the central government can redistribute income? We will show that for such a reform to be supported by the whole country, the level of redistribution must be lower than it would be required by the maximization of the social welfare function. This result shows that a trade-off between decentralization and redistribution emerges especially in transitions towards decentralization.

We assume that for the reform to be adopted, both regions must agree. Hence they must be both better off under decentralization than under centralization. However, we know that if the equilibrium decentralization is implemented, the rich region is worse off if inequality is great enough (and the poor region better off). Hence, the only way to implement decentralization is to limit redistribution in such a way that both regions benefit from it.

In particular the reform is feasible if the transfer between regions is limited by a cap on the redistributive tax-transfer $\bar{\theta}$. If this is binding, regions will consider $\bar{\theta}$ as not affected by their choices and will maximize:

$$U_A = u_A\left[kw(1-t_A)-\bar{\theta}\right]+H_A\left(kwt_A\right)$$

and

$$U_B = u_B\left[w(1-t_B)+\bar{\theta}\right]+H_B\left(wt_B\right)$$

from which we obtain the optimality conditions:

$$u'_A\left[kw(1-t_A)-\bar{\theta}\right]=H'_A\left(kwt_A\right)$$

$$u'_B\left[w(1-t_B)+\bar{\theta}\right]=H'_B\left(wt_B\right)$$

that is the functions $t_i = t_i(\bar{\theta})$ for $i=A,B$.

A credible value of the cut-off $\bar{\theta}$ is actually such that the rich region is just indifferent between centralization and decentralization. More precisely $\bar{\theta}$ is defined by the indifference relationship:

$$u_A\left[kw(1-t_A(\bar{\theta}))-\bar{\theta}\right]+H_A\left(kwt_A(\bar{\theta})\right)=u_A\left[kw(1-t)-\theta_1\right]+H_A\left(kwt\right)$$

Under our assumption regarding functional forms, we have:

$$\log\left(\frac{kw-\bar{\theta}}{2}\right)^2=\log\left[\frac{(k+1)kw^2}{8}\right]\;\Leftrightarrow\;\bar{\theta}=w\left[k-\sqrt{\frac{(k+1)k}{2}}\right]$$

To make sure that this is a right equilibrium we need to verify that also the poor region is always better off under such a form of partial decentralization, and that $\overline{\theta} < \theta^{DNC}$ so that the constraint is indeed binding; these two conditions can be easily verified to always hold.

Finally notice that this form of partial redistribution implies tax rates $t_A = \sqrt{\frac{(k+1)k}{8}}$ and $t_B = \frac{k+1}{2} - \sqrt{\frac{(k+1)k}{8}}$ from which we derive the allocations:

$$c_A = g_A = \frac{w}{2}\sqrt{\frac{(k+1)k}{2}} \qquad c_B = g_B = \frac{w}{2}\sqrt{\frac{(k+1)k}{8}}\left[\sqrt{\frac{2(k+1)}{k}} - 1\right]$$

This makes clear that an efficient allocation of resources is realized in each region, but redistribution from the rich to the poor is only partial. It is important to notice that the limit to redistribution is beneficial because it avoids the time-inconsistency problem emerging under decentralization: $\overline{\theta}$ is clearly independent from the tax rates chosen by the two regions. For this reason, despite the fact that best allocation of resources under constrained decentralization improves on constrained centralization only when inequality is high (as shown in the previous section), the feasible decentralization with partial redistribution always improves on centralization.

Summarizing, we have shown that *a transition from constrained centralization to a decentralization is supported by both regions if and only if the central government abandons the objective of complete redistribution.*

Fiscal federalism with a House of Regions

We now describe a system of fiscal federalism in which a House of Regions has the power to decide the redistributive parameter of the central government income tax (in previous notation, the parameter θ). Let us consider the decentralized system modified according to the following timing:

1. the regions decide on the redistributive transfer θ by bargaining;
2. each region chooses its own tax rate t_A and t_B a) taking the choice of the other region as given; b) knowing the transfer θ chosen in stage 1).

Given the transfer, the two regions maximize:

$$U_A = u_A\left[kw(1-t_A) - \theta\right] + H_A\left(kwt_A\right)$$

and

$$U_B = u_B\left[w(1-t_B) + \theta\right] + H_B\left(wt_B\right)$$

from which we obtain the optimality conditions:

$$u'_A\left[kw(1-t_A)-\theta\right]=H'_A\left(kwt_A\right)$$

$$u'_B\left[w(1-t_B)+\theta\right]=H'_B\left(wt_B\right)$$

that is the reaction functions $t_i=t_i(\theta)$ for *i=A,B.* This implies the utilities:

$$U_A\left(\theta\right)=u_A\left[kw(1-t_A(\theta))-\theta\right]+H_A\left(kwt_A(\theta)\right)$$

$$U_B\left(\theta\right)=u_B\left[w(1-t_B(\theta))-\theta\right]+H_B\left(kwt_B(\theta)\right)$$

We now assume that the redistributive policy is chosen by Generalized Nash Bargaining, assuming that the relative bargaining power of the rich region is η. Hence the equilibrium transfer solves the problem:

$$\max U_A\left(\theta\right)^{\eta}U_B\left(\theta\right)^{1-\eta}$$

Under our functional form assumptions, the problem becomes:

$$\max \eta\left[2\log\left(\frac{kw-\theta}{2}\right)\right]+(1-\eta)\left[2\log\left(\frac{w+\theta}{2}\right)\right]$$

whose solution is:

$$\theta(\eta)=w\left[(1-\eta)k-\eta\right]$$

Notice that when the bargaining power of the two regions is the same ($\eta=1/2$), the transfer obtained under unconstrained decentralization emerges and the first best solution follows; when the relative bargaining power of the two regions is k ($\eta=k/(k+1)$), the bargained transfer is zero and the independence outcome follows. The final utilities are:

$$U_A(\eta)=\log\left(\frac{(k+1)^2w^2\eta^2}{4}\right) \text{ and } U_B(\eta)=\log\left(\frac{(k+1)^2w^2(1-\eta)^2}{4}\right)$$

and they depict the set of efficient outcomes. The country's utility is:

$$W(\eta)=2\log\left(\frac{(k+1)^2w^2\eta(1-\eta)}{4}\right)$$

and it describes the welfare frontier.[10] The bargaining index is likely to range between the lower bound of η = ½ (equal regional power) and the upper bound

10 Notice that such a result is better than the result under constrained decentralization from the point of view of the social welfare if one of the two regions has not too much bargaining power (in particular $W(\mathsf{h})\geq W^D$ if $\mathsf{h}\in[1/3,2/3]$).

of $\eta = k/(k+1)$ (power proportional to income). The first case delivers the welfare of the full centralization case (the first best solution $W(\eta) = W^{FC}$); the second delivers the welfare of the independence situation $W(\eta) = W^{I}$ with no redistribution. Within the interval, values of η generate cases of partial redistribution.

Notice that this system of extreme budget decentralization solves the time inconsistency problem generated when the regions do not bargain. An extreme decentralization setting integrated by interregional bargaining (as it might happen if a House of Regions were in effect) induces efficient policies with redistribution outcomes depending upon distribution of powers among the bargaining agents. In a representative democracy, distribution of power in a House of Regions will depend on population size, electoral rules and, behind them, on the relative wealth of the regions.[11]

Regional cooperation as a reputational outcome

Another way in which decentralization can achieve an efficient allocation of resources is available in a dynamic setting or, in other words, when the budgetary "game" is repeatedly played over time. Under such circumstances, well known reputational arguments allow the regions to obtain any efficient outcome (as characterized in the previous subsection) if they are 'patient' enough. For simplicity we will just show how the first best allocation can be obtained if the discount factor δ is high enough. The important result is that, despite what one may think about the way inequality should affect the incentives of both regions to cooperate, in our example such a reputational argument works independently with the degree of inequality.

More precisely let us focus on an infinite periods extension of the model and on 'trigger strategies'. In equilibrium each region chooses the tax rates $t_A = (k+1)/4k$ and $t_B = (k+1)/4$ and the government implements the perfect redistributive scheme so as to guarantee the first best utilities in each period U_i^{FB} for $i = A,B$. To verify that this is indeed an equilibrium, imagine that one of the two regions deviates and chooses an alternative tax rate. It is easy to show that the best deviation generates the one-period utility:

$$U_i^{Dev} = \log\left(\frac{3}{8}\right)^2 \frac{(k+1)^2 w^2}{2}$$

for each one of the regions $i = A,B$. Now, such a deviation induces the other region to revert to the sub-optimal taxation strategy forever and so it induces the inefficient outcome with utilities U_i^D for $i = A,B$ in all the remaining periods. Hence, it is not profitable to deviate if and only if:

$$\frac{U_i^{FB}}{1-\delta} \geq U_i^{Dev} + \delta \frac{U_i^D}{1-\delta} \text{ for } i = A,B \Leftrightarrow \delta \geq \delta_i^* \equiv \frac{U_i^{Dev} - U_i^{FC}}{U_i^{Dev} - U_i^D} \text{ for } i = A,B$$

11 For early discussions of such matters, see again Pantaleoni (1891).

Simple algebra establishes that:

$$\delta_A^* = \delta_B^* = \frac{1}{2}$$

hence, the first best allocation of resource can be achieved when the regions repeatedly decide on their local tax policy and care enough about the future and such a result is not affected by the extent of income inequality.

Transparency and redistribution once again

At the outset, we studied a model of decentralization with a central government in full control that uses only one policy instrument, a progressive income tax. We showed that redistribution is attained and that resource allocation is efficient. Efficient redistributional outcomes can be reached, more generally, by employing a linear combination of central government tax rate (θ_1) and interregional transfers (θ_2). In equilibrium $\theta_1 + \theta_2 = w(k-1)/2$. Depending upon central government choices on θ_1 and θ_2, different values of t_A and t_B are generated. If $\theta_1 = \theta_2$, each equals *w(k-1)/4* and $t_A = t_B = 1/2$. If $\theta_2 = 0$, the burden of equalization is put entirely on central government taxation which becomes $\theta_1 = w(k-1)/2$. If $\theta_1 = 0$ the burden of efficient equalization is put entirely on interregional transfer which becomes equal to $\theta_2 = w(k-1)/2$. This latter situation requires that regional governments adopt different tax rates from those they would choose in the two previous cases.

With $\theta_1 = 0$ the tax rate in the rich region would be higher than ½ and higher than in the case $\theta_2 = 0$. However, in the model, the three possible cases, implying the same allocation of resources, are welfare equivalent.

Abstracting from the model, one may try to evaluate the two polar situations from the point of view of the rich region: in one extreme case, taxes are low in the rich region and high in the poor region, with the central government raising positive taxes in the former and negative taxes in the latter. In the other extreme case, taxes are high in the rich region and low in the poor region, with high interregional transfers from the former to the latter. In the first case the Parliament of central government bears the burden of high taxes in the rich region; in the second case, the burden is on the Governor of the rich region.

Let us briefly consider a real world situation where politics and the public have a positive evaluation on a *quid-pro-quo* balance of regional taxes and regional public goods output. It is likely that the rich region will show stronger opposition to the interregional equalization program (that forces it to levy high regional taxes) than to the transfer of income via progressive taxation (that belongs to the powers of the central government).

In our model, the interregional equalization transfer (θ_2) can be thought of as an ideal representation of the so called "*horizontal equalization plans*" discussed in the literature on fiscal federalism (see Giarda, 2001). These plans require that the regions themselves decide on the extent of equalization by setting up the rules according to which money is going to be transferred from rich to poor regions.

In the model, the central government tax-transfer (θ_1) can be taken to represent the family of equalization plans that go under the heading of '*vertical equalization plans*', plans whereby the central government decides on the extent of equalization by raising taxes on rich region citizens to finance central transfers (grant programs) to poor regions.

Horizontal equalization plans are more transparent than vertical equalization plans. They are often considered to be better than vertical plans, because they impose a sort of peer control upon the way poor regions spend money that comes from interregional transfer programs. No-tax revenue is suspected to be used less efficiently than tax revenue. The incentive to use it efficiently is expected to be greater when the grant comes from the neighbour region than when it comes from a presumably distant central government. According to this view, transparency of the relationship between donor and recipient – via peer control – will improve technical efficiency, lower production costs of local public goods and produce more value for money in the execution of expenditure programs. At least according to this optimistic view, transparency promotes efficiency.

However, for the reasons outlined above, popular sentiments against equalization objectives are likely to be of bigger relevance under 'horizontal equalization plans' than under 'vertical equalization plans'. In the real world, it is rich regions that more loudly call for programs of interregional transfer to substitute for central government equalizing grants: they want to show that it is their generosity that makes solidarity rules working. It is rich regions that call attention to regional fiscal residua.

In intergovernmental fiscal relations, transparency (as it attained with horizontal equalization plans) may limit the scope of equalization as tax rates autonomously determined in rich region tend to be lower than optimal values as the donor region (a) will determine the size of the grant on unrealistic assumptions on the cost savings it expects the recipient region to generate and, (b) will resist levying the required taxes and not want to bear explicitly the full cost of equalization, though the same taxes might be imposed by the central government.

To our knowledge, no country involved in regional equalization plans relies entirely on horizontal equalization plans. Even where House of Regions exist (as in Germany), the Constitution provides strict rules for the concrete construction of equalization plans. *In practice, transparency and its offspring – horizontal equalization plans based on direct interregional transfers – have left the place to more obfuscated vertical equalization plans based on central government grants financed by (progressive) income taxes.*

Conclusions

This chapter has discussed the relationship between income redistribution and the organization of a system of intergovernmental fiscal relations. Here, we want to briefly mention the possible extensions of our simple model designed to tackle a variety of interesting questions.

The assumption of equal population in the two regions can easily be relaxed introducing a bias in the composition of public expenditure: indeed, in this case the

production of the public good becomes more convenient in the region with higher population, because of greater scale economies.

The assumption of equal preferences in the two regions can also be relaxed: such a further heterogeneity would create a new advantage for decentralized systems compared to the centralized ones as long as decentralization allows the regions to adapt their policies to their preferences.

When a national public good is considered, the choice between independence, centralization and decentralization becomes more sophisticated: independence would generate benefits associated to the complete decisional autonomy, but also a cost associated to the loss of scale economies in the production of the public good (since independence would require the production of two separate public goods in each region). Between the extreme assumptions of two local public goods and a national public good there is a set of intermediate situations in which each public good is regional but exerts spillovers on the other region. In this more general case (in part investigated by Alesina, Angeloni and Etro, 2005) a further coordination issue emerges because independence and decentralization imply that the regions take their choices without considering the externalities induced on the other region.

The assumption that all regions have equal opportunity to participate in the economy-wide budget making may be substituted by the assumption that one of the two regions is a first mover. A substantial leadership of certain regions is often realistic. In our model of constrained decentralization such an hypothesis would shift the burden of the inefficiency on the follower region inducing higher public expenditure and higher utility in the leader region and lower welfare for the country than in the case of regional parity. In decentralized settings, care must thus be taken to preserve regional parity in the decision process and in budget making.

Finally, the assumption of non-distortive taxation can be relaxed by introducing endogenous labour supply or exogenous costs of tax collection, so that a second best solution becomes the benchmark to compare with the third best outcomes available under constrained centralization and decentralization. Bordignon, Manasse and Tabellini (2001) have already explored issues of optimal regional redistribution in a similar context, but the approach presented in this chapter will probably allow a more general treatment.

Altogether, a case can be made for analysing institutional rules of fiscal federalist systems in stylised models such as the one presented in the chapter. Forcing legal provisions into simple analytical models may provide fresh insights onto old questions. Equally important is empirical research on the relationship between the structure of nations between level of governments vertically related and policy choices. In this chapter we have advanced the hypothesis of a trade-off between decentralization and redistribution. The robustness of such a relation and the sense of causality remain to be carefully checked.

References

Aghion, P., Alesina, A. and Trebbi, F. (2004), 'Endogenous Political Institutions', *Quarterly Journal of Economics* 119:2, 565-611.

Alesina, A. (2003), 'The Size of Countries: Does it Matter?', *Journal of the European Economic Association*, 301-16.

_____, Angeloni, I. and Etro, F., (2001), 'Institutional Rules for Federations', *National Bureau of Economic Research Working Paper No. 8646*.

_____, (2005), 'International Unions', *American Economic Review* 95:3, 602-615.

Besley, T. and Coate, S. (2003), 'Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach', *Journal of Public Economics* 87:12, 2611-37.

Bolton, P. and Roland, G. (1997), 'The Break up of Nations: A Political Economy Analysis', *Quarterly Journal of Economics* 112:4, 1057-90.

Bordignon, M. and Minelli, E. (2001), 'Rules Transparency and Political Accountability', *Journal of Public Economics.*

_____, Manasse P. and Tabellini, G. (2001), 'Optimal Regional Redistribution', *American Economic Review* 91:3, 709-23.

Giarda, P. (2001), 'Fiscal Federalism in the Italian Constitution', *Nation, Federalism and Democracy* (Bologna: Editrice Compositori), 105-118.

Meltzer, A. and Richard, S. (1981), 'A Rational Theory of the Size of Government', *Journal of Political Economy* 89:5, 914-927.

Oates, W. (1972), *Fiscal Federalism* (New York: Harcourt Brace Jovanovich).

_____, (1999), 'An Essay on Fiscal Federalism', *Journal of Economic Literature* 37:3, 1120-49.

Pantaleoni, M. (1891), 'Delle regioni d'Italia in ordine alla loro ricchezza ed al loro carico tributario', *Giornale degli Economisti*, Reprinted in Pantaleoni, M. (1938), *Studi di finanza e di statistica* (Bologna: Zanichelli).

Persson, T. and Tabellini, G. (1996a), 'Federal Fiscal Constitutions: Risk Sharing and Redistribution', *Journal of Political Economy* 104:5, 979-1009.

_____, (1996b), 'Federal Fiscal Constitutions: Risk Sharing and Moral Hazard', *Econometrica*, 64, 623-46.