Discussion Papers
Department of Economics
University of Copenhagen

No. 17-18

Fiscal Federalism, Grants, and the U.S. Fiscal Transformation in the 1930s

Martín Gonzalez-Eiras and Dirk Niepelt

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Fiscal Federalism, Grants, and the U.S. Fiscal Transformation in the 1930s*

Martín Gonzalez-Eiras† Dirk Niepelt‡

July 20, 2017

Abstract

We propose a theory of tax centralization and intergovernmental grants in politico-economic equilibrium. The cost of taxation differs across levels of government because voters internalize general equilibrium effects at the central but not at the local level. The equilibrium degree of tax centralization is determinate even if expenditure-related motives for centralization considered in the fiscal federalism literature are absent. If central and local spending are complements, intergovernmental grants are determinate as well. Our theory helps to explain the centralization of revenue, introduction of grants, and expansion of federal income taxation in the U.S. around the time of the New Deal. Quantitatively, the model can account for the postwar trend in federal grants, and a third of the dramatic increase in the size of the federal government in the 1930s.

JEL Classification: D72, E62, H41, H77

Keywords: Fiscal policy, Federalism, Politico-economic equilibrium, Markov equilibrium, Public goods, Grants, Political Economy

*For useful comments and discussions, we thank Asger Lau Andersen, Maximilian von Ehrlich, David Dreyer Lassen and seminar participants at CEMFI, IIBEO workshop, SED, Universidad Carlos III, and University of Copenhagen.

†University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: mge@alum.mit.edu

‡Study Center Gerzensee and University of Bern. P.O. Box 21, CH-3115 Gerzensee, Switzerland. E-mail: dirk.niepelt@szgerzensee.ch.

1 Introduction

Whether control over fiscal policy decisions should rest with national, regional or local governments depends on how effective they make use of their authority. A broad body of literature on fiscal federalism has emphasized that depending on the policy task at hand, the efficiency of policy choices may differ according to which level of government is in charge. It has concluded that some decisions are best taken de-centrally to minimize informational or other frictions which render it difficult to cater to heterogeneous needs, while others should be taken centrally to ensure that all important consequences of policy are internalized.

In deriving these results the literature has almost exclusively focused on government spending as the source of externalities or object of heterogeneous needs. That the efficiency of *tax collections* might also differ across levels of government has attracted much less attention although it is of equal relevance. By abstracting from differences in funding efficiency the fiscal federalism literature has abstracted from an important motive for decoupling revenue collection and spending across levels of government. In fact, it has typically ruled out such decoupling by assuming that governments individually balance their budgets.

In this paper, we take issue with the implicit assumption that the cost of taxation is identical across levels of government. We argue that there are reasons to expect the opposite and explore the implications. Most importantly, we show that if certain levels of government are in a better position to tax then this determines an equilibrium degree of centralization of tax collections. And if, in addition, other levels of government are in a better position to spend—for example for the reasons argued by the fiscal federalism literature—then this provides a straightforward explanation for the presence of inter governmental transfers or grants. The specific source of cost differences we focus on is inherently dynamic and its implications appear consistent with the data.[1]

The model features a central, or federal, government and many regional governments that impose labor income taxes to finance the provision of public services.[2] Taxation slows down capital accumulation and thus has general equilibrium effects: It drives up interest rates and lowers future wages which reduces the tax base in the future. Policy makers and voters at the federal level—rationally—internalize these general equilibrium

[1] Another obvious source relates to increasing returns in tax collections. Yet another one relates to negative externalities of taxation; see below for a discussion of the literature on tax competition.

[2] We refer to a state with a multi-tier political organization as a "federal" state, and to a government that makes decisions at the central level as a "federal" government. We refer to governments making decisions at the local level as "regional" governments.

effects to the extent that they are of relevance for them.[3] In contrast, policy makers and voters at the regional level—rationally—do not perceive general equilibrium effects of their decisions since regions are small relative to the nation and markets are not segmented. As a consequence, the net cost of a federal tax hike as perceived by a voter participating in national elections differs from the net cost of a regional tax hike as perceived in local elections.

Against this background, we answer the questions of which level of government taxes more or less, why inter governmental grants exist, and why they have risen in prominence in the U.S. since the 1930s. Our positive analysis in the context of a dynamic model of politico-economic equilibrium contrasts with the normative approach adopted in much of the fiscal federalism literature. While the latter typically identifies the welfare maximizing exclusive assignment of control to either the federal government or regional governments, we allow both levels of government to tax and spend and solve for the politico-economic equilibrium with grants in a standard macroeconomic framework.[4]

In the model, the quantity or quality of public services in a region depends on spending at the federal and regional level. We first consider a specification where federal and regional spending are perfect substitutes and traditional fiscal federalism considerations are absent: Government spending does not generate externalities and preferences for public services are uniform across the population. In such an environment, the equilibrium degree of centralization is indeterminate if the economy is static. But in the dynamic economy we consider, differential net costs of taxation render the equilibrium composition of tax collections across levels of government determinate.

To render grants determinate as well, we relax the assumption that regional and federal government spending are perfect substitutes. With complementarities, the federal government is handicapped in its ability to increase the provision of public services because regional spending is essential. This gives value to the ability to employ federal grants in order to increase regional spending. When tax revenue at the federal level is "cheap," grants allow to channel that revenue into the most productive use (regional, not only federal spending). Cross-regional externalities from public service provision as emphasized in the fiscal federalism literature have a similar effect in combination with the complementarities. When the externalities are positive, grants allow to increase regional spending above the level that regional governments would choose in the absence of such

[3]The welfare consequences for yet unborn cohorts who are not represented in the political process are not internalized.

[4]Of course, the mechanisms we emphasize would also be present in a model that adopts a normative perspective.

grants. In either case, the equilibrium degree of centralization of both taxes and spending is determinate and as a consequence, the size of inter governmental grants is uniquely determined as well. Moreover, grants crowd out local taxation, in line with empirical evidence.[5]

Our results are robust along several dimensions. We check whether the type of grant—uniform or matching—makes a difference; we find that for the most part, it does not. We consider the implications of labor mobility across regions and argue that introducing mobility does not fundamentally alter our findings. We allow for elastic labor supply, tax distortions, and additional policy instruments and find that the results are robust since the perceived cost differences due to general equilibrium effects working through capital accumulation, which are the key drivers of our results, are orthogonal to the effects of tax distortions. Finally, our results are robust to introducing policy instruments for intergenerational redistribution, such as public debt or social security.

More interestingly, we also consider the effects of capital income taxation. In contrast to labor income taxes which depress workers' savings, capital income taxes do not affect future capital accumulation because they are chosen ex post and reduce the income of the old. From the perspective of federal and regional voters, the net costs of taxation thus are the same. This implies that with spending complementarities, grants only are present when positive spending externalities outweigh deadweight losses associated with the grants.

We use the model to shed light on the transformation of the U.S. fiscal system around the time of the Great Depression and the New Deal that is illustrated in figure 1.[6] In the early 1930s the federal government accounted for about a third of total government revenues; virtually no inter governmental grants were present. Within a few years, this arrangement changed dramatically. The revenue share collected by the federal government doubled and inter governmental grants, which had been introduced to fund major New Deal programs, emerged as a central source of revenue for state and local governments.[7]

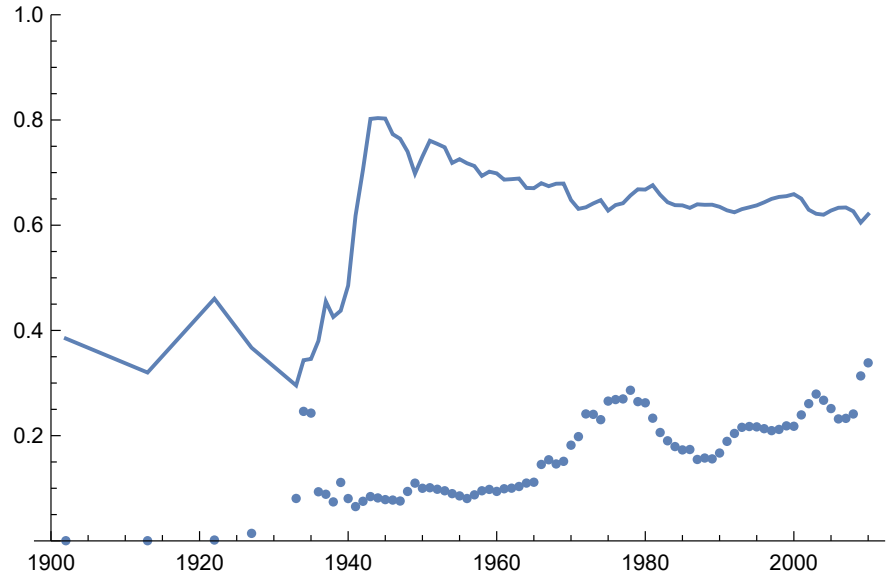
An even more dramatic transformation occurred with respect to the federal tax base, see figure 2. The income tax share of federal revenues increased from 28% in 1934 to 59% in 1940 and 84% in 1945, and the share of tax units who paid federal income taxes

[5]For example, Knight (2002) finds statistically and economically significant crowding out for the Federal Highway Aid Program in the U.S. He addresses identification problems (an omitted variable bias due to the positive correlation between grant levels and unobserved preferences for public spending) by using the political power of state congressional delegations as instruments.

[6]See Wallis (2000) for a discussion of this and earlier transformations.

[7]See Wallis and Oates (1998) for a description of New Deal programs and their impact on American federalism. Wallis and Oates (1998) discuss that the federal government ran large deficits during the 1930s; this is consistent with the fact that the rise of grants started before tax revenues were centralized.

Figure 1: Fiscal trends in the U.S.



Federal relative to total government revenues (solid), and federal grants relative to state and local revenues (dots). Sources: Wallis (2000) for years 1902, 1913, 1922, 1927; NIPA tables for subsequent years.

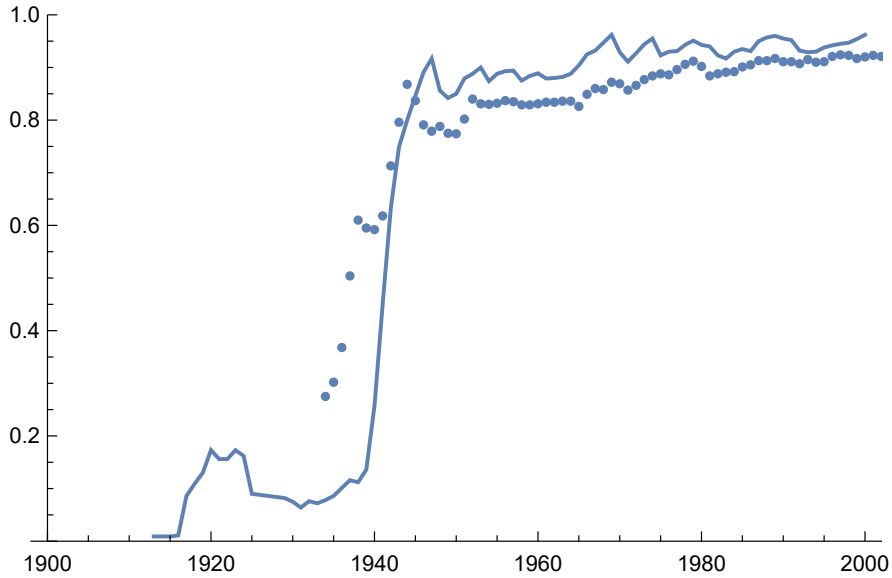
increased similarly rapidly, from only 7% in 1933 to 26% in 1940 and 85% in 1945.

The model rationalizes the centralization of tax revenue and the emergence of grants as the equilibrium response to a single exogenous event, the Sixteenth Constitutional Amendment that introduced the possibility for the federal government to tax income.[8] In an environment where voters at the local level were indifferent between (and thus, used) various revenue sources, voters in federal elections perceived the newly available federal labor income tax to have the cost advantage described earlier. As a consequence, revenue collection was not only centralized but income taxation gained prominence as a source of revenue for the federal government, as documented above.

The U.S. fiscal transformation did not occur immediately after the ratification of the Sixteenth Amendment in 1913 because administrative, legal and political hurdles had to be overcome. The Treasury underwent a major reorganization; the number of employees at the Bureau of Internal Revenue increased from roughly 4,000 in 1913 to more than 17,000 in 1921; the Supreme Court upheld the constitutionality of tax legislation enacted based on the Amendment (in 1916) and made a series of decisions relating to the proper

[8]The Sixteenth Amendment states: "The Congress shall have power to lay and collect taxes on incomes, from whatever source derived, without apportionment among the several States, and without regard to any census or enumeration."

Figure 2: Federal revenues and income taxation



Number of tax returns relative to number of tax units (solid), and federal income tax (including OASDI) relative to total revenue of the federal government (dots). Sources: Piketty and Saez (2003) Table A0, and Office of Management and Budget, Fiscal Year 2016, Historical Tables, Table 2.2.

definition of income and the fairness of its taxation (after the First World War), see Mehrotra (2013). All of this happened against the backdrop of political conflict as to who should pay income tax and how progressive the system should be.[9]

Our model does not feature within cohort conflict and therefore cannot account for the political struggle that accompanied the introduction of federal income taxation, and delayed its preponderance among revenue sources. But we argue in Section 6 that our explanation of the regime change dominates alternative possible explanations, including one in the tradition of the fiscal federalism literature that emphasizes a change in externalities on the spending side.

Calibrated to match the size of government, the model is able to explain the trend increase in grants in the postwar period as the result of increased urbanization, under the assumption that urban regions have a higher preference for public services than rural

[9]Instituted first under a Democratic administration, only less than 20% of the population was tax liable in the beginning. In 1921, a Republican administration gradually reduced the highest marginal tax rate from 77% in 1920 to 25% in 1925. After Democrats returned to power in 1933 the highest tax rate rose to 79% in 1937, without significant changes to the number of tax payers. The Second World War cemented the income tax as a mass tax. By 1945, 85% of the population paid income taxes, and both Republican and Democratic administrations have since kept the federal income tax system largely unchanged.

regions. Out of sample, the model predicts grants to continue to increase up to approximately 4.9% of GDP by 2060. Our analysis also establishes that without the possibility to exploit general equilibrium effects, the share of the federal government is roughly 10 percentage points smaller in the 1930s. The model thus explains a third of the observed dramatic increase in the relative size of the federal government in that period.

Related Literature We build on the classic analysis of fiscal federalism that features a trade-off between forces favoring centralization and decentralization. Oates (1972) emphasizes externalities in the provision of public goods on the one hand and heterogeneous preferences across regions on the other. He finds that absent spillovers and cost-savings from centralized provision, decentralization is preferable to uniform provision. But without information frictions, nothing prevents differentiated provision of services even in a centralized system (Oates, 1999).

Similar arguments are discussed in the theoretical political science literature (e.g., Kincaid, 2011). A federalist governance structure allowing for multiple centers of power is considered best suited for diverse countries, in particular if diversity is geographically based. Treisman (2007) critically discusses the rationales for and against political decentralization. He argues that administrative efficiency only requires administrative, not political decentralization and he questions the argument that local governments generally are better able or motivated to extract local information.[10] Our argument is related in so far as it stresses the possible decoupling of tax and spending decisions.

The literature subsequent to Oates (1972) has offered various explanations based on political economy frictions for the uniformity of centralized policy choices. For example, legislative bargaining among regional representatives at the federal level may imply reduced sensitivity of policy to regional needs (Lockwood, 2002); differentiated central service provision can give rise to costly bargaining and delay and may thus be avoided (Harstad, 2007);[11] credibility problems in signaling local tastes to the central government may generate inefficient federal policy choices (Kessler, 2014); and centralization may increase accountability but must be accompanied by policy uniformity because otherwise, the central government would implement policies favoring regions that monitor more extensively (Boffa, Piolatto and Ponzetto, 2016).[12] Alesina and Spolaore (1997) analyze the effect of international integration on the costs and benefits of centralization and thus, the

[10] Treisman (2007) suggests that the most convincing arguments for the relevance of decentralization are that it tends to increase policy stability and to lead to failures of fiscal coordination. See the discussion on tax competition below.

[11] See also Besley and Coate (2003).

[12] Related, Seabright (1996) argues that centralization limits the control rights of voters.

number of countries.[13]

Wallis (2000) documents that the U.S. passed through distinct regimes of government finance and suggests that federal, state and local governments may face differential costs to raise revenues from specific sources. Our model provides an explanation for such cost differences and it rationalizes the change of regime during the 1930s. The notion of differential costs of taxation due to the internalization or not of general equilibrium effects relates to Soares (2005) and Gonzalez-Eiras and Niepelt (2008) where the political support for education or social-security financing, respectively, depends on such effects as well.[14]

Our work also relates to the literature on tax competition (e.g., Gordon, 1983) which points out that uncoordinated regional taxation of mobile factors gives rise to revenue (and other) externalities across regions. A federal government concerned with welfare at the national level may correct these externalities by imposing federal taxes or transferring resources to regional governments through grants, among others. Our paper shares the focus on general equilibrium effects of taxation but its perspective is positive rather than normative. We emphasize that different perceptions about the cost of taxation at the federal and regional level give rise to a positive theory of fiscal federalism and may explain federal grants. Hatfield and Padró i Miquel (2012) study an economy where some public goods are funded and provided regionally and others federally. They show that the federal government imposes capital income taxes while regions resort to lump sum taxes, due to tax competition. In our setting, the grant instrument decouples funding from public good provision.

Uniform federal grants combined with non-uniform federal taxes (or vice versa) redistribute between regions and may constitute a form of inter-regional risk sharing (see, for example, Persson and Tabellini, 1996). The fact that such risk-sharing is very common does not provide a rationale for federal grants, however, since risk sharing in the joint interest of regions can be implemented without federal intervention. In our model, fiscal policy does not redistribute, and grants are used to achieve an allocation of resources that regions would not choose by themselves. Furthermore, since we are interested in explaining long run trends, risk sharing considerations are of second order, and the absence of fiscal equalization is a natural assumption as the U.S. federal government does not use grants for this purpose.[15]

[13] See also Bolton and Roland (1997).

[14] See also Kotlikoff and Rosenthal (1990).

[15] An exception is the joint financing of social insurance and welfare programs since the federal share of those costs increases as state income falls. See Gruber (2011, p. 266).

On the methodological side, our paper relates to the literature on dynamic politico-economic equilibrium (Krusell, Quadrini and Ríos-Rull, 1997). While most work in this literature studies equilibria with a single political decision maker Song, Storesletten and Zilibotti (2012) analyze politico-economic equilibrium in a setting with a continuum of governments that take factor prices as given. We solve a dynamic game with a continuum of regional governments and a central government that internalizes general equilibrium effects.

Outline The remainder of the paper is structured as follows. In section 2 we describe the model, and in section 3 we define equilibrium. Sections 4 and 5 contain the main analysis and extensions, respectively. In section 6, we contrast the model's implications with empirical evidence on fiscal trends in the U.S. and discuss its quantitative implications. Section 7 concludes.

2 The Model

2.1 Demographics and Institutions

We consider an economy inhabited by overlapping generations: workers and retirees. Workers supply labor, pay taxes, consume and save. In the subsequent period, they retire, consume the return on their savings, and die. The ratio of workers to retirees in period t equals ν_t.

The economy is composed of a continuum of regions of measure one over the unit interval. Each region is populated by a continuum of agents. The mass of agents and their age profile is identical across regions but the preferences of agents for publicly provided services may vary. Formally, regions are indexed by i and partitioned into J groups with groups indexed by j. All agents in all regions in group j share the parameter γ_t^j in their preference for publicly provided services.[16] The mass of regions in group j is given by θ_t^j, and in every period $\sum_{j=1}^{J} \theta_t^j = 1$. The demographic, preference parameters, as well as their cross-regional distribution follow deterministic processes.

Policy decisions are taken by governments at the federal and the regional level. Federal and regional governments act in the interest of voters participating in nationwide and regional elections, respectively. None of the governments can commit, and in each period

[16]With costless sorting there is no need to specify the initial allocation of households to regions. See Tiebout (1956).

they take decisions simultaneously.[17]

2.2 Production of Final Good

A continuum of competitive firms transforms capital and labor into output. Capital is owned by retirees—it corresponds to the savings of workers in the preceding period—and fully depreciates after a period. The economy-wide capital stock per worker, k_t, therefore corresponds to the economy-wide per-capita savings of workers in the previous period, s_{t-1}, normalized by ν_t. Labor is supplied inelastically (we show later the results are robust when labor supply is elastic). The gross interest rate R_t and the wage w_t are determined competitively.

We assume that the production function displays constant returns to scale such that factor prices in period t only depend on k_t,

$$R_t = R(k_t),\ w_t = w(k_t). \tag{1}$$

Moreover, we assume that the elasticities of the factor prices with respect to the capital-labor ratio are independent of the latter, $\epsilon_{Rk} \equiv \mathrm{d}\ln(R_t)/\mathrm{d}\ln(k_t) \perp k_t,\ \epsilon_{wk} \equiv \mathrm{d}\ln(w_t)/\mathrm{d}\ln(k_t) \perp k_t$. Examples of production functions that satisfy these assumptions include the Cobb-Douglas production function with capital share α where factor prices equal $R_t = \alpha k_t^{\alpha-1}$ and $w_t = (1-\alpha)k_t^{\alpha}$, the Ak production function, or a small open economy with exogenous factor prices.

The independence assumption can be disposed of at the cost of loosing the ability to derive closed-form solutions.

2.3 Production and Financing of Public Services

The quantity or quality of publicly provided services (or public services, for short) in a region i in group j, g_t^{ij}, depends on public spending at the regional level and nationwide. Let e_t^{ij} denote spending at the regional level and e_t the—uniform—spending by the federal government.[18]

We allow for positive or negative externalities across regions. Let e_t^j denote spending in

[17] In the data, this is not strictly true as state and federal elections of the executive and legislative branches are not perfectly synchronized. Our choice of timing assumption is motivated by our interest in the long run determinants of fiscal federalism, and the fact that one period in the model corresponds to several decades.

[18] Thus, we assume that both levels of government tax and spend. For rationalizations of policy uniformity at the federal level, see the literature review in the introduction.

a typical region in group j and let $\vec{e}_t \equiv (e_t^1, \ldots, e_t^J)$ denote the vector that collects regional spending across the J typical regions. (Throughout the paper, we use this notation for cross sections.) Publicly provided services in region i in group j are a function of $(e_t^{ij}, \vec{e}_t, e_t)$. We specify this function as

$$g_t^{ij} = a(e_t^{ij}, e_t) \times A(\vec{e}_t, e_t)^\lambda \quad \forall i, j. \tag{2}$$

The aggregator $a(\cdot)$ at the regional level and the cross-regional aggregator $A(\cdot)$ are increasing in all their arguments and the exponent λ measures the strength of the externality. In subsequent sections, we will adopt simple—and mutually consistent—functional form assumptions for $a(\cdot)$ and $A(\cdot)$. We will first consider the case without externalities, $\lambda = 0$, and with perfect substitutability between spending by the regional and federal governments, $a(e_t^{ij}, e_t) = e_t^{ij} + e_t$, before studying more general settings. In appendix A we provide one possible micro foundation for these aggregators based on constitutional restrictions that prescribe which services must be provided (but not necessarily financed) by regional or federal governments.[19]

Spending by the federal government is financed by a labor-income tax at rate τ_t and spending by region i in group j is financed by a tax at rate τ_t^{ij} as well as a uniform grant from the federal government, x_t. (See section 5 for a discussion of matching grants.) We allow for proportional deadweight losses of grants at rate $1-\sigma \geq 0$. Since all governments balance their budget in each period this implies

$$e_t = w_t(\tau_t - x_t), \quad e_t^{ij} = w_t(\tau_t^{ij} + \sigma x_t), \quad e_t^j = w_t(\tau_t^j + \sigma x_t) \quad \forall i, j, \tag{3}$$

where τ_t^j denotes the regional tax rate in a typical region in group j.

Tax rates and grants are non-negative.

2.4 Preferences and Household Choices

Workers and retirees in period t value private consumption, $c_{1,t}$ and $c_{2,t}$ respectively, as well as public services. They discount the future at factor $\beta \in (0, 1)$. For analytical tractability, we assume that period utility functions are logarithmic. Welfare of a worker

[19] This division could, for example, reflect externalities, spillovers, or the strength of tax-benefit linkages for local voters, as highlighted by Tiebout (1956). See also Hatfield and Padró i Miquel (2012).

in region i in group j who chooses savings s_t^{ij} is given by

$$\begin{aligned} & \ln(c_{1,t}^{ij}) + \gamma_t^j \ln(g_t^{ij}) + \beta \left(\ln(c_{2,t+1}^{ij'}) + \mathbb{E}_t[\gamma_{t+1}^{j'} \ln(g_{t+1}^{ij'})] \right) \\ \text{s.t.} \quad & c_{1,t}^{ij} = w_t(1 - \tau_t - \tau_t^{ij}) - s_t^{ij}, \quad c_{2,t+1}^{ij'} = s_t^{ij} R_{t+1}. \end{aligned}$$

The expectation accounts for the risk that region i of type j in period t turns into a region of type j' in period $t+1$.[20] To streamline notation, we define $\varphi_t^{ij} \equiv (1 - \tau_t - \tau_t^{ij})$ and correspondingly, $\varphi_t^j \equiv (1 - \tau_t - \tau_t^j)$.

Taking prices and taxes as given the worker optimally chooses

$$s_t^{ij} = \frac{\beta}{1+\beta} w_t \varphi_t^{ij}. \tag{4}$$

Conditional on prices, taxes and savings in the preceding period the welfare of a worker and a retiree, respectively, thus equal (dropping constants)

$$U_t^{ij,w} = (1+\beta)(\ln(w_t) + \ln(\varphi_t^{ij})) + \beta \ln(R_{t+1}) + \gamma_t^j \ln(g_t^{ij}) + \beta \mathbb{E}_t[\gamma_{t+1}^{j'} \ln(g_{t+1}^{ij'})], \tag{5}$$

$$U_t^{ij,r} = \ln(s_{t-1}^{ij'}) + \ln(R_t) + \gamma_t^j \ln(g_t^{ij}), \tag{6}$$

where we allow for the possibility that region i of type j in period t was of type j' in period $t-1$. Welfare of a worker and a retiree in a typical region in group j, $U_t^{j,w}$ and $U_t^{j,r}$ respectively, are defined accordingly.

2.5 Elections

Elections take place at the beginning of each period, simultaneously in all regions and nationwide. Workers and retirees may vote on candidates whose electoral platforms specify values for the policy instruments as well as other characteristics like "ideology" that are orthogonal to the fundamental policy dimensions of interest. These other characteristics are permanent and cannot be credibly altered in the course of electoral competition. Moreover, their valuation differs across voters (even if voters agree about the preferred policy platform) and is subject to random aggregate shocks, realized after candidates have chosen their platforms. This "probabilistic-voting" setup renders the probability of winning a voter's support a continuous function of the competing policy platforms. It implies that equilibrium policy platforms smoothly respond to changes in the demographic

[20] With costless sorting the expectation would account for the risk that a household of type j turns into type j' in period $t+1$.

structure and other fundamentals.

In the Nash equilibrium of the game with two competing candidates in a constituency choosing platforms to maximize their expected vote shares, both candidates propose the same policy platform.[21] This platform maximizes a convex combination of the objective functions of all groups of voters, where the weights reflect the groups' sizes and sensitivity of voting behavior to policy changes. Those groups that care the most about policy platforms rather than other candidate characteristics are the most likely to shift their support from one candidate to the other in response to small changes in the proposed platforms. In equilibrium, such groups of "swing voters" thus gain in political influence and tilt policy in their own favor. If all voters are equally responsive to changes in the policy platforms, electoral competition implements the utilitarian optimum with respect to voters. We assume that across groups of typical regions, voters are equally responsive to proposed changes in policy platforms. However, we allow for age related variation in responsiveness, reflected in a per capita political influence weight of unity for young voters and a per capita weight of $\omega \geq 0$ for retired voters.

3 Equilibrium

3.1 Competitive Equilibrium

We focus on symmetric equilibria where all regions within the same group behave identically, except possibly a set of regions of measure zero. The state is given by z_t, which includes the exogenous demographic and preference parameters as well as the endogenous state $\vec{s}_{t-1}$.[22] Conditional on z_t, the production function as well as competition among firms determine factor prices, w_t and R_t. A financing policy $(\vec{\tau}_t, \tau_t, x_t)$ (or policy for short) then determines public services, $\vec{g}_t$, capital accumulation, $\vec{s}_t$, and thus z_{t+1}. Proceeding recursively, a policy sequence $\{\vec{\tau}_s, \tau_s, x_s\}_{s \geq t}$ fully determines an allocation and price system.

Definition 1. A *competitive equilibrium* conditional on z_0 and a policy sequence $\{\vec{\tau}_t, \tau_t, x_t\}_{t \geq 0}$ is given by an allocation and price system such that

[21] See Lindbeck and Weibull (1987) and Persson and Tabellini (2000) for discussions of probabilistic voting.

[22] In general, the state also includes the level of assets in each individual region. Logarithmic preferences imply that the capital stock in an individual region does not affect the trade-offs faced by any political decision maker, see below.

i. capital evolves according to $k_t = s_{t-1}/\nu_t$, with $s_{t-1} \equiv \sum_{j=1}^{J} \theta_{t-1}^j s_{t-1}^j$, and factor prices are determined according to (1) for all t;

ii. the government budget constraints (2) and (3) are satisfied for all t; and

iii. households optimize: (4) is satisfied for all i, j, t.

3.2 Politico-Economic Equilibrium

In politico-economic equilibrium political decision makers optimally choose the values of the policy instruments under their control, taking all implications of their actions into account and forming rational expectations about future policy choices. We assume that these choices are Markov that is, they are functions of the fundamental state variables. We conjecture and later verify that policy choices are independent of the endogenous state variables, $\vec{s}_{t-1}$. Future policy choices therefore are unaffected by current policy choices.

Political decision makers at the regional and federal level perceive the economic environment differently. On the regional level they take policy choices by the federal government and in other regions, as well as factor prices and externalities, as given. On the federal level they take regional policy choices as given and account for the endogeneity of factor prices as well as externalities.

Formally, under the conjecture a regional decision maker in period t takes $(w_t, w_{t+1}, R_t, R_{t+1})$ as well as s_{t-1}^{ij} and $(\vec{\tau}_t, \tau_t, x_t, \tau_{t+1}^{ij}, \vec{\tau}_{t+1}, \tau_{t+1}, x_{t+1})$ as given and her objective is $\omega U_t^{ij,r}/\nu_t + U_t^{ij,w}$. Effectively, she maximizes

$$V_t^{ij} \equiv \left(\frac{\omega}{\nu_t} + 1\right) \gamma_t^j \ln(a(e_t^{ij}, e_t)) + (1+\beta) \ln(\varphi_t^{ij}) \quad \text{s.t.} \quad (3). \tag{7}$$

In contrast, the federal decision maker in period t takes (w_t, R_t) as well as $\vec{s}_{t-1}$ and $(\vec{\tau}_t, \vec{\tau}_{t+1}, \tau_{t+1}, x_{t+1})$ as given and she is concerned with the average of $\omega U_t^{j,r}/\nu_t + U_t^{j,w}$ across the J groups. Effectively, she maximizes

$$\begin{aligned} V_t \equiv & \sum_{j=1}^{J} \theta_t^j \left\{ \left(\frac{\omega}{\nu_t} + 1\right) \gamma_t^j \ln(g_t^j) + (1+\beta) \ln(\varphi_t^j) + \beta \ln(R_{t+1}) + \beta \mathbb{E}_t[\gamma_{t+1}^{j'} \ln(g_{t+1}^{j'})] \right\} \\ & \text{s.t.} \quad (1), (2), (3), (4), k_{t+1} = s_t/\nu_{t+1}. \end{aligned} \tag{8}$$

We can now define politico-economic equilibrium (under the conjecture).[23]

[23] In general, politico-economic equilibrium requires that political decision makers anticipate future policy choices to be determined according to policy functions (mappings from the state into policy)

Definition 2. A *politico-economic equilibrium* conditional on z_0 is given by a policy sequence $\{\vec{\tau}_t, \tau_t, x_t\}_{t\geq 0}$ and an allocation and price system such that

i. $\tau_t^{ij} \geq 0$ maximizes V_t^{ij} and $\tau_t^{ij} = \tau_t^j$ for all i, j, t;

ii. $(\tau_t, x_t) \geq 0$ maximizes V_t for all t; and

iii. the allocation and price system constitute a competitive equilibrium conditional on z_0 and $\{\vec{\tau}_t, \tau_t, x_t\}_{t\geq 0}$.

4 Analysis

4.1 Substitutability, No Traditional Fiscal Federalism Motives

To build intuition, we start with the case where spending by the federal and the regional governments are perfect substitutes, $a(e_t^{ij}, e_t) = e_t^{ij} + e_t$, and neither externalities nor heterogeneity are present, $\lambda = 0$ and $\gamma_t^j = \gamma_t \ \forall j$. In this case, (2) and (3) imply

$$g_t^{ij} = w_t(\tau_t^{ij} + \sigma x_t) + w_t(\tau_t - x_t) = w_t(\tau_t^{ij} + \tau_t + (\sigma - 1)x_t) \ \ \forall i, j.$$

Absent heterogeneity in regional preferences and without externalities across regions, none of the traditional fiscal federalism motives for decentralization or centralization is present. Nevertheless, the equilibrium degree of centralization of tax collection generally *is* determinate. To see this, consider the derivative of the regional objective function V_t^{ij} with respect to the regional tax rate, τ_t^{ij} (which equals τ_t^j in equilibrium), and the derivative of the federal objective function V_t with respect to the federal tax rate, τ_t. Since tax rates must be non-negative the derivative of V_t^{ij} in (7) and of V_t in (8) must be weakly negative in equilibrium,

$$\left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t}{\tau_t^j + \tau_t + (\sigma - 1)x_t} - \frac{1 + \beta}{\varphi_t^j} \leq 0 \ \ \forall j, \tag{9}$$

$$\left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t}{\tau_t^j + \tau_t + (\sigma - 1)x_t} - \frac{1 + \beta}{\varphi_t^j} + \mathcal{F}_t \leq 0, \tag{10}$$

respectively, where $\mathcal{F}_t \equiv -\frac{\beta}{\varphi_t^j}(\epsilon_{Rk} + \epsilon_{wk}\gamma_{t+1})$ denotes the factor price effect that is internalized at the federal level. In addition, the corresponding complementary slackness conditions must be satisfied.

and that optimal policy choices are consistent with policy functions evaluated at the state. Under the conjecture this consistency requirement is trivially satisfied.

The terms in the first inequality represent the marginal benefit and cost, respectively, of a higher regional tax rate as perceived by voters at the regional level. The marginal benefit derives from higher public services which both old and young voters appreciate, and the marginal cost reflects reduced wealth and thus, consumption of workers.

In the second inequality, the first two terms represent the marginal benefit of higher public services and the direct marginal cost of lower consumption as perceived by voters in nationwide elections. The marginal benefit and the direct marginal cost are the same as those perceived on the regional level because of the uniformity of preferences and the absence of externalities.

The third term in the second inequality, $\mathcal{F}_t$, represents the *indirect* net benefit of higher taxes that young voters at nationwide elections internalize. This net benefit materializes in the subsequent period (thus the discounting) and works through the tax induced reduction in savings in all regions (note that $\mathrm{d}\ln(s_t^j)/\mathrm{d}\tau_t = -1/\varphi_t^j$, see equation (4)). The benefit arises in the form of higher interest rates (reflected in ϵ_{Rk}, which is negative), and the cost in the form of a lower tax base to fund public services in the future (reflected in ϵ_{wk}, which is positive) weighted by the preference for public services in the subsequent period, γ_{t+1}.

A comparison of the two inequalities implies that the equilibrium degree of centralization of tax collection, and the amount of taxes that are collected both are *determinate* unless $\mathcal{F}_t = 0$. Since at least one of the tax rates τ_t^j and τ_t must be strictly positive in equilibrium (otherwise $g_t^j = 0$), at least one of the two first-order conditions must hold with equality. But $\mathcal{F}_t \neq 0$ implies that at most one first-order condition can hold with equality. It follows that either τ_t or τ_t^j equals zero. If $\mathcal{F}_t > 0$ then the first-order condition with respect to τ_t holds with equality, that is τ_t is interior and $\tau_t^j = 0$. If $\mathcal{F}_t < 0$, in contrast, the first-order condition with respect to τ_t^j holds with equality, that is τ_t^j is interior $\forall j$ and $\tau_t = 0$.

Importantly, this result holds although no traditional fiscal federalism motives are present. Determinacy results because voters at nationwide elections perceive different net benefits of taxation than voters in regional elections. For example, when lower savings drive up interest rates sufficiently strongly to render $\mathcal{F}_t > 0$, then the federal government levies taxes because voters at nationwide elections internalize that taxation improves their inter temporal terms of trade. In contrast, when lower savings depress next period's wages sufficiently strongly and the preference for public services in the subsequent period is sufficiently high to render $\mathcal{F}_t < 0$, then regional governments levy taxes because only voters at nationwide elections internalize the cost of taxation that results from lowering

next period's tax base. A binding commitment for regions not to raise taxes would improve voters' welfare in that case.

Turning to grants, the derivative of V_t in (8) with respect to x_t must be non-negative as well,

$$\left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t(\sigma-1)}{\tau_t^j+\tau_t+(\sigma-1)x_t}\leq 0,$$

and the corresponding complementary slackness condition must be satisfied. This implies that $x_t = 0$ when grants entail deadweight losses ($\sigma < 1$). If $\sigma = 1$, in contrast, the equilibrium level of grants is indeterminate since perfect substitutability of spending across levels of government then implies that inter governmental transfers do not affect the allocation.

We have characterized equilibrium policy. Note that we have verified our earlier conjecture that the policy functions are orthogonal to the endogenous state variables. Although the capital stock does not enter the first-order (and complementary slackness) conditions the trade-offs underlying the conditions are dynamic as they relate contemporaneous tax revenue and spending with future factor prices and revenue. The gain in tractability does not arise from suppressing this dynamic interaction, as in static models, but from specifying functional forms that render the factor price elasticities orthogonal to the capital stock. As shown elsewhere, in a related setting, different functional form assumptions (which render equilibrium policy a function of the capital stock) generate very similar numerical predictions for equilibrium outcomes.[24] We summarize these findings as our first main result:

Proposition 1. Consider the case with perfect substitutability and with no traditional fiscal federalism motives. Suppose that $\epsilon_{Rk}+\epsilon_{wk}\gamma_{t+1}\neq 0$ such that $\mathcal{F}_t \neq 0$. Then, in equilibrium, only one level of government levies taxes. In particular, for $\epsilon_{Rk}+\epsilon_{wk}\gamma_{t+1}<0$ (such that $\mathcal{F}_t > 0$) only the federal government levies taxes and for $\epsilon_{Rk}+\epsilon_{wk}\gamma_{t+1}>0$ (such that $\mathcal{F}_t < 0$) only the regional governments levy taxes. Grants equal zero unless $\sigma = 1$ in which case they are indeterminate.

In deriving proposition 1 we have assumed that labor is supplied inelastically. This assumption is *not* important for the results. To see this, suppose that households value leisure in addition to consumption and government services such that household prefer-

[24] In Gonzalez-Eiras and Niepelt (2005) we numerically solve for the equilibrium in a model with intergenerational transfers. We find that quantitatively, the numerical solution for equilibrium policy in the model version with CRRA preferences is very similar to the analytical solution in the version with logarithmic preferences.

ences are given by

$$\ln(c_{1,t}^{ij}) + v(l_t^{ij}) + \gamma_t^j \ln(g_t^{ij}) + \beta \left(\ln(c_{2,t+1}^{ij'}) + \mathbb{E}_t[\gamma_{t+1}^{j'} \ln(g_{t+1}^{ij'})] \right),$$

where l_t^{ij} and $v(\cdot)$ denote leisure and a smooth utility function, respectively.[25] The budget constraint of a worker now reads

$$c_{1,t}^{ij} = w_t(1 - l_t^{ij})(1 - \tau_t - \tau_t^{ij}) - s_t^{ij}.$$

It is easy to check that in this extended model labor supply does not respond to contemporaneous taxes, and proposition 1 therefore continues to hold without changes.

Maybe more interestingly, one may wonder whether in an environment with endogenous labor supply voters would employ additional distorting policy instruments to manipulate prices for their benefit. In appendix B, we analyze this in more detail. We consider an environment where voters at the federal and regional level may impose additional taxes whose proceeds are fully refunded to workers. These taxes therefore only serve to distort labor supply (which they do because the proceeds are refunded). At the regional level, voters do not benefit from creating such distortions. But at the federal level, where general equilibrium effects are internalized, the tax might be perceived to be valuable.

As we show in appendix B, introduction of these new instruments does *not* change the first-order conditions for τ_t or x_t, but adds a distortion term, $-\mathcal{X}_t^l \leq 0$ say, to the first-order condition for τ_t^j. The results of proposition 1 thus continue to hold subject to $\mathcal{F}_t + \mathcal{X}_t^l$ replacing $\mathcal{F}_t$: Taxation at the federal level constitutes an equilibrium outcome as long as $\mathcal{F}_t + \mathcal{X}_t^l > 0$. Intuitively, under the equilibrium choice of the new tax instrument at the federal level, the net benefit in general equilibrium from distorting labor supply equals zero. The choice of τ_t thus reflects the same considerations as in the model without elastic labor supply.

As another extension, consider a model where young households supply labor inelastically, but are mobile across regions. After voting, but before taking up work and being taxed, they may move at a utility cost. In a symmetric equilibrium, regional governments then still do not perceive general equilibrium price effects of their tax choices. But they do account for the fact that a marginal tax increase fosters emigration and reduces the tax base, driving up taxes for the remaining population in the region. Denoting by $\mathcal{X}_t^e$ the welfare cost of such emigration, results similar to those of proposition 1 follow, with

[25] We assume that $v(\cdot)$ is continuously differentiable, strictly increasing, concave and satisfies $\lim_{l \downarrow 0} v'(l) = \infty$.

taxation at the federal level an equilibrium outcome as long as $\mathcal{F}_t + \mathcal{X}_t^e > 0$.

Finally, it is useful to consider the implications of introducing government debt or social security. In our setup, voters at the federal level only internalize the general equilibrium effects that affect themselves; they disregard the income losses of future workers that go hand in hand with their own gains due to higher interest rates. One may therefore suspect that the availability of instruments for intergenerational redistribution—government debt or pay-as-you-go financed social security—could undermine our results.

To see that this is not the case, suppose that the federal government also levies a social security tax at rate η_t whose proceeds are distributed among retirees.[26] The first-order conditions that characterize public services provision, equations (9) and (10), then are unchanged except that the tax wedge now includes the new tax rate, $\varphi_t^j = 1 - \tau_t^j - \tau_t - \eta_t$. This might affect the magnitude of the general equilibrium term, $\mathcal{F}_t$, but not its sign.[27] The main message of proposition 1 therefore is robust: The level of government that funds public services is determined by the sign of $\mathcal{F}_t$, and grants are irrelevant.

Proposition 1 establishes that the equilibrium level of government that collects taxes in a federal state is determined even if none of the traditional fiscal federalism motives for decentralization or centralization of spending is present; in contrast, grants are irrelevant. Cross-regional externalities, or heterogeneous regional preferences for public services, would affect the incentives of the federal government to tax, but they would still not provide a rationale for grants.[28] We study next how complementarities between federal and regional spending do provide such a rationale.

4.2 Complementarity, Traditional Fiscal Federalism Motives

As is clear from the discussion leading to proposition 1 the source of the irrelevance of grants is the assumption of perfect substitutability. We now relax this assumption and introduce strict concavity in the aggregator function $a(\cdot)$, reflecting complementarities in government spending, or in the preferences for public services, in the presence of consti-

[26] Our setup satisfies the conditions for politico-economic equivalence (Gonzalez-Eiras and Niepelt, 2015, condition 4). This implies that absent commitment, the politico-economic equilibrium allocation in an environment with public debt and another one with pay-as-you-go financed social security are identical. We leave an extension with public debt issued by *both* levels of government for further work.

[27] The additional first-order condition determining the level of social security tax rate, η_t, is given by

$$\frac{\omega}{\nu_t} \frac{1}{\frac{\alpha}{1-\alpha} + \eta_t} - \frac{1+\beta}{\varphi_t^j} + \mathcal{F}_t = 0.$$

With intergenerational redistribution, the taxes levied to fund public services thus fall. Similarly, social security taxes are lower than in a model without public services.

[28] For an analysis, see the working paper (Gonzalez-Eiras and Niepelt, 2016).

tutional restrictions on service provision.[29] This provides a natural role for grants: When the federal government perceives a more favorable trade-off between taxation and spending than regional governments then it transfers resources to the latter, such that regional spending increases hand in hand with federal spending.

We also allow for heterogeneous regional preferences for public services, and for cross-regional externalities of government spending. In combination with the complementarities the latter also give a role for grants. If the federal government wants to raise spending because of positive spillover effects across regions then it can do so most efficiently (due to the complementarities) by transferring some of the additional funds to the regional governments.

For tractability, we assume that the aggregator function $a(\cdot)$ is of the Cobb-Douglas form with exponent δ on regional spending, $a(e_t^{ij}, e_t) = (e_t^{ij})^{\delta}(e_t)^{1-\delta}$. Given our micro foundation based on constitutional restrictions, δ is the fraction of goods that must be provided by regional governments. The conforming cross-regional aggregator $A(\vec{e}_t, e_t)$ is given by $A(\vec{e}_t, e_t) = \prod_{n=1}^{J} \left((e_t^n)^{\delta}(e_t)^{1-\delta}\right)^{\lambda\theta_t^n}$ and public services therefore equal

$$g_t^{ij} = w_t^{1+\lambda}(\tau_t^{ij} + \sigma x_t)^{\delta}(\tau_t - x_t)^{(1-\delta)(1+\lambda)} \prod_{n=1}^{J}(\tau_t^n + \sigma x_t)^{\delta\lambda\theta_t^n} \quad \forall i, j.$$

The first-order conditions with respect to a regional tax rate and the federal tax rate, respectively, now read

$$\left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t^j \delta}{\tau_t^j + \sigma x_t} - \frac{1+\beta}{\varphi_t^j} \leq 0 \ \forall j, \tag{11}$$

$$\sum_{j=1}^{J} \theta_t^j \left\{ \left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t^j (1-\delta)}{\tau_t - x_t} - \frac{1+\beta}{\varphi_t^j} \right\} + \mathcal{E}_t + \mathcal{F}_t \leq 0, \tag{12}$$

where $\mathcal{E}_t \equiv \lambda\bar{\gamma}_t(\omega/\nu_t + 1)(1-\delta)/(\tau_t - x_t)$ captures the marginal benefit from higher federal taxes due to cross-regional externalities. The factor price effect now equals $\mathcal{F}_t \equiv -\beta/\bar{\varphi}_t \left(\epsilon_{Rk} + \epsilon_{wk}(1+\lambda)\bar{\gamma}_{t+1}\right)$.

In addition to these first-order conditions and the complementary slackness conditions

[29] Spending complementarities could reflect informational frictions. If some governments can better provide certain public services than others then allocating the responsibility for spending to specific levels of government affects the total provision. For a critique of the notion of insurmountable information frictions across levels of government, see Treisman (2007). In appendix A, we provide micro foundations for complementarities that derive from constitutional restrictions.

the following first-order condition for grants must be satisfied in equilibrium:

$$\sigma\delta\sum_{j=1}^{J}\frac{\theta_t^j(\gamma_t^j+\lambda\bar{\gamma}_t)}{\tau_t^j+\sigma x_t}-\frac{(1-\delta)(1+\lambda)\bar{\gamma}_t}{\tau_t-x_t}\leq 0. \tag{13}$$

Again, the conjecture that policy functions are orthogonal to the endogenous state is verified. We have the following result:

Proposition 2. Consider the case with complementarity and traditional fiscal federalism motives.

(i) The federal government always levies taxes.

(ii) Let $\Omega_t \equiv (\omega/\nu_t+1)$ and $\Lambda_{t+1}\equiv(\epsilon_{Rk}+\epsilon_{wk}(1+\lambda)\bar{\gamma}_{t+1})=-\mathcal{F}_t\bar{\varphi}_t/\beta$. If

$$1+\beta+\delta\Omega_t\bar{\gamma}_t+\frac{\beta}{1+\beta}\Lambda_{t+1}\left(\sum_j\frac{\theta_t^j}{1+\beta+\delta\gamma_t^j\Omega_t}\right)^{-1}\geq$$
$$\sigma\left(1+\beta+\delta\Omega_t\bar{\gamma}_t+\lambda\bar{\gamma}_t\sum_j\frac{\theta_t^j(1+\beta+\delta\gamma_t^j\Omega_t)}{\gamma_t^j}\right), \tag{14}$$

then all regions levy taxes as well and grants generically equal zero.

(iii) If the opposite condition holds, then grants are strictly positive and fully crowd out taxes in regions with a low valuation of public services.

(iv) A mean preserving spread of the γ_t^j's reduces the set of parameters for which condition (14) holds, rendering grants more likely.

Proof. See appendix C. □

Intuitively, part (i) of the proposition follows from the fact that federal spending is necessary for $g_t^{ij}>0$ and grants are non-negative. The federal tax rate therefore must be positive in equilibrium. In contrast, the regional tax rates need not be positive unless grants equal zero.

To understand part (ii) note that with interior federal and regional tax rates the corresponding first-order conditions hold with equality. When combined with the equilibrium requirement that the marginal benefit of grants is weakly negative, this implies the parametric condition (14).[30] With homogeneous preferences, (14) reduces to

$$1+\frac{\beta}{1+\beta}\Lambda_{t+1}\geq\sigma(1+\lambda),$$

[30] In the non-generic case where the parametric condition is satisfied with equality, grants are indeterminate.

reflecting a trade-off between static and dynamic externalities as well as deadweight losses. Large deadweight losses (a small value for σ) make it more likely that the condition is met such that all regions levy strictly positive taxes while grants are absent. Positive externalities from government spending ($\lambda \geq 0$) or positive general equilibrium effects of taxation ($\mathcal{F}_t \geq 0$ and thus, $\Lambda_{t+1} \leq 0$) work in the opposite direction. They give the federal government an incentive to transfer resources to regional governments and as a consequence, crowd out taxation in regions with a low valuation of public services.

If the reverse of inequality (14) holds then some regional tax rates must be zero. But since regional spending also is necessary for $g_t^{ij} > 0$ grants must be strictly positive in this case, establishing part (iii).

Finally, concerning the size of grants and part (iv) of the proposition note that when preferences are homogeneous, only taxes, deadweight losses, and the importance of regional vs. federal spending (δ) directly affects the size of grants. With heterogeneous preferences, the *dispersion* of preferences for public services (but not their average level) and the strength of externalities (λ) affect x_t as well. From condition (13), by the implicit function theorem, we have $\partial x_t / \partial \lambda \geq 0$.

The results of proposition 2 extend to environments with labor mobility, elastic labor supply, or the inclusion of public debt or pay-as-you-go social security, for the same reasons as those discussed in section 4.1.

5 Extensions

5.1 Capital Income Taxes

It is instructive to consider the implications of the alternative assumption that governments resort to capital rather than labor income taxes. At the time when capital income taxes are decided upon and implemented, they only affect consumption of the old, but not savings of the young. As a consequence, the federal government perceives no equilibrium factor price effects, $\mathcal{F}_t = 0$.

This is reflected in the first-order condition for taxes. Moreover, the weight that governments place on the cost of taxation changes as well: $(1 + \beta)$ is replaced by $\frac{\omega}{\nu_t}$. Otherwise, the first-order conditions for taxes remain unchanged (with φ_t^j now denoting the tax wedge due to capital income taxes).[31] The first-order condition for grants is

[31] Although tax bases of labor and capital income taxes are different, implying different levels of spending for a given tax rate, voters face a similar trade-off between the marginal costs and benefits of taxation since preferences are logarithmic.

unaffected since general equilibrium effects do not enter it.

When public spending is perfectly substitutable across levels of government, spending externalities are absent, and preferences for public services are homogenous across regions, then total taxes are determinate in equilibrium but the degree of centralization is not. Grants are indeterminate unless they are wasteful, in which case $x_t = 0$. This follows immediately from a variant of proposition 1.

With complementarity, in contrast, a variant of proposition 2 implies that the federal government always levies taxes as do a subset of regional governments. The parametric condition (14) applies with $\Lambda_{t+1} = \epsilon_{Rk} = \epsilon_{wk} = 0$.

5.2 Matching Grants

In contrast to uniform, or block, grants, federal governments sometimes use matching grants that provide local governments with resources in proportion to what they themselves spend on a particular outlay. We consider the case in which matching grants are provided at the same rate on all expenditures, such that if local government i in a region of type j spends $w_t\tau_t^{ij}$, the federal government transfers $x_t w_t \tau_t^{ij}$ and spending is given by

$$e_t = w_t(\tau_t - x_t\bar{\tau}_t), \quad e_t^{ij} = w_t\tau_t^{ij}(1+\sigma x_t), \quad e_t^j = w_t\tau_t^j(1+\sigma x_t) \quad \forall i,j. \tag{15}$$

In Appendix D, we prove the following result:

Proposition 3. With matching grants, perfect substitutability and no traditional fiscal federalism motives, the results of proposition 1 apply. With complementarity, all tax rates are positive; if

$$\sigma(1+\lambda)\bar{\gamma}_t < \sum_j \frac{\theta_t^j \gamma_t^j}{1+\beta+\Omega_t\delta\gamma_t^j}\left(1+\beta+\delta\Omega_t\bar{\gamma}_t + \frac{\beta}{1+\beta}\frac{\Lambda_{t+1}}{\sum_j \frac{\theta_t^j}{1+\beta+\Omega_t\delta\gamma_t^j}}\right), \tag{16}$$

then grants are not used.

We conclude from proposition 3 that the predictions of the model with uniform grants and with matching grants are qualitatively the same. Note that conditions (14) and (16) are identical when preferences are homogeneous.

6 Model Implications and Empirical Evidence

6.1 The U.S. Fiscal Transformation in the 1930s

Wallis (2000) documents that the United States passed through three different eras of government finance. From 1790 until about 1842 state governments were the most active and financed themselves through asset income, primarily from tolls on canals, dividends from bank stock, and revenue from land sales. By 1842 several states were in default on their debts.

To meet this crisis, state governments resorted to property taxes and retreated from infrastructure investments. This was met with an increase in importance of local governments that also mainly relied on property taxes for their funding. By the eve of the Great Depression, local governments collected over half of the tax revenues raised by all levels of government and property taxes accounted for 42% of revenues at all levels. During these first two eras, the federal government's main source of revenue were tariffs, and on a smaller scale, property taxes.

The Great Depression and New Deal marked the birth of the current fiscal system, see figures 1 and 2 in the Introduction. On the revenue side this system is characterized by the reliance on income taxes at the federal level.[32] While in 1934, income taxes accounted for less than 28% of federal government revenues, this share rose to 59% in 1940 and to 84% by 1945. In parallel, grants from the federal to state or local governments grew from negligible levels before 1933 to 8% of state and local revenues in 1940.

The model offers an explanation for the fiscal transition in the U.S. around the time of the Great Depression and New Deal. This explanation for the shift from tariffs and property taxes to income taxation, alongside an increase in taxation by the federal government and the introduction of federal grants, emphasizes the ratification of the Sixteenth Constitutional Amendment in 1913. The Amendment opened the way for a form of federal taxation—income taxes collected from workers—with stronger general equilibrium effects.[33] As a consequence, federal taxation and grants became politically more attractive and thus, more prevalent.

The shift did not occur immediately after the Amendment because within-cohort po-

[32]Income taxes include individual, corporate, and payroll taxes. States also levy sales taxes.

[33]While income taxes are collected from workers and retirees (and from corporations, but the share of revenue collected from the latter has never exceeded 15% since 1978), the fraction of individuals paying taxes falls rapidly with age, see for example Greenstone and Looney (2012). Applying Piketty and Saez's (2003) methodology to classify respondents we find that in the 2015 March Current Population Survey 87% of tax units aged 65 or below paid taxes, roughly twice the share of those aged 65 years or older.

litical conflicts (which are not present in our setup) had to be resolved.[34] Policy debates on income taxation, both before and after the ratification of the Sixteenth Amendment, reflected a strong ideological clash between progressives who pushed the idea that those with greater ability to pay had a moral obligation to do so, and conservatives who criticized income taxes because they punished thrift and were meant to "soak the rich."[35] In the years between World War I and the mid 1930s, marginal tax rates were repeatedly raised and lowered again (fluctuating between 25% and 79%), by varying majorities in Congress. While until the eve of World War II the fraction of the population that paid income taxes always lay below 20%, it increased to 85% by the end of the war, signalling the transformation of a class tax into a mass one.[36] Both Republican and Democratic administrations have since kept the federal tax system largely unchanged.

Of course, local, and state governments in the U.S. do not exclusively rely on labor income taxes; in fact, some heavily rely on property and sales taxes. However, the crucial step in our proposed explanation—that regional governments do not internalize the general equilibrium effects of their taxes—is unaffected by the exact type of tax they levy. With a property or sales tax, a regional government still would not internalize such effects.[37]

The explanation of the U.S. experience offered by the model is attractive not least because of its simplicity—it assumes a single, permanent change in the policy environment. Competing explanations require stronger assumptions. For example, Wallis (2000) suggests that the increase in federal relative to state and local receipts might reflect a reduction of tax collection costs for the federal government. Indeed, the introduction of Social Security payroll taxes might have given federal authorities an information advantage that could in turn have led to lower costs for it to raise income taxes. But an argument along those lines has difficulties explaining why the drastic changes were not undone later in time. After all, with rapid advances in information technology throughout the second half of the twentieth century, an information advantage for federal authorities is unlikely to have persisted over decades.

An explanation based on traditional fiscal federalism motives similarly has problems

[34] As the first-order conditions (11) and (12) make clear, regional taxes only adjust once federal taxes change. The model thus predicts a simultaneous change of taxes at the regional and federal level, once the adjustment hurdles at the federal level are overcome, in line with the evidence presented in figures 1 and 2.

[35] See Mehrotra (2013).

[36] Vélez (2014) argues that across a sample of OECD countries, progressive income taxation was instituted as a response to war effort and not due to changes in income inequality.

[37] Indeed, as we showed in subsection 5.1, *no* government would internalize the general equilibrium effects of a capital income tax.

matching the data. Such an explanation based on positive externalities of public services that call for spending by the federal government would have to argue that these externalities permanently increased around the 1930s. Maybe the most plausible candidate in that respect would be public infrastructure investment to support the major technological innovations that transformed production. But these innovations (in particular, electric light and the internal combustion engine) already occurred at the end of the nineteenth century (Gordon, 2012). And by the time of the Great Depression, most of the infrastructure investments based on them were already undertaken, at least in urban areas.[38] The shifts in the fiscal landscape thus should have been observed earlier. Moreover, even if spending externalities had increased around that time, federal spending should have spiked rather than permanently increased since the higher externalities would have triggered a federal public investment boom followed by more moderate maintenance spending. This is not what we see in the data.

6.2 Quantitative Assessment

Figures 1 and 2 document two key aspects of the U.S. fiscal transformation: First, a dramatic increase in the share of the federal government in the 1930s. And second, after World War II, a trend increase in the use of grants which grew from around 0.6% of GDP in 1950 to 2.8% in 2014. In this subsection, we show that the model is able to quantitatively replicate a substantial part of the former, dramatic increase and most of the latter trend.

We do this in two steps. First, we focus on the time after World War II, calibrate the model to match the size of the regional and federal governments in that period, and analyze drivers of federal grants. Second, we verify that the calibrated model predicts a much smaller size of the federal government (in the period before World War II) if the federal government is prevented from exploiting the general equilibrium price effects of labor income taxation.

Our quantitative analysis is based on the model with spending complementarities. We assume that one period in the model corresponds to 30 years in the data, posit a Cobb-Douglas production function for the final good, and use the following parameter values: Based on findings in Piketty and Saez (2003) we let the capital share in the production function be 0.2815. We set ν_t to the 30-year gross U.S. population growth rate and use Census Bureau data. From Gonzalez-Eiras and Niepelt (2008) we take $\omega = 0.9176$. In

[38] By 1952 almost all U.S. farms had electricity, and by 1970 the main network of interstate highways was in place, see Gordon (2012).

the benchmark calibration we assume no externalities ($\lambda = 0$) and 7.5% deadweight losses ($\sigma = 0.925$). For the remaining parameters, we rely on moment conditions of the model as described in more detail below.

From proposition 2, time variation in several model variables, including deadweight losses, externalities, the importance of federal vs. regional spending, and preferences for public spending, could in principle explain the observed increase of grants. Since there is little tangible evidence of systematic variation of the former three factors we focus on the role of preferences. (Recall that preference *heterogeneity* but not the average preference for public services affects the trade-off governing the choice of grants.) Specifically, we assume that the shares of regions with a high or low valuation of public services changed over time.

We assume that there are two types of regions, "urban" and "rural," and we proxy the share of urban regions, θ_t^1, with the average urbanization rate as reported by the Census Bureau. The motivation to distinguish regions by urban vs. rural character is twofold. On the one hand, the distinction seems relevant for observed patterns of political support. For example, Frank (2004) argues that low-income Americans living in rural areas vote strongly Republican even though the Republican party's economic platform cuts against their economic interests. We interpret this behavior as reflecting a lower preference for government spending in rural areas.[39] On the other hand, the distinction also seems to be borne out by survey evidence. Data on attitudes towards public spending collected by the General Social Survey in the years 1985, 1990, 1996, and 2006 indicates that respondents in rural areas supported government spending to a lesser extent than respondents in urban areas, see table 1.[40]

The urban versus rural distinction also is consistent with indirect evidence that blends data on state level spending and an implication of the model. Recall that the model (with uniform grants) predicts regions with lower valuations for public services to have a higher ratio of grants relative to regional tax revenues. If urbanization is positively correlated with the valuation of public services it should be negatively correlated with that ratio. We check this prediction using U.S. state level data for 1969 and 2008 in a panel regression of the ratio of federal grants to direct general revenue in state and local governments on

[39]Other observers have argued that voters care more about moral than economic issues. See Ansolabehere, Rodden and Snyder Jr. (2006) for a critical discussion of the "culture war" interpretation of these voting patterns.

[40]The survey is conducted yearly by The National Data Program for the Social Sciences. Respondents in the years 1985, 1990, 1996, and 2006 were asked if they favored or not cuts in government spending. We take the fraction of those answering "strongly in favor of" and "in favor of" as a measure of the intensity of preferences for lower spending.

Table 1: Attitudes towards government spending cuts

	1985		1990		1996		2006	
	Obs.	% in favor	Obs.	% in favor	Obs.	% in favor	Obs.	% in favor
Total	666	0.820	1182	0.782	1293	0.834	1483	0.633
Urban	540	0.807	1014	0.775	1163	0.831	1293	0.627
Rural	126	0.873	168	0.821	130	0.861	190	0.674

The table shows the fraction of respondents answering "strongly in favor of" and "in favor of" spending cuts. Data from the General Social Survey. Counties having no towns of 10,000 or more inhabitants are classified as rural.

urbanization (controlling for state income per capita), see table 2. As expected, we find a negative relationship.[41] Table 2 also reports the results of a regression of grants per capita on urbanization. The negative coefficient indicates that grants are higher rather than lower in rural areas, providing support for our modelling choice of uniform rather than matching grants.[42]

To calibrate the preference for public services as well as β and δ, we use the first-order conditions for τ_t and τ_t^1 in 1950 and for τ_t, τ_t^1 and x_t in 2000. Specifically, we match the GDP-share of grants, the size of government (federal, state, and local spending relative to GDP), and the size of the federal government in the year 2000, as well as the size of government in the year 1950.[43] We assume that grants are used in 1950 and 2000. An implication of our model is then that in those two periods regions with low valuation—rural regions—do not tax. (This could be modified by introducing an exogenous component of regional government spending.) We allow preferences for public

[41]We use 2008 data to avoid problems with the Great Recession, and we use data for the year 1969 rather than 1970 since the table in Dales (1971) appears to contain a typo in the entry for Colorado. We exclude the District of Columbia as its urbanization is 100% in both periods.

[42]Strictly speaking, uniform grants should imply that the coefficient equals zero. The partitioning of regions into either rural or urban might be too coarse to capture all relevant determinants of grants. In any case, the assumption of uniformity is easier to reconcile with the data than the assumption of matching grants, which would imply a positive coefficient.

[43]Data comes from the National Income and Product Accounts (NIPA) from the Bureau of Economic Analysis. In the model there is no public debt, thus revenues and expenditure are equivalent (except for deadweight losses) measures of the size of government. To account for the use of debt we take the average of current revenues and current expenditures as our measure of the size of government. For state and local governments we subtract federal grants.

Table 2: Urbanization and grants

	Grants to local revenue	Grants per capita
Urbanization	−0.309	−2.66
	(0.183)	(0.986)
Income per capita	0.295	2.44
	(0.247)	(0.704)
State FE	YES	YES
Time FE	YES	YES
R^2	0.726	0.899
Observations	100	100

Note: The table shows panel OLS regressions with two measure of grants as outcomes: grants to state and local revenue, and grants per capita. The explanatory variables used are state-level urbanization and state income per capita relative to national. Sources: The ratio of federal grants to state and local direct general revenue, and grants per capita for 1969 are taken from Dales (1970); grants for 2008 from the Census Bureau's Consolidated Federal Funds Report for Fiscal Year 2008, Table 4 (www.census.gov/prod/2009pubs/cffr-08.pdf); and state and local government finances for 2008 from the Census Bureau (www.census.gov/govs/local/historical_data_2008.html). State income per capita relative to national is taken from the Bureau of Economic Analysis (www.bea.gov/itable). Population and urbanization data comes from the Census Bureau (www.census.gov). Robust standard errors are in parenthesis.

Table 3: Calibration

β	δ	γ^1_{2000}	γ^2_{2000}	$\gamma^j_{t+1}/\gamma^j_t$
0.6133	0.4830	0.8032	0.0223	1.1779

services to change over time at a constant and equal rate. In addition, we use a moment condition for the Euler equation in steady state (see Gonzalez-Eiras and Niepelt, 2008).[44]

Table 3 lists the calibrated parameters. The β value corresponds to an annual discount factor of approximately 0.984. The calibration for δ suggests an almost equal role for federal and regional spending in the provision of public services. To meet the requirement that rural areas do not levy taxes they must have a low preference for the public service, of approximately 2.8% of the corresponding parameter value for urban areas.[45] The ratio of spending in urban relative to rural areas equals 6.95. In the data, spending disparities at the state level extend up to roughly 2.5[46] with disparities at lower levels likely to be higher.[47]

To replicate the increase in the size of governments between 1950 and 2000, the model requires the preference for public services to grow at about 0.55% per year. This is qualitatively consistent with Wagner's law and with the evolution over time of attitudes towards spending cuts, as reported in table 1.[48]

We are now in a position to assess the quantitative performance of the model. Concerning the first step of this assessment, figure 3 illustrates that the model does a good job

[44]We impose the 30-year gross interest rate $R = 2.443$. See Gonzalez-Eiras and Niepelt (2008) for details.

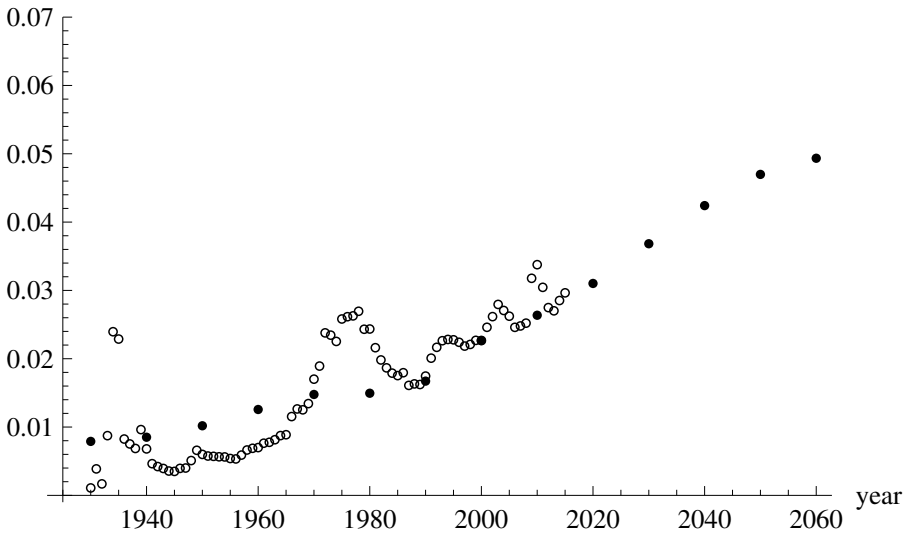
[45]This feature is robust to assuming matching grants. A calibration under that assumption yields very similar parameters, with the preference parameter in rural areas equalling approximately 5.5% of the corresponding value in urban areas. The simulated data based on this calibration has grants growing at a significantly lower rate than we see in the data.

[46]Data comes from the Census Bureau.

[47]If we impose that regions must exogenously collect a tax of 5% for other uses, then the ratio of spending in urban relative to rural areas is 2.5, and we find similar predictions for the trend increase in grants than in our baseline calibration. Grants are predicted to peak at 5.3% of GDP at the end of the simulation horizon.

[48]By construction, the model perfectly matches the trend growth across all levels of government. It produces a flatter trend for the ratio of federal to total government than in the data with a predicted value of 63.6% for 1950 instead of the actual 72.6%. An alternative explanation for the rising size of governments could rely on public services being a luxury good, and higher incomes. Our assumption of logarithmic preferences rules out income effects on tax rates.

Figure 3: Federal grants, share of GDP



Data from NIPA (circles), model predictions (dots).

at explaining the long-term evolution of grants. At the same time, the model is not able to match the observed short-run fluctuations in grant provision (notably during the periods of the oil shocks in the 1970s and the Great Recession). This suggests, not surprisingly, that grants are also used for redistributive and risk sharing purposes which our model does not speak to. The increase in grants since the 1930s (which the model correctly predicts) reflects rising urbanization and thus, preference heterogeneity—not stronger average preferences for public services, which explain the growing size of governments but not grants. Furthermore, when we simulate the model holding θ^1 constant at its year 2000 value it predicts a slightly negative trend for grants.

Out of sample, the model predicts that grants continue to increase up to approximately 4.9% of GDP by 2060. The predictions are robust to changes in all parameters except λ and σ. When externalities are assumed to be negative, say -3%, or deadweight losses higher, say 10%, grants peak at 2.8-2.9% of GDP between 2040 and 2050 and then revert back to lower values. When spending externalities are positive, say 2% or deadweight losses lower, say 6%, grants are predicted to increase to 9.3-10% of GDP in 2060. The intuition for these results is straightforward: with lower deadweight losses or higher externalities, the federal government has a stronger incentive to provide grants (see proposition 2).

Concerning the second step of our quantitative assessment, we verify that the model

indeed predicts a much smaller size of the federal government in the period before World War II, if the government is prevented from exploiting the general equilibrium effects of labor income taxation. Subject to the calibrated parameter values listed in table 3 the predicted relative share of the federal government is roughly 10 percentage points smaller in 1930 than otherwise if we let $\mathcal{F}_t = 0$. The model thus accounts for roughly one third of the dramatic increase in the size of the federal government during the 1930s.[49]

Next, we compute a second backward out-of-sample forecast. We establish that according to the model, grants would not have been employed in the year 1910 given that mostly property taxes were in place and thus, general equilibrium price effects were minor. Specifically, we solve for politico-economic equilibrium in an economy with capital income taxes and show that in this economy, in the year 1910, grants would have been absent, in line with the data.[50]

Finally, we conduct two counterfactual analyses. First, we compute the choice of fiscal instruments in the year 2000 under the assumption that political decision makers at the federal level do not perceive the general equilibrium effects of labor income taxation. We find that in this case, the federal government does not use grants and federal taxes are more than 6 percentage points lower, partially compensated by regional taxes that are more than 3 percentage points higher. These changes result in an 11.5 percentage points lower share of federal government revenue.

Second, we re-calibrate the model assuming that the federal government does not perceive the general equilibrium effects of labor income taxation. Under this assumption, the model requires large static externalities to match the moments described earlier. Imposing the δ value reported in table 3, the newly calibrated λ value is very large (0.3029).[51] We consider this value implausible and interpret the result as further evidence that our explanation for the centralization of revenue and the use of inter governmental grants is more credible than an alternative based on spending externalities.

[49] Assuming $\lambda = 0.04$ and $\sigma = 0.8$ the model can explain almost a half of the change in relative governments' sizes, but at the cost of having grants almost flat in the postwar period. Larger externalities increase the response of federal taxes, and larger dead-weight losses are needed for grants not to be used in the absence of general equilibrium effects.

[50] One may argue that the Sixteenth Amendment indirectly also affected the division of tasks between levels of government on the spending side. Accordingly, the enactment of the Amendment in 1913 might have lowered δ relative to its value in 1910. When we re-run the backward out-of-sample forecast under the assumption of a larger spending share for regional governments in 1910 ($\delta = 0.65$ rather than 0.48) we find the same result: Grants are still not used.

[51] This alternative calibration implies a negative value for γ^2_{2000}. Eliminating the moment condition for grants in the year 2000 and imposing for γ^2_{2000} the value reported in table 3 results in $\lambda = 0.1246$ and grants in 2000 a third higher than in the data.

7 Concluding Remarks

What determines the degree of centralization of tax collections in a federal union? We have argued that differences in the perceived cost of taxation across levels of government constitute a candidate explanation. While such differences may arise from various sources we have emphasized one that is inherently dynamic, relating to the fact that income taxation at the national level induces general equilibrium effects on interest rates and wages. We have also argued that cost differences of taxation in combination with complementarities of government spending across levels of government, and possibly traditional fiscal federalism motives for the centralization or not of government spending, provide a natural motive for inter governmental grants.

Cost differences of taxation have the potential to explain secular changes in the U.S. fiscal landscape since the 1930s towards more centralized revenue collection, more widespread use of grants, and increased reliance on income taxation. In our framework these changes result in response to a single shock—the ratification of the Sixteenth Amendment—which opened the way for labor income taxation at the federal level. Alternative explanations of the U.S. experience, based on an information advantage of federal authorities or traditional fiscal federalism motives on the spending side, do not offer comparable explanatory power.

Our simple framework abstracts from cross-regional insurance, redistribution, and many other features that are present in federalist states. Given this simplicity, the predictive power of the model when calibrated to match U.S. data is reassuring. The trend increase in urbanization can account for the increase in federal grants during the post-World-War II period, and factor price effects can account for a third of the dramatic increase in the relative size of the federal government during the 1930s.

Three extensions of the model presented in this paper appear to be of particular interest. First, the setup could be enriched to admit productivity differences across regions, generating a role for cross-regional insurance and redistribution. Such an extension could be useful to study the determinants of redistributive federal grants and the consequences of cross-regional inequality, in the post-World-War II U.S. but also, for example, in the context of European integration.

Second, the option to issue government debt for tax smoothing or tax burden shifting purposes could be introduced both at the federal and the regional level. Governments would hold conflicting views about the costs and benefits of public debt since regional policymakers would not internalize the general equilibrium effects of deficits on factor prices. As a consequence, the federal government might opt to employ grants (and deficits)

to influence both regional taxes and deficits. This model extension could address questions regarding debt and deficit policies in federal states.

Finally, the setup could be used to study the constitutional assignment of spending responsibilities. We treated the Sixteenth Amendment as an exogenous event and took constitutional restrictions regarding the division of tasks across levels of government (as reflected by the parameters in the aggregator functions) as given. Both the division of responsibilities and their reform could be endogenized with the aim to better understand why at the constitutional level, certain tasks are left in the realm of federal policy makers while others are assigned to regional governments.

A Derivation of Aggregator Functions

The functional form assumptions we adopt are special cases of the specification

$$g_t^{ij} = \left(\int_0^{\delta} \left(\frac{e_t^{ij}(l)}{\delta} \right)^{\kappa} \mathrm{d}l + \int_{\delta}^{1} \left(\frac{e_t(l)}{1-\delta} \right)^{\kappa} \mathrm{d}l \right)^{\frac{1}{\kappa}} \times \left(\sum_{n=1}^{J} \theta_t^n \left(\int_0^{\delta} \left(\frac{e_t^n(l)}{\delta} \right)^{\kappa} \mathrm{d}l + \int_{\delta}^{1} \left(\frac{e_t(l)}{1-\delta} \right)^{\kappa} \mathrm{d}l \right) \right)^{\frac{\lambda}{\kappa}},$$

where $\kappa \equiv (\eta - 1)/\eta$ and $\eta \geq 0$. The interpretation is as follows: The public service is an aggregate that reflects spendings at the federal and the regional level on a continuum of goods indexed by $l \in [0, 1]$. We assume a constant elasticity of substitution, $\eta \geq 0$, between types of goods (reflecting voters' preferences or technology in the production of the public service). The constitution prescribes that goods with index $l \in [0, \delta]$ must be provided (but not necessarily financed) by the regional government while goods with index $l \in (\delta, 1]$ must be provided by the federal government where $0 < \delta < 1$.[52] We take constitutional restrictions as given. They might reflect an allocation of spending powers that takes into account externalities, spillovers, and the strength of tax-benefit linkages for local voters, as highlighted by Tiebout (1956).[53]

Efficiency requires that every government provides the same amount of each good under its control, $e_t^{ij}(l) \perp l$ for all $l \in [0, \delta]$ and $e_t(l) \perp l$ for all $l \in (\delta, 1]$ implying

$$g_t^{ij} = \left(\delta \left(\frac{e_t^{ij}}{\delta} \right)^{\kappa} + (1-\delta) \left(\frac{e_t}{1-\delta} \right)^{\kappa} \right)^{\frac{1}{\kappa}} \times \left(\sum_{n=1}^{J} \theta_t^n \left(\delta \left(\frac{e_t^n}{\delta} \right)^{\kappa} + (1-\delta) \left(\frac{e_t}{1-\delta} \right)^{\kappa} \right) \right)^{\frac{\lambda}{\kappa}}.$$

The first term on the right-hand side of the preceding equation is defined as $a(e_t^{ij}, e_t)$ and the second term as $A(\vec{e}_t, e_t)^{\lambda}$.

For $\eta \to \infty$ ($\kappa \to 1$) the case of perfect substitutes follows,

$$g_t^{ij} = \left(e_t^{ij} + e_t \right) \times \left(\sum_{n=1}^{J} \theta_t^n e_t^n + e_t \right)^{\lambda}.$$

[52] If there were a third category of goods to which both governments could simultaneously contribute then only the government for which benefits outweigh the costs the most would contribute, due to perfect substitutability. As a consequence, the formulation in the text applies.

[53] We could additionally allow for differences in the efficiency of regional governments, relative to the federal government, in providing the different services. Denoting the relative productivity by $\nu(l)$ with $\nu' > 0$, the regional terms in the above specification would change to $\frac{e_t^{ij}(l)}{\delta\nu(l)}$ and $\frac{e_t^n(l)}{\delta\nu(l)}$. Such an extension could rationalize that regional governments (and the federal government) specialize on providing specific public services. It could be used to study the constitutional assignment of spending responsibilities. We leave such an extension for further work.

The constitutional restriction (encapsulated in the parameter δ) is irrelevant in this case. For $\eta \to 0$, the Leontieff case results.[54] For $\eta = 1$, the Cobb-Douglas specification follows,

$$g_t^{ij} = (e_t^{ij})^{\delta}(e_t)^{(1-\delta)(1+\lambda)} \prod_{n=1}^{J} (e_t^n)^{\delta\lambda\theta_t^n} \times \frac{1}{\Psi}.$$

In the main text we drop the constant term $\Psi \equiv \delta^{\delta(1+\lambda)}(1-\delta)^{(1-\delta)(1+\lambda)}$ as it is irrelevant, due to the logarithmic utility assumption.

B Elastic labor supply

We introduce additional taxes on labor income, levied at rates $\eta_t \geq 0$ and $\eta_t^{ij} \geq 0$ by the federal and regional governments respectively, whose proceeds are reimbursed to workers.[55] The program of a worker in region i of type j is given by

$$\begin{aligned} & \ln(c_{1,t}^{ij}) + v(l_t^{ij}) + \gamma_t^j \ln(g_t^{ij}) + \beta \left(\ln(c_{2,t+1}^{ij'}) + \mathbb{E}_t[\gamma_{t+1}^{j'} \ln(g_{t+1}^{ij'})] \right) \\ \text{s.t.} \quad & c_{1,t}^{ij} = w_t(1-l_t^{ij})(1-\tau_t-\tau_t^{ij}-\eta_t-\eta_t^{ij}) + T_t^j - s_t^{ij}, \quad c_{2,t+1}^{ij'} = s_t^{ij} R_{t+1}, \end{aligned}$$

where T_t^j denotes the lump sum transfer to workers. In equilibrium, $\tau_t^{ij} = \tau_t^j$ and $\eta_t^{ij} = \eta_t^j$. Moreover, since preferences for consumption and leisure do not vary across regions, $l_t^{ij} = l_t$ and $T_t^j = (\eta_t + \eta_t^j) w_t (1 - l_t)$. Workers' optimal savings and labor supply choices therefore imply

$$\frac{(1-\tau_t-\tau_t^j-\eta_t-\eta_t^j)(1+\beta)}{(1-\tau_t-\tau_t^j)(1-l_t)} = v'(l_t).$$

Thus, as long as $\eta_t + \eta_t^j > 0$, taxation distorts labor supply.

In addition to the terms present before, the objective functions of regional and federal voters now also account for the effect of leisure on utility. Moreover, the objective function of voters at the federal level also accounts for the general equilibrium implications of endogenous labor supply for contemporaneous and future interest rates and wages (the latter mediated through changes in capital accumulation). The objective functions of

[54] In this case we find $g_t^{ij} = w_t^{1+\lambda} \min\left[\frac{\tau_t - x_t}{1-\delta}, \frac{\tau_t^{ij} + \sigma x_t}{\delta}\right] \times \min\left[\frac{\tau_t - x_t}{1-\delta}, \frac{\min_n \tau_t^n + \sigma x_t}{\delta}\right]^{\lambda}$.

[55] For a related analysis in another context, see Gonzalez-Eiras and Niepelt (2008).

regional and federal voters, $\mathcal{V}_t^{ij}$ and $\mathcal{V}_t$ respectively, therefore are

$$\mathcal{V}_t^{ij} = V_t^{ij} + v(l_t^{ij}) + (1+\beta)\ln(1-l_t^{ij}),$$
$$\mathcal{V}_t = V_t + g(l_t) \equiv V_t + v(l_t) + \ln(1-l_t)\left[(1-\alpha)\left(1+\alpha\beta+\frac{\omega}{\nu_t}+\left(\frac{\omega}{\nu_t}+1\right)\gamma_t+\alpha\beta\gamma_{t+1}\right)\right],$$

where V_t^{ij} and V_t are defined in (7) and (8).

Because η_t^j is distorting and regional governments do not perceive general equilibrium effects, in equilibrium $\eta_t^j = 0$.[56]

At the federal level, the first-order condition with respect to η_t gives

$$\frac{dg(l_t)}{dl_t}\frac{\partial l_t}{\partial \eta_t} \leq 0.$$

If the equilibrium η_t is interior, then $\frac{\partial l_t}{\partial \eta_t} > 0$ implies $\frac{dg(l_t)}{dl_t} = 0$. Alternatively, if the equilibrium η_t is in a corner such that $\eta_t + \eta_t^j = 0$, then labor supply is unaffected by η_t (and $\frac{\partial l_t}{\partial \tau_t} = \frac{\partial l_t}{\partial \tau_t^j} = 0$).

Turning to the equilibrium choice of taxes that fund public services, we have

$$\frac{\partial \mathcal{V}_t^{ij}}{\partial \tau_t^{ij}} - \frac{\partial V_t^{ij}}{\partial \tau_t^{ij}} = \left(v'(l_t^{ij}) - \frac{1+\beta}{1-l_t^{ij}}\right)\frac{\partial l_t^{ij}}{\partial \tau_t^{ij}} \equiv -\mathcal{X}_t^l \leq 0 \quad \forall j,$$
$$\frac{\partial \mathcal{V}_t}{\partial \tau_t} - \frac{\partial V_t}{\partial \tau_t} = 0,$$

because, as shown above, either $\frac{dg(l_t)}{dl_t} = 0$ or $\frac{\partial l_t}{\partial \tau_t} = 0$ (and thus $\frac{dg(l_t)}{dl_t}\frac{\partial l_t}{\partial \tau_t} = 0$) when η_t is chosen optimally. Intuitively, the equilibrium choice of η_t "absorbs" all political cost-benefit considerations that relate to the distortion of labor supply, and the choice of τ_t therefore reflects the same considerations as in the model without elastic labor supply.

In conclusion, whether taxes to fund public services are raised at the regional or federal level thus depends on the strength of the general equilibrium effects on capital accumulation, $\mathcal{F}_t$, and the deadweight losses of taxation perceived by regional governments, $\mathcal{X}_t^l$.

C Proof of Proposition 2

(i) The marginal benefit of federal taxes includes the term $\sum_{j=1}^{J}\theta_t^j\left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t^j(1-\delta)}{\tau_t - x_t}$. Since spending complementarities imply that $0 < \delta < 1$, taxes cannot be zero since otherwise

[56] The derivative of the regional objective function with respect to η_t^j yields $-\frac{1+\beta}{1-l_t} + v'(l_t)$ which is negative if $\eta_t + \eta_t^j > 0$.

this term would diverge.

(ii) Suppose that all regional tax rates are interior such that

$$\tau_t^j = \frac{(1-\tau_t)\delta\gamma_t^j\Omega_t - (1+\beta)\sigma x_t}{1+\beta+\delta\gamma_t^j\Omega_t} \ \forall j.$$

This implies $\bar{\varphi}_t = (1+\beta)(1-\tau_t+\sigma x_t)\sum_j \theta_t^j/(1+\beta+\delta\gamma_t^j\Omega_t)$ and $\sum_j \theta_t^j/\varphi_t^j = (1+\beta+\delta\bar{\gamma}_t\Omega_t)/[(1+\beta)(1-\tau_t+\sigma x_t)]$.

With an interior federal tax rate the corresponding first-order condition holds with equality. Substituting the expressions above into this first-order condition yields

$$\frac{\Omega_t(1-\delta)(1+\lambda)\bar{\gamma}_t}{\tau_t - x_t} = \frac{1+\beta+\delta\Omega_t\bar{\gamma}_t + \frac{\beta}{1+\beta}\Lambda_{t+1}\left(\sum_j \frac{\theta_t^j}{1+\beta+\delta\gamma_t^j\Omega_t}\right)^{-1}}{1-\tau_t+\sigma x_t}.$$

Similarly, substituting the expressions above into the equilibrium condition for grants yields

$$\frac{\sigma}{\Omega_t}\frac{1+\beta+\delta\Omega_t\bar{\gamma}_t + \lambda\bar{\gamma}_t\sum_j \frac{\theta_t^j(1+\beta+\delta\gamma_t^j\Omega_t)}{\gamma_t^j}}{1-\tau_t+\sigma x_t} \leq \frac{(1+\lambda)(1-\delta)\bar{\gamma}_t}{\tau_t - x_t}.$$

Combining the last two relations, we conclude that interior tax rates at the federal level and in all regions constitute an equilibrium if the following parametric inequality condition is satisfied:

$$1+\beta+\delta\Omega_t\bar{\gamma}_t + \frac{\beta}{1+\beta}\Lambda_{t+1}\left(\sum_j \frac{\theta_t^j}{1+\beta+\delta\gamma_t^j\Omega_t}\right)^{-1} \geq$$
$$\sigma\left(1+\beta+\delta\Omega_t\bar{\gamma}_t + \lambda\bar{\gamma}_t\sum_j \frac{\theta_t^j(1+\beta+\delta\gamma_t^j\Omega_t)}{\gamma_t^j}\right).$$

In the non-generic case when the condition holds with equality positive tax rates constitute an equilibrium and grants are indeterminate. If the condition holds strictly then the marginal benefit of grants is negative; positive tax rates constitute an equilibrium as well in this case and grants equal zero.

(iii) If the parametric condition does not hold it must be the case that at least in one region the tax rate is zero. A similar reasoning as in (i) then implies that grants are positive, since otherwise the term $\left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t^j\delta}{\sigma x_t}$, capturing the marginal benefit of regional taxation (when $\tau_t^j = 0$), would diverge. Since the marginal benefit of regional taxation is increasing in the preference for the public service, γ_t^j, grants crowd out taxes in the regions with the lowest valuation.

(iv) A mean preserving spread of the γ_t^j's reduces the left hand side of (14) (see also above inequality) since it increases the geometric average of the $\frac{1}{\gamma_t^j}$'s. For the same reason this increases the right hand side of (14). Thus, a mean preserving spread of the γ_t^j's reduces the set of parameters for which (14) holds, rendering grants more likely.

D Matching Grants

With perfect substitutability and no traditional fiscal federalism motives, the first-order conditions are given by

$$\begin{aligned}\left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t(1+\sigma x_t)}{\tau_t^j(1+\sigma x_t)+\tau_t-x_t\bar{\tau}_t}-\frac{1+\beta}{\varphi_t^j}&\leq 0,\\ \left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t}{\bar{\tau}_t(1+(\sigma-1)x_t)+\tau_t}-\frac{1+\beta}{\varphi_t^j}+\mathcal{F}_t&\leq 0,\\ \frac{(\sigma-1)\bar{\tau}_t\gamma_t}{\bar{\tau}_t(1+(\sigma-1)x_t)+\tau_t}&\leq 0.\end{aligned}$$

The sign of $\mathcal{F}_t$ determines whether the national or regional governments tax, as with uniform grants. The first-order condition for grants implies that grants are not used if they entail deadweight losses. When $\sigma = 1$ and $\mathcal{F}_t > 0$ grants are irrelevant as it is immaterial whether spending is done by the federal or regional governments.

With complementarity the first-order conditions are given by

$$\begin{aligned}\Omega_t\frac{\gamma_t^j\delta}{\tau_t^j}-\frac{1+\beta}{\varphi_t^j}&\leq 0,\\ \sum_j\theta_t^j\left\{\Omega_t\frac{\gamma_t^j(1-\delta)}{\tau_t-x_t\bar{\tau}_t}-\frac{1+\beta}{\varphi_t^j}\right\}+\mathcal{E}_t+\mathcal{F}_t&\leq 0,\\ \sigma\delta\frac{(1+\lambda)\bar{\gamma}_t}{1+\sigma x_t}-\frac{(1-\delta)(1+\lambda)\bar{\tau}_t\bar{\gamma}_t}{\tau_t-x_t\bar{\tau}_t}&\leq 0,\end{aligned}$$

where $\mathcal{E}_t = \lambda\Omega_t\sum_j\theta_t^j\frac{\gamma_t^j(1-\delta)}{\tau_t-x_t\bar{\tau}_t}$ measures the federal government's marginal benefit of taxation due to cross-regional externalities. Note that now all tax rates are interior, regardless of whether there are grants in place or not, and the grant rate does not directly affect the incentives for regional taxes.

From the first-order condition for grants, $x_t = 0$ implies

$$\delta\sigma-\frac{(1-\delta)\bar{\tau}_t}{\tau_t}<0. \tag{17}$$

From regional governments' first order conditions we get $\tau_t^j = \frac{\Omega_t \delta \gamma_t^j (1-\tau_t)}{1+\beta+\Omega_t \delta \gamma_t^j}$, and $\varphi_t^j = \frac{(1+\beta)(1-\tau_t)}{1+\beta+\Omega_t \delta \gamma_t^j}$. Thus, we can calculate the following averages, and the factor price effect

$$\sum_j \frac{\theta_t^j}{\varphi_t^j} = \frac{1+\beta+\Omega_t \delta \bar{\gamma}_t}{(1+\beta)(1-\tau_t)},$$

$$\bar{\varphi}_t = (1+\beta)(1-\tau_t) \sum_j \frac{\theta_t^j}{1+\beta+\Omega_t \delta \gamma_t^j},$$

$$\mathcal{F}_t = -\frac{\beta \Lambda_{t+1}}{(1+\beta)(1-\tau_t)} \frac{1}{\sum_j \frac{\theta_t^j}{1+\beta+\Omega_t \delta \gamma_t^j}}.$$

Using the above expressions to solve for τ_t and $\bar{\tau}_t$ (assuming $x_t = 0$) we find

$$\tau_t = \frac{(1-\delta)(1+\lambda)\Omega_t \bar{\gamma}_t}{\Omega_t \bar{\gamma}_t \left[\delta + (1+\lambda)(1-\delta)\right] + 1 + \beta + \frac{\beta}{1+\beta} \frac{\Lambda_{t+1}}{\sum_j \frac{\theta_t^j}{1+\beta+\Omega_t \delta \gamma_t^j}}},$$

$$\bar{\tau}_t = (1-\tau_t)\Omega_t \delta \sum_j \frac{\theta_t^j \gamma_t^j}{1+\beta+\Omega_t \delta \gamma_t^j}.$$

Replacing in (17) we have that grants will not be used if

$$\sigma(1+\lambda)\bar{\gamma}_t < \sum_j \frac{\theta_t^j \gamma_t^j}{1+\beta+\Omega_t \delta \gamma_t^j} \left(\Omega_t \bar{\gamma}_t \delta + 1 + \beta + \frac{\beta}{1+\beta} \frac{\Lambda_{t+1}}{\sum_j \frac{\theta_t^j}{1+\beta+\Omega_t \delta \gamma_t^j}} \right).$$

With homogeneous preferences, this simplifies to

$$\sigma(1+\lambda) < 1 + \frac{\beta}{1+\beta} \Lambda_{t+1},$$

the same condition as in the case with uniform grants.

References

Alesina, A. and Spolaore, E. (1997), 'On the number and size of nations', *Quarterly Journal of Economics* **112**(4), 1027–1056.

Ansolabehere, S., Rodden, J. and Snyder Jr., J. M. (2006), 'Purple America', *Journal of Economic Perspectives* **20**(2), 97–118.

Besley, T. and Coate, S. (2003), 'Centralized versus decentralized provision of local public goods: A political economy approach', *Journal of Public Economics* **87**(12), 2611–2637.

Boffa, F., Piolatto, A. and Ponzetto, G. A. M. (2016), 'Political centralization and government accountability', *Quarterly Journal of Economics* **131**(1), 381–422.

Bolton, P. and Roland, G. (1997), 'The breakup of nations: A political economy analysis', *Quarterly Journal of Economics* **112**(4), 1057–1090.

Dales, S. R. (1970), 'Federal grants to state and local governments 1968–1969', *Social Security Bulletin* **33**, 31–38.

Dales, S. R. (1971), 'Federal grants to state and local governments 1969–1970', *Social Security Bulletin* **34**, 15–23.

Frank, T. (2004), *What's the Matter with Kansas?*, Henry Holt, New York.

Gonzalez-Eiras, M. and Niepelt, D. (2005), Sustaining social security, Working Paper 1494, CESifo.

Gonzalez-Eiras, M. and Niepelt, D. (2008), 'The future of social security', *Journal of Monetary Economics* **55**(2), 197–218.

Gonzalez-Eiras, M. and Niepelt, D. (2015), 'Politico-economic equivalence', *Review of Economic Dynamics* **18**(4), 843–862.

Gonzalez-Eiras, M. and Niepelt, D. (2016), Fiscal federalism, taxation, and grants, Discussion Paper 11482, CEPR.

Gordon, R. H. (1983), 'An optimal taxation approach to fiscal federalism', *Quarterly Journal of Economics* **98**(4), 567–586.

Gordon, R. J. (2012), Is U.S. economic growth over? Faltering innovation confronts the six headwinds, Working Paper 18315, NBER, Cambridge, Massachusetts.

Greenstone, M. and Looney, A. (2012), The truth about taxes: Just about everyone pays them, The Hamilton Project, http://goo.gl/dvnr0d, Brookings, Washington.

Gruber, J. (2011), *Public Finance and Public Policy*, 3rd edn, Worth, New York.

Harstad, B. (2007), 'Harmonization and side payments in political cooperation', *American Economic Review* **97**(3), 871–889.

Hatfield, J. W. and Padró i Miquel, G. (2012), 'A political economy theory of partial decentralization', *Journal of the European Economic Association* **10**(3), 605–633.

Kessler, A. S. (2014), 'Communication in federal politics: Universalism, policy uniformity, and the optimal allocation of fiscal authority', *Journal of Political Economy* **122**(4), 766–805.

Kincaid, J., ed. (2011), *Federalism*, Sage Publications, Los Angeles.

Knight, B. (2002), 'Endogenous federal grants and crowd-out of state government spending: Theory and evidence from the federal highway aid program', *American Economic Review* **92**(1), 71–92.

Kotlikoff, L. J. and Rosenthal, R. W. (1990), Some inefficiency implications of generational politics and exchange, Working Paper 3354, NBER, Cambridge, Massachusetts.

Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1997), 'Politico-economic equilibrium and economic growth', *Journal of Economic Dynamics and Control* **21**(1), 243–272.

Lindbeck, A. and Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public Choice* **52**, 273–297.

Lockwood, B. (2002), 'Distributive politics and the costs of centralization', *Review of Economic Studies* **69**(2), 313–337.

Mehrotra, A. K. (2013), *Making the Modern American Fiscal State: Law, Politics, and the Rise of Progressive Taxation, 1877–1929*, Cambridge University Press, New York.

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

Oates, W. E. (1999), 'An essay on fiscal federalism', *Journal of Economic Literature* **37**(3), 1120–1149.

Persson, T. and Tabellini, G. (1996), 'Federal fiscal constitutions: Risk sharing and redistribution', *Journal of Political Economy* **104**(5), 979–1009.

Persson, T. and Tabellini, G. (2000), *Political Economics*, MIT Press, Cambridge, Massachusetts.

Piketty, T. and Saez, E. (2003), 'Income inequality in the United States, 1913–1998', *Quarterly Journal of Economics* **118**(1), 1–39.

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

Soares, J. (2005), 'Public education reform: Community or national funding of education?', *Journal of Monetary Economics* **52**(3), 669–697.

Song, Z., Storesletten, K. and Zilibotti, F. (2012), 'Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt', *Econometrica* **80**(6), 2785–2803.

Tiebout, C. M. (1956), 'A pure theory of local expenditures', *Journal of Political Economy* **64**(5), 416–424.

Treisman, D. (2007), *The Architecture of Government*, Cambridge University Press, Cambridge.

Vélez, J. L. (2014), War and progressive income taxation in the 20th century, Working Paper 2014-3, Berkeley Economic History Laboratory, Berkeley, California.

Wallis, J. J. (2000), 'American government finance in the long run: 1790 to 1990', *Journal of Economic Perspectives* **14**(1), 61–82.

Wallis, J. J. and Oates, W. E. (1998), The impact of the New Deal on American federalism, *in* M. D. Bordo, C. Goldin and E. N. White, eds, 'The Defining Moment: The Great Depression and the American Economy in the Twentieth Century', University of Chicago Press, Chicago, chapter 5, pp. 155–180.