Niepelt, Dirk

Working Paper

Dynamic tax externalities and the U.S. fiscal transformation in the 1930s

Discussion Papers, No. 18-03

Provided in Cooperation with:
Department of Economics, University of Bern

Suggested Citation: Niepelt, Dirk (2018) : Dynamic tax externalities and the U.S. fiscal transformation in the 1930s, Discussion Papers, No. 18-03, University of Bern, Department of Economics, Bern

This Version is available at:
https://hdl.handle.net/10419/204899

UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Dynamic Tax Externalities and the U.S. Fiscal Transformation in the 1930s

Dirk Niepelt

18-03

März 2018

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Dynamic Tax Externalities and the U.S. Fiscal Transformation in the 1930s*

Martín Gonzalez-Eiras† Dirk Niepelt‡

March 21, 2018

Abstract

We propose a theory of tax centralization in politico-economic equilibrium. Taxation has dynamic general equilibrium implications which are rationally internalized at the federal, but not at the regional level. The political support for taxation therefore differs across levels of government. Complementarities on the spending side decouple the equilibrium composition of spending and taxation and create a role for inter governmental grants. The model provides an explanation for 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, it accounts for between 30% and 100% of the federal revenue share's doubling in the 1930s, and for the long-term increase in federal grants.

JEL Classification: D72, E62, H41, H77

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

1 Introduction

Whether control over fiscal policy should rest with national, regional or local governments depends on how effective these agents make use of their authority. When information frictions render it difficult to cater to heterogeneous needs, fiscal policy is best chosen de-centrally. When it is key to internalize spillover effects, in contrast, centralized policy choices are advantageous. A broad body of fiscal federalism literature has studied the normative and positive implications of this fundamental trade-off. The focus of that literature has generally been on static sources of the cost-benefit differences across governments.

*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.

In this paper, we propose a complementary—dynamic—source that arises from general equilibrium effects of taxation. We show that this dynamic source of cost differences across governments is present under a wide set of assumptions, and we argue that it can help explain the dramatic fiscal transformation in the United States during the 1930s when federal tax collections increased strongly and federal grants started their long-term rise.

The model features overlapping generations that work, save, consume, and vote as well as a central, or federal, government and many regional governments that impose labor income taxes to finance the provision of public services.[1] In politico-economic equilibrium, households make optimal savings choices conditional on current and expected future policies; and they vote for their preferred political candidates, taking the policy functions of other political decision makers as well as the competitive equilibrium conditions into account.

Taxation slows down capital accumulation and thus has general equilibrium effects: It drives up interest rates and lowers future wages. Voters and policy makers at the federal level—rationally—internalize these general equilibrium effects, at least to the extent that they affect them.[2] In contrast, voters and policy makers 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 regional elections.

In the baseline specification, federal and regional spending are perfect substitutes and all static sources of cost-benefit differences across governments are absent: Government spending does not generate externalities; preferences for public services are uniform across the population; and regional tax bases are immobile, undermining any motive for tax competition. We show that, nevertheless, the incentives to raise taxes at the federal level differ from those at the regional level. Depending on the sign of the dynamic general equilibrium effects, federal taxation enjoys stronger or weaker support in politico-economic equilibrium than regional taxation, and the equilibrium composition of tax collections across governments is determinate.

This result is robust along many dimensions. We introduce labor mobility across regions and find that it does not fundamentally alter our findings. We allow for elastic labor supply, tax distortions, and additional policy instruments and show that the results are robust since the perceived cost differences due to general equilibrium effects are orthogonal to the effects of tax distortions. We also find that our results are robust to introducing policy instruments for intergenerational redistribution, such as public debt or social security, or longer-lived households.

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

[1] 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.

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

accumulation because they are chosen ex post and reduce the income of the old.[3] From the perspective of federal and regional voters, the net cost of taxation thus is the same and the asymmetry in the political support for federal versus regional taxation disappears. A shift from capital to labor income taxation therefore can trigger a major change in the composition of tax collections, favoring federal or regional income taxation.

We argue that this mechanism offers an explanation for the dramatic fiscal transformation that the United States underwent during the 1930s, see figures 1 and 2.[4] On the eve of the Great Depression, local governments collected the majority of tax revenues and property taxes accounted for nearly half of all revenues. The federal government's main source of revenue were tariffs, and on a smaller scale, property taxes.[5] In the 1930s this arrangement changed completely.[6] As indicated by the solid line in figure 1 the revenue share of the federal government nearly doubled and inter governmental grants (indicated by dots and discussed below) emerged as a central source of revenue for state and local governments. An even more dramatic transformation occurred with respect to the federal tax base, see figure 2. The income tax share of federal revenues more than doubled within a few years and continued to grow quickly for another decade, and the share of tax units who paid federal income tax similarly exploded.

Our model explains this transformation as the equilibrium response to the ratification of the Sixteenth Constitutional Amendment which introduced the *possibility* for the federal government to tax income.[7] Starting from a situation with exclusive competence for income taxation at the level of the states and strong reliance on property taxes (which generate limited general equilibrium effects), the ratification opened the door for the federal government to tax labor income and to exploit the general equilibrium effects we emphasize.[8] Stronger demand for government outlays, specifically for New Deal policies and spending related to World War II, subsequently strengthened the *incentives* to employ the newly available tax instrument, and the fiscal transformation took place.

Importantly, our model does not aim at explaining the *increase* in total government

[3]When cohorts live longer than for two periods then capital income taxes do affect capital accumulation but by less than labor income taxes. The perceived difference in tax collection costs across governments then is smaller than with labor income taxes.

[4]See Wallis (2000) for a discussion of this and two earlier transformations of the American fiscal architecture.

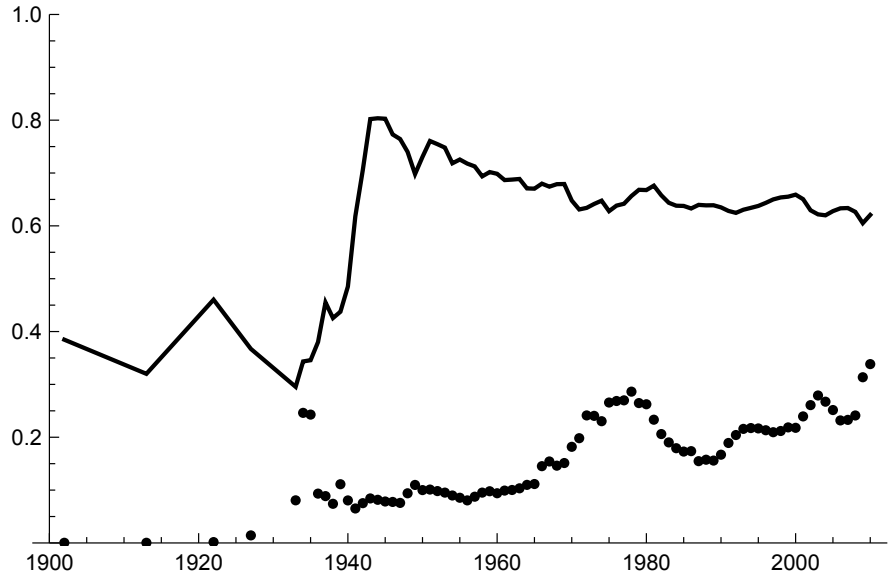
[5]Federal tariffs were set against the background of intense interstate conflict between more industrialized states that demanded protection from imports, and states that relied more heavily on such imports. Proponents of tariffs emphasized infant industry arguments and stressed the need for national self sufficiency. See Taussig (1910). Note that the Constitution rules out interstate tariffs.

[6]See Wallis and Oates (1998) for a description of New Deal programs and a discussion of federal deficits that accompanied the transformation.

[7]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."

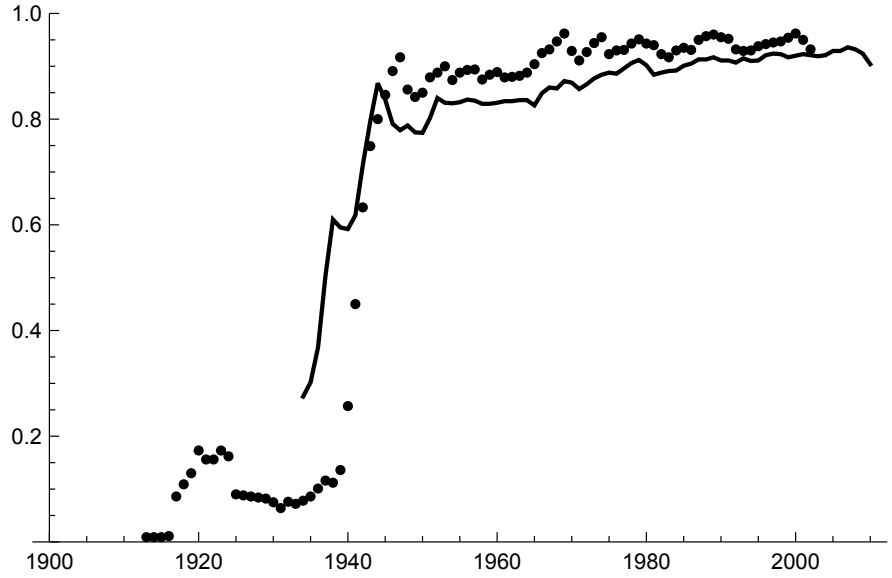
[8]Income taxes predominantly affect savers. While they are collected from both workers (who save) and retirees (who do not), the fraction of individuals paying taxes sharply falls with age, see for example Greenstone and Looney (2012). Applying Piketty and Saez's (2003) methodology to classify respondents 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.

Figure 1: Fiscal transformation in the United States: Federal revenues and grants



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.

Figure 2: Fiscal transformation in the United States: Federal 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.

revenues. Instead, it aims at explaining the shift in the *composition* of government financing, the centralization of revenue collection, as well as the decoupling of revenue collection and government spending that is reflected in the rise of inter-governmental grants (discussed below).

Our model abstracts from within-cohort distributive conflicts as well as administrative and legal hurdles that had to be overcome before the federal government could start levying a comprehensive income tax, and which were brushed aside by the need to increase expenditure for New Deal policies and World War II spending.[9] [10] While the model thus is silent about the time lag between the Amendment's ratification (1913) and the actual transformation in the 1930s it does account for the magnitude of the transformation. A calibrated baseline model with two groups of regions that differ in terms of their preference for public services predicts a rise in the federal government's tax share in line with the data once a tax without general equilibrium effects is replaced by a tax with such effects.

To study the second feature of the fiscal transformation illustrated in figure 1, and to further discipline the quantitative analysis, we introduce a role for federal grants. This requires that the composition of both taxes and spending across governments is determinate. We therefore relax the assumption that regional and federal spending are perfect substitutes and instead assume complementarities between the two spending components. The grant instrument has value when it is beneficial to channel revenue from the federal government to regions, either because tax revenue at the federal level is "cheap" or because positive cross-regional externalities from public service provision imply inefficiently low regional spending in the absence of grants.

The extended framework is as analytically tractable as the baseline model. In line with empirical evidence, it predicts grants to crowd out local taxation.[11] More importantly, it is able to explain the trend increase of grants as the result of time varying regional disparities which we associate (supported by empirical evidence) with preference heterogeneity regarding the size of government between urban and rural regions. The model's explanatory power for the federal tax share remains substantial; the extended framework explains roughly 30% of the observed increase between 1930 and 1950. In both the basic and the extended framework, we find that an alternative explanation for the fiscal transformation that relies on static spending externalities as they are discussed in the fiscal federalism literature cannot satisfactorily quantitatively explain the data.

[9]The Treasury underwent a major reorganization; the number of employees at the Bureau of Internal Revenue increased fourfold; and the Supreme Court had to uphold the constitutionality of tax legislation enacted based on the Amendment (in 1916) and made a series of decisions relating to the proper 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, reflected in sharp swings in the highest marginal tax rate and the share of the population that was tax liable.

[10]Vélez (2014) argues that across a sample of OECD countries, progressive income taxation was instituted as a response to war efforts.

[11]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.

Related Literature We build on the classic analysis of fiscal federalism that stresses trade-offs between forces favoring centralization and decentralization. Oates (1972) finds that absent spillovers and cost-savings from centralized tax collection or public good provision, heterogeneous preferences render decentralization preferable. Without information frictions, a centralized system may in principle support differentiated provision (Oates, 1999). But various political economy frictions favor uniform centralized policy choices.[12] Alesina and Spolaore (1997) analyze the effect of international integration on the costs and benefits of centralization and thus, the number of countries.

Similar arguments are discussed in the political science literature (e.g., Kincaid, 2011) which tends to favor federalist governance structures for diverse countries. Treisman (2007) questions many rationales for and against political decentralization. He argues that administrative efficiency requires administrative, not political decentralization and he criticizes the view that local governments better manage local information.[13] Our argument is related in so far as it stresses the decoupling of tax and spending decisions.

Horizontal and vertical tax competition in federal structures gives rise to important static externalities. Gordon (1983) discusses revenue (and other) externalities due to uncoordinated regional taxation of mobile factors, and Keen and Kotsogiannis (2004) analyze the interplay between horizontal and vertical tax externalities. A federal government concerned with welfare at the national level may correct some of these externalities by imposing federal taxes or extending federal grants. Our paper also emphasizes tax externalities but of a different type, namely dynamic externalities due to general equilibrium effects, and it builds a positive theory of fiscal federalism and federal grants.[14]

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.

Wallis (2000) documents that the U.S. passed through distinct regimes of government finance and suggests that the costs of raising revenue differ across governments. Our model provides an explanation for such cost differences that stresses general equilibrium effects.[15]

[12]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); credibility problems in signalling local tastes to the central government may generate inefficient federal policy choices (Kessler, 2014); and centralization to increase accountability may have to be accompanied by policy uniformity because otherwise, the central government would implement policies favoring regions that monitor more extensively (Boffa, Piolatto and Ponzetto, 2016).

[13]According to Treisman (2007), decentralization is important for policy stability and centralization is important for fiscal coordination.

[14]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.

[15]In other contexts, Kotlikoff and Rosenthal (1990), Soares (2005), and Gonzalez-Eiras and Niepelt

This explanation complements alternative, static theories that rely on permanently lower information processing costs for the federal government;[16] permanently higher externalities from public infrastructure investment; or interstate mobility and tax competition which have been criticized.[17] In addition to offering a novel source of differences in the cost of taxation our model can quantitatively account—at least in parts—for the dramatic transformation of the U.S. fiscal system during the 1930s and afterwards.

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. Section 4 contains the analysis of the baseline model and several extensions. In section 5, we contrast the model's quantitative implications with empirical evidence on the fiscal transformation in the U.S. during the 1930s and subsequent fiscal trends. Section 6 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. This demographic parameter follows a deterministic process.

The economy is composed of a continuum of regions of measure one over the unit interval. Each region is populated by a continuum of homogeneous agents. The population structure including the preferences of agents are the same across regions (we relax this assumption later). Regions are indexed by i.

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

(2008) have emphasized the role of general equilibrium effects in politico-economic equilibrium.

[16] Wallis (2000) suggests that the introduction of Social Security payroll taxes could have lowered federal tax collection costs.

[17] Rhode and Strumpf (2003) document that households' migration decisions mostly reflect personal factors rather than Tiebout (1956) sorting and they find that the secular decrease in mobility costs in the United States was not accompanied by stronger policy or preference heterogeneity across communities. Rhode and Strumpf (2003) conclude that "any theoretical or empirical model that adopts a pure Tiebout framework ... is misspecified" (p. 160). Similarly, Baicker, Clemens and Singhal (2012) find that "patterns in mobility, seem to have little power to explain observed changes in the landscape of fiscal federalism" (p. 1080).

regional elections, respectively. None of the governments can commit, and in each period they take decisions simultaneously.[18]

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 relax this assumption later). 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, g_t^i, depends on public spending at the regional level and nationwide. Let e_t^i denote spending at the regional level and e_t the—uniform—spending by the federal government.[19] In the baseline model, we assume that federal and regional spending are perfect substitutes in the production of public services and we abstract from externalities across regions (we relax both assumptions below). Accordingly,

$$g_t^i = e_t^i + e_t. \tag{2}$$

[18] 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.

[19] Thus, we allow for both levels of government to tax and spend. For rationalizations of policy uniformity at the federal level, see the literature review in the introduction.

Spending by the federal government is financed by a labor income tax at rate τ_t and spending by region i is financed by a tax at rate τ_t^i. (Below, we introduce federal grants as an additional source of regional revenue.) Since all governments balance their budget in each period this implies

$$e_t = w_t \tau_t, \quad e_t^i = w_t \tau_t^i. \tag{3}$$

Tax rates are non-negative.

2.4 Preferences and Household Choices

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

$$\begin{aligned} & \ln(c_{1,t}^i) + \gamma_t \ln(g_t^i) + \beta \left(\ln(c_{2,t+1}^i) + \gamma_{t+1} \ln(g_{t+1}^i) \right) \\ \text{s.t.} \quad & c_{1,t}^i = w_t(1 - \tau_t - \tau_t^i) - s_t^i, \quad c_{2,t+1}^i = s_t^i R_{t+1}. \end{aligned}$$

Parameter γ_t measures the preference for public services.

Taking prices and taxes as given the worker optimally chooses

$$s_t^i = \frac{\beta}{1+\beta} w_t (1 - \tau_t - \tau_t^i). \tag{4}$$

Since equilibrium consumption and saving of a worker are proportional to the after tax wage, lifetime utility is proportional to the log after tax wage, the log gross interest rate, and the logarithms of current and future public services. We summarize this information in the indirect utility functions of workers and retirees, respectively, which equal (dropping constants)

$$\begin{aligned} U_t^{i,w} &= (1+\beta)(\ln(w_t) + \ln(1 - \tau_t - \tau_t^i)) + \beta \ln(R_{t+1}) + \gamma_t \ln(g_t^i) + \beta \gamma_{t+1} \ln(g_{t+1}^i), & (5) \\ U_t^{i,r} &= \ln(s_{t-1}^i) + \ln(R_t) + \gamma_t \ln(g_t^i) & (6) \end{aligned}$$

subject to (2), (3). Here, s_{t-1}^i denotes the inherited stock of savings by retirees.

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 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.[20] 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 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

The state is given by z_t, which includes the exogenous demographic parameter as well as the cross section of savings levels across regions which we denote by $\vec{s}_{t-1}$. (Throughout the paper, we use this notation for cross sections.) Conditional on z_t, the production function as well as competition among firms determine factor prices, w_t and R_t. A financing policy (or policy for short) of all regions and the federal government, $(\vec{\tau}_t, \tau_t)$, then determines public services, $\vec{g}_t$, capital accumulation, $\vec{s}_t$, and thus z_{t+1}. Conditional on z_t, a policy sequence $\{\vec{\tau}_s, \tau_s\}_{s \geq t}$ then fully determines an allocation and price system.

We focus on symmetric equilibria where all regions behave identically, except possibly a set of regions of measure zero. We denote the "typical" regional tax by τ_t^j, and "typical" public services by g_t^j.

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

i. capital evolves according to $k_t = s_{t-1}/\nu_t$, 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, t.

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

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, $\vec{s}_{t-1}$, such that future policy choices are unaffected by current policy choices. This conjecture is motivated by two observations. First, the indirect utility functions are additively separable in prices and policy (reflecting our assumption about preferences and the production of public services); and second, the elasticities of factor prices with respect to the capital-labor ratio are orthogonal to the latter (reflecting our assumption about the aggregate production function).

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.

Formally, under the conjecture a regional decision maker at date t takes $(w_t, w_{t+1}, R_t, R_{t+1})$ as well as s^i_{t-1} and $(\tau^j_t, \tau_t, \tau^i_{t+1}, \tau^j_{t+1}, \tau_{t+1})$ as given and her objective is $\omega U^{i,r}_t/\nu_t + U^{i,w}_t$. Effectively, she maximizes

$$V^i_t \equiv \left(\frac{\omega}{\nu_t}+1\right)\gamma_t \ln(g^i_t) + (1+\beta)\ln(1-\tau_t-\tau^i_t) \quad \text{s.t.} \quad (2),(3). \tag{7}$$

In contrast, the federal decision maker at date t takes (w_t, R_t) as well as s_{t-1} and $(\tau^j_t, \tau^j_{t+1}, \tau_{t+1})$ as given and she is concerned with $\omega U^{j,r}_t/\nu_t + U^{j,w}_t$. Effectively, she maximizes

$$\begin{aligned} V_t \quad \equiv \quad & \left(\frac{\omega}{\nu_t}+1\right)\gamma_t \ln(g^j_t) + (1+\beta)\ln(1-\tau_t-\tau^j_t) + \beta\ln(R_{t+1}) + \beta\gamma_{t+1}\ln(g^j_{t+1}) \quad (8) \\ & \text{s.t.} \quad (1),(2),(3),(4), k_{t+1} = s_t/\nu_{t+1}. \end{aligned}$$

We can now define politico-economic equilibrium (under the conjecture).[21]

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

i. $\tau^i_t \geq 0$ maximizes V^i_t and $\tau^i_t = \tau^j_t$ for all i, t;

ii. $\tau_t \geq 0$ maximizes V_t for all t; and

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

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

4 Analysis

4.1 Main Result

Absent heterogeneity in regional preferences or static externalities from spending or taxation across regions, none of the traditional static fiscal federalism motives for decentralization or centralization is present. Nevertheless, the equilibrium degree of centralization of tax collections generally *is* determinate. To see this, consider the derivative of the regional objective function V_t^i with respect to the regional tax rate, τ_t^i (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 derivatives of V_t^i in (7) and of V_t in (8) must be weakly negative in equilibrium,

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

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

In addition, the corresponding complementary slackness conditions must be 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 static spending externalities.

The third term in the second inequality,

$$\mathcal{F}_t\equiv-\frac{\beta}{1-\tau_t-\tau_t^j}\left(\epsilon_{Rk}+\epsilon_{wk}\gamma_{t+1}\right),$$

represents the *indirect* net benefit of higher taxes due to general equilibrium factor price effects 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 $\mathrm{d}\ln(s_t)/\mathrm{d}\tau_t=-1/(1-\tau_t-\tau_t^j)$, 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 collections, 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 and thus, 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 and $\tau_t = 0$.

Intuitively, 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.

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 and the derivatives of the indirect utility functions 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.[22] We summarize these findings in the following proposition:

Proposition 1. 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.

The "bang-bang" property of the equilibrium policy is a direct consequence of the fact that only $\mathcal{F}_t$ drives a wedge between the regional and federal first-order conditions, and that the sign of this wedge does not vary with taxes. In the quantitative analysis, we introduce preference heterogeneity. When general equilibrium effects are positive but small, federal taxes then only crowd out regional taxation in low valuation regions; this smoothes the response of average regional taxes to an increase in $\mathcal{F}_t$.

Before turning to the quantitative analysis, however, we discuss the robustness of the finding in proposition 1.

[22] 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.

4.2 Endogenous Labor Supply

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 preferences are given by

$$\ln(c^i_{1,t}) + v(l^i_t) + \gamma_t \ln(g^i_t) + \beta \left(\ln(c^i_{2,t+1}) + \gamma_{t+1} \ln(g^i_{t+1}) \right),$$

where l^i_t and $v(\cdot)$ denote leisure and a smooth utility function, respectively.[23] The budget constraint of a worker now reads

$$c^i_{1,t} = w_t(1 - l^i_t)(1 - \tau_t - \tau^i_t) - s^i_t.$$

It is easy to check that in this more general 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 A, 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 A, introduction of these new instruments does not change the first-order condition for τ_t, but adds a distortion term, $-\mathcal{X}^{ls}_t \leq 0$ say, to the first-order condition for τ^j_t. The results of proposition 1 thus continue to hold subject to replacing $\mathcal{F}_t$ by $\mathcal{F}_t + \mathcal{X}^{ls}_t$: Taxation at the federal level constitutes an equilibrium outcome as long as $\mathcal{F}_t + \mathcal{X}^{ls}_t > 0$. 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.

4.3 Labor Mobility

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 voters 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 $\mathcal{X}^{lm}_t$ the welfare cost of such emigration, results similar to those of proposition 1 follow, with taxation at the federal level an equilibrium outcome as long as $\mathcal{F}_t + \mathcal{X}^{lm}_t > 0$.

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

4.4 Government Debt and 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 the main result.

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.[24] The first-order conditions that characterize public services provision, conditions (9) and (10), then are unchanged except that the tax wedge now includes the new tax rate. This might affect the magnitude of the general equilibrium term, $\mathcal{F}_t$, but not its sign.[25] The main message of proposition 1 therefore is robust: The level of government that collects taxes to fund public services is determined by the sign of $\mathcal{F}_t$.

4.5 Longer-Lived Households

In the baseline model, voters at the federal level fully account for the general equilibrium effect on interest rates while they internalize the general equilibrium effect on wages only partly, to the extent that it affects the public service provision in the subsequent period. This asymmetry is a consequence of the assumption that agents live for just two periods; if households lived, and supplied labor for more than two periods then some of the voters would also internalize the effect of contemporaneous taxes on their own subsequent wage income.

This feature is irrelevant for the results summarized in proposition 1, though, since these results hold independently of the weight attached to the effect on future wages. In the baseline model, the weight reflects the preference for public services. But nothing substantive would change if the weight also reflected future labor income. Note also that some asymmetry of the type described above would remain in place even if agents lived for many periods. This is because independently of agents' life span, current voters always fully internalize the effects of policy on future capital income while they only partly internalize the effects on future labor income which also benefits some workers who are yet unborn when policy is chosen.

[24] 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.

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

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

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.

4.6 Capital Income Taxes

Finally, consider the implications of changing the tax base from labor to capital income. 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, $\mathcal{F}_t = 0$. Moreover, the weight that political candidates attach to the cost of taxation changes from $(1+\beta)$ to ω/ν_t since the old rather than the young bear the tax burden. Otherwise, the first-order conditions for taxes remain unchanged.[26]

Total taxes thus are determinate in equilibrium but the degree of centralization of tax collections is not. Intuitively, unlike in the baseline model with labor income taxes, voters at the federal and regional level in an economy with capital income taxes perceive exactly the same trade-off when weighing the pros and cons of a tax hike. Note that the indeterminacy of the equilibrium composition of capital income taxes is closely related to the bang-bang property of equilibrium labor income taxes in the baseline model (see proposition 1). For example, with region-specific preferences over public services, regions with a high preference for public services may levy labor income taxes in the baseline model even if $\mathcal{F}_t > 0$; and the same regions may levy capital income taxes in the modified setup although $\mathcal{F}_t = 0$ and although the federal government levies taxes as well. We discuss the case with region-specific preferences (and static externalities and complementarities) below.

5 Quantitative Analysis

We analyze the model's predictive power in two steps. In the next subsection, we focus on the relative size of the federal government and how it grew during the 1930s. As discussed in the introduction, we relate the observed growth to the new policy options afforded by the ratification of the 16th Constitutional Amendment and the increased demand for fiscal resources against the background of the New Deal and World War II. Subsequently, we include grants in the analysis and extend the forecast period. We do this for two reasons. First, because the steady increase in the importance of federal grants constitutes a key element of the fiscal transformation that began in the 1930s. And second, because motives to provide federal grants and to tax at the federal level naturally relate to each other such that introducing the former imposes additional discipline in our quantitative analysis.

5.1 The Rise of Federal Taxation

Between 1900 and 1930 the federal government's share in tax collections averaged roughly 38.3%; by 1950 this share had risen to 72.6%, with most of the increase occurring during the 1930s (see figure 1 in the introduction). To assess whether dynamic general equilibrium effects of taxation can explain this increase, we simulate a simple extension of

[26] 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 cost and benefit of taxation since preferences are logarithmic.

the baseline model that allows for preference heterogeneity across regions. The need to introduce heterogeneity (or some other modification) arises because the baseline model generates "bang-bang" solutions (see proposition 1) while in the data, the inter governmental composition of tax collections is interior.

We assume that there are two types of regions, with high and low preference for public services, γ_t^1 and $\gamma_t^2 < \gamma_t^1$ respectively. The share of high and low preference regions is denoted θ_t^1 and θ_t^2. In symmetric equilibrium all regions within the same group behave identically, except possibly a set of regions of measure zero. The endogenous state therefore contains the savings of retirees in the typical regions, $\vec{s}_{t-1} = (s_{t-1}^1, s_{t-1}^2)$, and policy at date t is given by $(\vec{\tau}_t, \tau_t)$. In the equilibrium we focus on, high preference regions levy taxes and provide public services even if the federal government does the same; low preference regions, in contrast, do not levy taxes and fully rely on federal provision.

The equilibrium conditions in the extended baseline model are given by

$$\left(\frac{\omega}{\nu_t}+1\right)\frac{\gamma_t^1}{\tau_t^1+\tau_t}-\frac{1+\beta}{1-\tau_t-\tau_t^1}=0,$$
$$\left(\frac{\omega}{\nu_t}+1\right)\left(\frac{\theta_t^1\gamma_t^1}{\tau_t^1+\tau_t}+\frac{\theta_t^2\gamma_t^2}{\tau_t}\right)-(1+\beta)\left(\frac{\theta_t^1}{1-\tau_t-\tau_t^1}+\frac{\theta_t^2}{1-\tau_t}\right)+\mathcal{F}_t=0, \quad (11)$$

with $\mathcal{F}_t \equiv -\beta/(1-\tau_t-\bar{\tau}_t)\left(\epsilon_{Rk}+\epsilon_{wk}\bar{\gamma}_{t+1}\right)$, where a "bar" denotes the average across regions. The first condition which reflects the tax choice in high preference regions is identical to condition (9). Condition (11), which reflects the choice of federal tax rate, generalizes condition (10) by summing and weighing region-specific marginal effects.

Our strategy to assess the importance of dynamic general equilibrium effects for the tax structure is as follows: We calibrate the model with general equilibrium effects from taxation ($\mathcal{F}_t \neq 0$) to exactly match the size of governments in 1950. (We use 1950 rather than 1940 as the base year because by that time government finances likely reflected a more regular, post-recovery and post-war mode.) We then shut down the general equilibrium effects from taxation ($\mathcal{F}_t = 0$) and solve for the equilibrium that we associate with the situation before the fiscal transformation when the federal government mostly relied on tariffs and property taxes. We do this assuming either the demographic structure in the year 1930 or 1950. In either case, we find that shutting down the general equilibrium channel implies a large drop in the federal tax share.

For the calibration, we assume that one period in the model corresponds to 30 years in the data. We 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 in 1950 and use Census Bureau data. From Gonzalez-Eiras and Niepelt (2008) we take $\omega = 0.9176$.

We associate regions that optimally levy taxes with "urban" regions and those that do not with "rural" regions, and we proxy the share of urban regions, θ_t^1, by 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.[27] On the other hand, the distinction also

[27]For example, Frank (2004) argues that low-income Americans living in rural areas vote strongly

Table 1: Support for government spending cuts (in %)

	1985	1990	1996	2006
Total	82.0	78.2	83.4	63.3
Urban	80.7	77.5	83.1	62.7
Rural	87.3	82.1	86.1	67.4

The table shows the percentage of respondents answering "strongly in favor of" or "in favor of" government spending cuts. Data from General Social Survey. Counties without towns of 10,000 or more inhabitants are classified as rural. There are between 540 and 1293 urban observations in the four samples and between 126 and 190 rural observations.

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 favored government spending cuts more strongly than respondents in urban areas, see table 1.[28]

To calibrate the preference for public services as well as β, we use the extended baseline model's first-order conditions for τ_{1950} and τ^1_{1950} (which feature γ^j_{1950} and γ^j_{1980}) and assume that between 1950 and 1980, preferences for public services increase at the same rate as the total size of government (federal, state, and local spending relative to GDP).[29] In addition, we use a moment condition for the Euler equation in steady state (see Gonzalez-Eiras and Niepelt, 2008).[30] Table 2 lists the calibrated parameters and figure 3 illustrates the demographic trends that we feed into the model of this and the subsequent subsection. The β value corresponds to an annual discount factor of approximately 0.9769. The low preference for public services in rural regions reflects our assumption that taxes equal zero in those regions.

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

[28]The annual survey is conducted by The National Data Program for the Social Sciences. Respondents in the years 1985, 1990, 1996, and 2006 were asked about their attitudes towards government spending cuts.

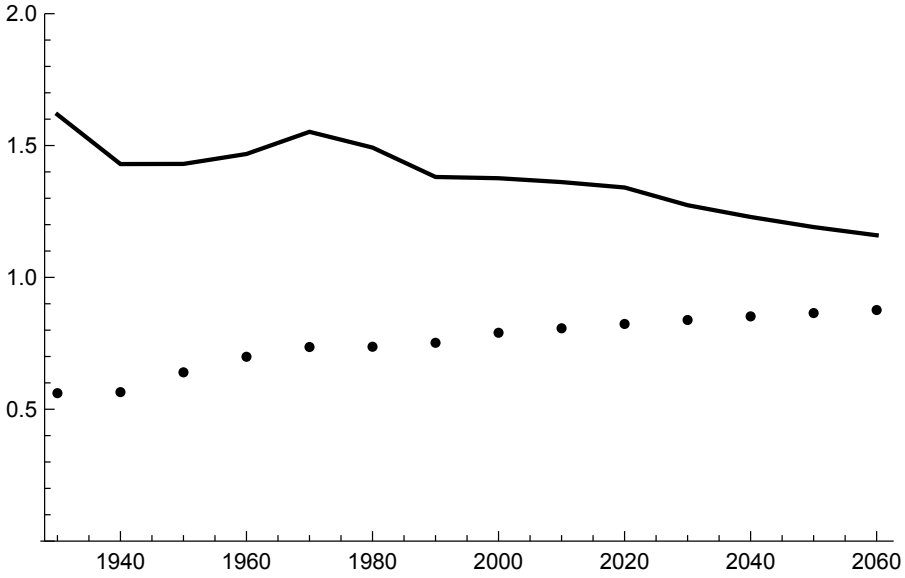
[29]Data comes from the National Income and Product Accounts (NIPA) from the Bureau of Economic Analysis. In the model there is no public debt; revenues and expenditure thus are equivalent measures of the size of government. To account for the absence of debt in the model we use the average of current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

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

Table 2: Calibration of extended baseline model

β	γ^1_{1950}	γ^2_{1950}	$\gamma^j_{1980}/\gamma^j_{1950}$
0.4954	0.5013	0.1086	1.3471

Figure 3: Demographics and urbanization



Population growth rate (solid) and urbanization (dots). Data from U.S. Census Bureau. Projections for population growth as reported by Census Bureau (middle series). Projections for urbanization interpolated based on United Nations (2014) forecast for 2050.

Based on this calibration we then let $\mathcal{F}_t = 0$ and solve for the counterfactual politico-economic equilibrium. We find that subject to the demographic structure in the year 1950, the relative share of the federal government drops from 72.6% (the actual value) to 40.1%. If we instead impose the demographic and regional structure in the year 1930 then the relative share of the federal government drops to 41.3% (compared with 38.3% in the data).[31] That is, independently of how exactly we define the counterfactual, switching off the general equilibrium channel of taxation explains almost all of the observed variation in the tax share of the federal government over the 1930s.

An alternative, more traditional explanation might motivate the rise in the federal government's revenue share with reference to a stipulated increase in static externalities from regional spending.[32] To assess the quantitative plausibility of this alternative mechanism we introduce static externalities in the extended baseline model and check by how much these static externalities would have had to increase to replicate the data.

Specifically, we assume that publicly provided services in region i are a function of (regional and federal) spending in other regions, in addition to the spending in the region itself,

$$g_t^i = (e_t^i + e_t) \cdot (\bar{e}_t + e_t)^\lambda,$$

where λ measures the strength of the cross-regional externalities. It is straightforward to verify that this modification leaves the equilibrium conditions unchanged except that a new term $\mathcal{E}_t$ enters into the optimality condition for federal taxes which captures the benefit from higher federal taxes and thus, the benefit from federal spending, that operates through cross-regional externalities,[33]

$$\mathcal{E}_t \equiv \left(\frac{\omega}{\nu_t} + 1\right) \frac{\lambda \bar{\gamma}_t}{\tau_t + \bar{\tau}_t}.$$

Assuming the parameter values displayed in table 2 and letting $\mathcal{F}_t = 0$, we find that the strength of static externalities, λ, would have had to increase strongly, from zero to more than 0.21, in order to have equivalent explanatory power as the general equilibrium effects, i.e. to generate an increase in the federal government's revenue share from 40.1% in 1930 to 72.6% in 1950. We consider this increase to be implausibly high and conclude that dynamic tax externalities in combination with the 16th Constitutional Amendment offer a more plausible explanation for the U.S. fiscal transformation.

[31] Intuitively, the higher share of rural regions in 1930 implies that decision makers at the federal level attach a stronger weight to those (rural) regions which benefit more strongly from federal taxation.

[32] Maybe the most plausible candidate in that respect would be public infrastructure investment to support major technological innovations. But many of 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. The shifts in the fiscal landscape thus should have been observed earlier. Even if spending externalities had increased around the 1930s, 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.

[33] Static externalities also affect the dynamic general equilibrium effects of labor income taxation (which we shut off for this counterfactual simulation). We now have $\mathcal{F}_t \equiv -\beta/(1-\tau_t-\bar{\tau}_t)\,(\epsilon_{Rk} + \epsilon_{wk}(1+\lambda)\bar{\gamma}_{t+1})$.

5.2 The Rise of Federal Grants

During the 1930s, federal grants started to gain importance as a source of funding for states. Since then, the share of federal grants in state and local revenues has continued to grow (see figure 1 in the introduction) and as a share of GDP, federal grants surpassed 0.5% in the 1950s and approached 2.8% in 2014. We now ask whether dynamic general equilibrium effects also can contribute towards explaining this feature of the data.

To study the determination of grants we assume that regional and federal spending are complements. Both the federal and the regional governments therefore must spend resources for public services to be provided, and efficiency calls for the two levels of government to spend resources in specific proportions. Federal grants serve the purpose to decouple the composition of government spending across levels of government from the composition of government revenue collection.

This decoupling is useful either if the net benefit of taxation, $\mathcal{F}_t$, renders it "cheaper" for the federal government to tax, or if it is advantageous to subsidize regional government spending due to benefits from cross-regional externalities that are reflected in $\mathcal{E}_t$. We find that the decoupling motive is sufficiently strong to explain most of the observed increase in grants while the model still explains about 30% of the increase in the relative size of the federal government during the 1930s.

5.2.1 Model with Spending Complementarities and Grants

As in the extended baseline model, the two groups of regions (which we index by j) are differentiated by their preference for public services, γ_t^j, and services in region i are a function of regional and federal spending as well as spending in other regions.[34] Letting $\vec{e}_t \equiv (e_t^1, e_t^2)$ collect spending in the two typical regions, publicly provided services in region i equal

$$g_t^i = a(e_t^i, e_t) \cdot A(\vec{e}_t, e_t)^\lambda.$$

Unlike in the extended baseline model where the spending components are perfect substitutes we now consider the case of complements,

$$a(e_t^i, e_t) = (e_t^i)^\delta (e_t)^{1-\delta} \quad \text{and} \quad A(\vec{e}_t, e_t) = \left(a(e_t^1, e_t)\right)^{\theta_t^1} \cdot \left(a(e_t^2, e_t)\right)^{\theta_t^2},$$

where $\delta \in (0, 1)$. In the working paper we discuss micro foundations for this specification (Gonzalez-Eiras and Niepelt, 2017).[35] We also provide detailed derivations for the general case with many types of regions and discussions of the equilibrium conditions.

The federal government pays a positive, uniform grant, x_t, to regional governments.[36]

[34] We allow for static externalities ($\lambda \neq 0$) in order to check the robustness of our findings but we set $\lambda = 0$ in the baseline calibration.

[35] We discuss constitutional restrictions that prescribe which services must be provided (but not necessarily financed) by regional or federal governments. The division could 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).

[36] In the working paper, we also analyze matching grants (Gonzalez-Eiras and Niepelt, 2017). We conclude that the model predictions are qualitatively unaffected by the type of grant.

We allow for proportional deadweight losses of grants at rate $1-\sigma \geq 0$.[37] Accordingly, condition (3) generalizes to

$$e_t = w_t(\tau_t - x_t), \quad e_t^i = w_t(\tau_t^i + \sigma x_t), \quad e_t^j = w_t(\tau_t^j + \sigma x_t), \ j = 1, 2, \tag{12}$$

and public services in region i equal

$$g_t^i = w_t^{1+\lambda}(\tau_t^i + \sigma x_t)^{\delta}(\tau_t - x_t)^{(1-\delta)(1+\lambda)} \prod_{j=1}^{2} (\tau_t^j + \sigma x_t)^{\delta\lambda\theta_t^j}.$$

The definition of equilibrium is modified in the obvious way and the political first-order conditions with respect to τ_t^j and τ_t, respectively, now read

$$\left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t^j \delta}{\tau_t^j + \sigma x_t} - \frac{1+\beta}{1-\tau_t-\tau_t^j} \leq 0, \ j = 1, 2, \tag{13}$$

$$\sum_{j=1}^{2} \theta_t^j \left\{ \left(\frac{\omega}{\nu_t} + 1\right) \frac{\gamma_t^j (1-\delta)}{\tau_t - x_t} - \frac{1+\beta}{1-\tau_t-\tau_t^j} \right\} + \mathcal{E}_t + \mathcal{F}_t \leq 0. \tag{14}$$

Condition (13) differs from condition (9) in the baseline model because the marginal benefit of taxation only depends on regional spending, reflecting the assumption of complementarities rather than substitutes (as well as logarithmic preferences). Condition (14) differs from condition (10) for the same reason. Moreover, it reflects both preference heterogeneity, as in the extended baseline model, and cross-regional externalities.[38]

In addition to (13), (14), and the complementary slackness conditions a first-order condition for grants holds in equilibrium:

$$\sigma\delta \sum_{j=1}^{2} \frac{\theta_t^j(\gamma_t^j + \lambda\bar{\gamma}_t)}{\tau_t^j + \sigma x_t} - \frac{(1-\delta)(1+\lambda)\bar{\gamma}_t}{\tau_t - x_t} \leq 0. \tag{15}$$

The first term reflects the benefit from higher regional spending and the second term represents the cost due to lower federal spending. Note that the degree of preference heterogeneity—not the average preference for public services—affects the choice of grants.[39]

In appendix B we prove that in this model with grants, the federal government always levies taxes while regional tax rates need not be positive unless grants equal zero. If a parametric condition is met (which is more likely when deadweight losses are positive, that is σ is small), all regions levy taxes and grants generically equal zero. If the condition is violated (which is more likely when static or dynamic externalities are positive, that is $\lambda > 0$ or $\mathcal{F}_t > 0$), then grants are strictly positive and fully crowd out taxes in the region with a low preference for public services. A mean preserving spread of the preference for public services also renders positive grants more likely.

[37] In addition to capturing resource costs of inter governmental transfers, the "deadweight losses" might serve as stand in for other frictions or specification errors, for instance related to the fact that the elasticity of substitution between federal and regional spending could differ from unity.

[38] Because of the static externalities, $\mathcal{F}_t \equiv -\beta/(1-\tau_t-\bar{\tau}_t)\,(\epsilon_{Rk} + \epsilon_{wk}(1+\lambda)\bar{\gamma}_{t+1})$. The spending complementarities affect the static cross-regional externalities; the latter now are given by $\mathcal{E}_t \equiv \lambda\bar{\gamma}_t(\omega/\nu_t + 1)(1-\delta)/(\tau_t - x_t)$.

[39] This follows from dividing (15) by $\bar{\gamma}_t$.

5.2.2 Model Predictions

In the baseline calibration, we assume no static externalities ($\lambda = 0$) and 7.5% deadweight losses ($\sigma = 0.9250$). To calibrate the preference for public services as well as β and δ, we use the political first-order conditions evaluated in 1950 and 2000. Specifically, we match the size of the federal government and of total government (federal, state, and local spending relative to GDP) as well as the GDP-share of grants in the year 2000, and the size of total government in the year 1950. The calibration does not impose any restrictions on the level of grants or the federal government's revenue share prior to the year 2000.

In principle, time variation in deadweight losses (σ); static externalities (λ); the importance of federal vs. regional spending (δ); or preferences (γ_t^j and θ_t^j) could explain the observed rise of federal grants (see proposition 2 in appendix B). Since there is little tangible evidence for variation of the former three factors we focus on changing regional preference disparities which we continue to associate with the shares of rural and urban regions. (Recall that preference heterogeneity—not the average preference for public services—affects the choice of grants. Changes in the average preference for public services therefore cannot explain a trend increase in grants.)

Our assumption that time variation in urbanization fostered the rise of grants also is supported by indirect evidence that blends the model with data on state level spending. Recall that the model predicts regions with a weaker preference for public services to choose a higher ratio of grants relative to regional tax revenue. If urbanization is positively correlated with the valuation of public services, as we argue, it should be negatively correlated with that ratio. This prediction is borne out in state level data over the period 1969 to 2008: A panel regression of the ratio of federal grants and direct general revenue of state and local governments on urbanization (and controls including state income per capita) yields the expected negative sign, see table 3.[40]

Table 4 displays the calibrated parameters. The β value corresponds to an annual discount factor of approximately 0.9838. The calibration for δ suggests an almost equal importance of federal and regional spending in the provision of public services. To meet the moment condition that rural areas do not levy taxes the calibration assigns a very low value to the preference for public services in these regions, at approximately 3% of the value in urban areas.[41] Accordingly, the model also predicts a counterfactually high ratio (7.0) of government spending in urban relative to rural areas, but introducing an exogenous component for regional tax collections resolves this problem.[42] Finally, to

[40] We use 2008 data to minimize measurement problems as a consequence of 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 rate is 100% in both periods.

[41] This feature is robust to assuming matching grants, see Gonzalez-Eiras and Niepelt (2017).

[42] In state level data from the Census Bureau, the ratio of government spending in urban relative to rural areas extends up to roughly 2.5. If we assume that regions exogenously collect a tax of 5% to fund spending unrelated to the provision of public services, then the predicted ratio of spending in urban relative to rural areas equals 2.5. This modification does not change the prediction of a trend increase in grants; in the modified model, grants peak at 5.3% of GDP at the end of the simulation horizon.

Table 3: Urbanization and grants

	Federal grants relative to state and local revenue
Urbanization	−0.3091
	(0.1834)
Income per capita	0.2948
	(0.2470)
State FE	YES
Time FE	YES
R^2	72.57
Observations	100

The table displays panel OLS regression results over the period 1969–2008 with federal grants relative to state and local revenue as the dependent variable. The explanatory variables are state-level urbanization and state income per capita relative to the national average. Sources: Federal grants relative to state and local direct general revenue 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). Relative state income per capita 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 4: Calibration of model with grants

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

replicate the increasing size of governments between 1950 and 2000, the model requires the preference for public services to grow at about 0.55% per year or 17.8% over thirty years. This increase is qualitatively consistent with Wagner's law and with the evolution over time of attitudes towards spending cuts, as reported in table 1.[43]

By construction, the calibrated model perfectly matches the level of total government in 1950 and 2000 as well as the federal share in tax collections in 2000. In contrast, the predicted evolution of the federal share prior to 2000 is not constrained by the calibration. For the post World War II period, the predicted share is nearly flat, compared to a slight decline in the data.[44] For the fiscal transformation phase (1930–1950), which is associated with slowing population growth, rising urbanization, and the switch from $\mathcal{F}_t = 0$ to $\mathcal{F}_t > 0$, the model predicts the share of the federal government to increase from 54.4% to 63.6%. That is, the model explains nearly 30% of the actual increase.[45]

As for federal grants, the model captures the long-term increase but not the short-run fluctuations (notably during the 1970s and the Great Recession), see figure 4. This reflects the fact that in the data, grants also are used for redistributive and risk sharing purposes which our model does not speak to. Importantly, the increase in grants since the 1930s reflects rising urbanization and thus, preference heterogeneity, not the stronger *average* preference for public services that the calibration imposes. This is particularly evident when we simulate the model subject to constant rural and urban shares (at their year 2000 values); the model then predicts a slightly negative trend for grants, see figure 4. In contrast to the important role played by urbanization, the changing demographics only are of minor importance. If we fix the population growth rate at its year 2000 value, the model predictions barely change.[46] Out of sample, the model predicts that grants

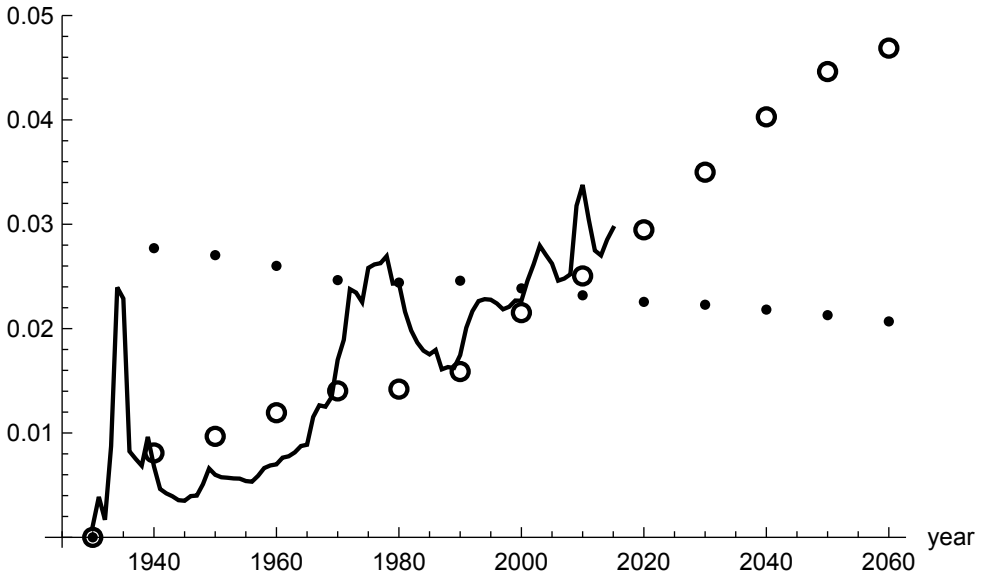
[43] An alternative explanation for the rising size of government could rely on public services being a luxury good, and higher incomes. Our assumption of logarithmic preferences rules out income effects on tax rates.

[44] Baicker et al. (2012) offer a potential explanation for the post World War II increase in the share of state governments.

[45] The difference in explanatory power across models is due to the fact that in the model with spending complementarities the direct marginal benefit of federal taxes depends only on decisions made by the federal government, while in the basic extended model this also depends on decisions made at the regional level (see conditions (14) and (11)). Thus, when regional governments in high valuation regions respond to a federal tax cut by raising taxes, in the basic extended model this dampens the effect that a tax change has on the direct marginal benefit of federal taxes. As a result the federal government in the basic extended model has to reduce taxes by more in order to account for a given shortfall in $\mathcal{F}_t$.

[46] This reflects the fact that population growth does not directly enter the equilibrium condition for

Figure 4: Federal grants, share of GDP



Data from NIPA (solid), model predictions (circles), model predictions with constant urbanization (dots).

continue to increase in the future up to approximately 4.9% of GDP in the year 2060.

The model predictions are robust to changes in all parameters except λ and σ. When static externalities are negative, $\lambda = -3\%$ say, or deadweight losses higher, $\sigma = 90\%$ say, then the predicted grants peak at between 2.8 and 2.9% of GDP between 2040 and 2050 before reverting back to lower values. When spending externalities are positive, $\lambda = 2\%$ say, or deadweight losses lower, $\sigma = 94\%$ say, then grants are predicted to increase to between 9.3 and 10% of GDP in 2060. Intuitively, with lower deadweight losses or higher static externalities, the federal government has a stronger incentive to provide grants (see proposition 2).

While the value for λ has a major effect on the model predictions an alternative specification that only allows for traditional, static externalities cannot plausibly explain the variation in the data as well as the model with dynamic tax externalities. When we re-calibrate the model under the assumption that the federal government does not perceive the dynamic general equilibrium effects of labor income taxation (or that they are not present) the results appear unreasonable, as in the previous subsection.[47] We interpret this as further evidence that our proposed explanation for the centralization of revenue and the use of inter governmental grants is more plausible than an alternative

grants, equation (15), but only indirectly through its effect on taxation.

[47]Very large static externalities are required to match the moments described earlier: Subject to the δ value reported in table 2, the calibrated λ value equals 0.3029 and the calibrated value for γ^2_{2000} is negative. When we instead eliminate the moment condition for grants in the year 2000 and impose for γ^2_{2000} the value reported in table 2 then the calibrated $\lambda = 0.1246$ and the predicted value for grants in the year 2000 exceeds the value in the data by 33%.

explanation that mainly focuses on spending externalities.

6 Concluding Remarks

What determines the degree of centralization of tax collections in a federal union? We propose a novel explanation that stresses differences in the perceived cost of taxation across levels of government due to dynamic general equilibrium effects. The dynamic externalities we emphasize complement static externalities that have traditionally been analyzed in the fiscal federalism literature, including spending externalities and externalities from horizontal or vertical tax competition.

When augmented with complementarities of government spending our model also offers an explanation for inter governmental grants. Grants have value because they allow to channel revenue from the federal government where tax revenue is "cheap" to regions, or because regions underspend as they do not internalize positive cross-regional externalities.

We find that dynamic general equilibrium effects can help explain the U.S. fiscal transformation during 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 the ratification of the Sixteenth Amendment, which opened the door for labor income taxation at the federal level, and higher demand for government expenditure, specifically New Deal policies and World War II spending.

Our simple framework abstracts from cross-regional insurance, redistribution, and many other features that are present in federalist states. Given this simplicity, its quantitative performance is reassuring. The baseline model fully accounts for the observed increase in the relative size of the federal government during the 1930s, and a plausibly calibrated extended model with spending complementarities accounts both for the trend increase in federal grants since 1930 and for roughly 30% of the compositional change during the 1930s.

Two extensions of the model appear to be of particular interest. First, the setup could be extended 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, for instance in the post-World-War II U.S. or 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 prices. As a consequence, the federal government might opt to employ grants (and deficits) to influence both regional taxes and deficits. This extension could address questions regarding debt and deficit policies in federal states.

A Elastic Labor Supply

We introduce additional taxes on labor income, levied at rates $\eta_t \geq 0$ and $\eta_t^i \geq 0$ by the federal and regional governments respectively, whose proceeds are reimbursed to workers.[48] The program of a worker in region i is given by

$$\begin{aligned} \max \quad & \ln(c_{1,t}) + v(l_t) + \gamma_t \ln(g_t^i) + \beta \left(\ln(c_{2,t+1}) + \gamma_{t+1} \ln(g_{t+1}^i) \right) \\ \text{s.t.} \quad & c_{1,t} = w_t(1-l_t)(1-\tau_t - \tau_t^i - \eta_t - \eta_t^i) + T_t^i - s_t, \quad c_{2,t+1} = s_t R_{t+1}, \end{aligned}$$

where T_t^i denotes the lump sum transfer to workers. In equilibrium, $\tau_t^i = \tau_t^j$ and $\eta_t^i = \eta_t^j$. Moreover, since preferences for consumption and leisure do not vary across regions, labor supply is constant across regions 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 in the baseline model, 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 regional and federal voters, $\mathcal{V}_t^i$ and $\mathcal{V}_t$ respectively, are

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

where V_t^i 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 $\eta_t^j = 0$.[49]

At the federal level, the first-order condition with respect to η_t is given by

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

If the equilibrium choice of η_t is interior, then $\partial l_t / \partial \eta_t > 0$; this implies that $dg(l_t)/dl_t = 0$. 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 (as well as by τ_t and τ_t^j).

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

[49] 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 $\eta_t + \eta_t^j > 0$.

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

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

The equality in the second line holds because, as shown above, either $dg(l_t)/dl_t = 0$ or $\partial l_t/\partial \tau_t = 0$ 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 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, $\mathcal{X}_t^{ls}$.

B Proposition 2

Proposition 2. In the general model:

(i) The federal government always levies taxes.

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

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

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 (16) holds, rendering grants more likely.

Proof. (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 $\delta \in (0,1)$, taxes cannot be zero since otherwise this term would diverge.

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

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

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

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 marginal benefit of regional taxation in that region 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 preference parameters reduces the left hand side of (16) (see also above inequality) since it increases the geometric average of $1/\gamma_t^j$. Similarly, it increases the right hand side of (16). A mean preserving spread therefore reduces the set of parameters for which (16) holds, rendering grants more likely. □

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.

Baicker, K., Clemens, J. and Singhal, M. (2012), 'The rise of the states: U.S. fiscal decentralization in the postwar period', *Journal of Public Economics* **96**(11–12), 1079–1091.

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

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. (2017), Fiscal federalism, grants, and the U.S. fiscal transformation in the 1930s, Discussion Paper 17-18, University of Copenhagen, Department of Economics.

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.

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.

Keen, M. J. and Kotsogiannis, C. (2004), 'Tax competition in federations and the welfare consequences of decentralization', *Journal of Urban Economics* **56**(3), 397–407.

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.

Rhode, P. W. and Strumpf, K. S. (2003), 'Assessing the importance of Tiebout sorting: Local heterogeneity from 1850 to 1990', *American Economic Review* **93**(5), 1648–1677.

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.

Taussig, F. (1910), *The Tariff History of the United States*, 5. edn, G.P. Putnam's Sons, New York and London.

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.

United Nations (2014), *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)*, Department of Economic and Social Affairs, Population Division, New York.

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.