Journal of Public Economics 239 (2024) 105240

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

The impossible trinity: Competitive markets, free entry, and efficiency[☆]

Halvor Mehlum[a], Gisle J. Natvik[b], Ragnar Torvik[c,d,*]

[a] University of Oslo, Department of Economics, P.O. Box 1095 Blindern, N-0317 Oslo, Norway
[b] BI Norwegian Business School, Department of Economics, 0442 Oslo, Norway
[c] Norwegian University of Science and Technology, Department of Economics, N-7491 Trondheim, Norway
[d] BI Norwegian Business School, CAMP, 0442 Oslo, Norway

ARTICLE INFO

JEL classification:
P16
P48
D72
H11

Keywords:
Political economy
Endogenous political interests
Size of government
Efficiency and popular support

ABSTRACT

We present a model in which workers make occupational choices and vote over a tax rate which determines the level of government spending. Workers in occupations whose services are in high (low) demand by the government favor high (low) taxes. We show that the socially efficient size of the public sector cannot be supported in a political economic equilibrium. The reason is that equilibrium tax rates always reward excessive entry into the politically most powerful sector, and thus the equilibrium size of government is always either too big or too small. We show that this is an example of a more general political economy result that extends well beyond the baseline model and holds quite generally: the combination of (i) competitive markets and (ii) free entry is inconsistent with (iii) allocative efficiency.

1. Introduction

A tenet of economics is that scarcity invites entry, as factors in short supply are generously remunerated. The mechanism is a desirable one: with free entry and competitive markets, the marginal productivities in different activities are equalized, ensuring allocative efficiency. In this paper we argue that this basic economic insight needs to be reconsidered when policy depends on popular support. Let, for instance, the political influence of citizens in a particular activity be increasing in their relative size. Factor allocations then not only affect market prices, but also economic policy. By implication, being one of the few comes with a cost, as policies are tilted toward majority interests. On the one hand, scarce activities yield high income. On the other hand, entering a scarce activity entails joining the politically weak. While the first incentive promotes efficiency, the second does not. As a consequence, when policy depends on popular support, the combination of (i) competitive markets and (ii) free entry cannot yield (iii) allocative efficiency.

In order to transparently develop our impossibility result and convey its empirical relevance, we start by studying the familiar question of what determines the size of government in the economy. First, we establish the optimal government size within a simple framework similar to the influential model of Barro (1990). Then, we study the consequences of introducing competitive markets and free entry when

☆ This paper has circulated under various titles. We are grateful to editor Jonathan Heathcote for suggestions, for inputs from two referees, and for comments and suggestions from Gani Aldashev, Erling Barth, Tim Besley, Christian Bjørnskov, Matilde Bombardini, Steven Callander, Francesco Caselli, Matteo Cervellati, Sylvain Chassang, Raquel Fernandez, Patrick Francois, Cecilia Garcia-Penalosa, Pawel Gola, Colin Green, Rune Jansen Hagen, Magnus Hatlebakk, Andreas Stranden Hoel-Holt, Sergei Izmalkov, Francesco Lippi, Laura Mayoral, Omer Moav, Espen Moen, Karl Moene, Plamen Nenov, Andy Newman, Nathan Nunn, Tone Ognedal, Torsten Persson, Jørn Rattsø, Debraj Ray, James A. Robinson, Michael Ross, Rajiv Sethi, Kjetil Storesletten, Shaun Vahey, Frederick van der Ploeg, and seminar participants at BI Norwegian Business School, CEPR Macroeconomics and Growth Conference (London School of Economics), EEA Conference (Bocconi University), New Economic School Moscow, Norwegian University of Science and Technology, Oslo Macro Group, Reserve Bank of Australia, ThReD Conference (Boston University), University of Bari, and University of Oslo. This paper is part of the research activities at the Centre for Applied Macroeconomics and commodity Prices (CAMP) at BI Norwegian Business School, and the centre of Equality, Social Organization, and Performance (ESOP) at the Department of Economics at the University of Oslo. CAMP and ESOP are supported by the Research Council of Norway.

* Corresponding author at: Norwegian University of Science and Technology, Department of Economics, N-7491 Trondheim, Norway.
E-mail addresses: halvor.mehlum@econ.uio.no (H. Mehlum), gisle.j.natvik@bi.no (G.J. Natvik), ragnar.torvik@ntnu.no (R. Torvik).

https://doi.org/10.1016/j.jpubeco.2024.105240
Received 12 June 2023; Received in revised form 6 June 2024; Accepted 16 September 2024

government size is determined by voting. This allows us to derive our impossibility result within a well-known environment.

We consider a model where citizens initially face an occupational entry decision. Either they specialize as a service sector worker, or they enter as a goods sector worker. Equilibrium with free entry involves an arbitrage condition where, on the margin, payoffs from either alternative are equal. Given this standard arbitrage condition, no agent regrets his or her occupational choice. Under well-known conditions, this is also the socially optimal allocation of workers between the two occupations. Now, introduce politics. Assume that relative to private demand, tax financed public demand is more intensive in services than in goods. Then, a higher tax rate to finance higher public consumption shifts relative demand towards services and away from goods. Absent instant and perfect mobility between occupations, such a shift increases wages in the service occupation and decreases wages in the goods occupation. Let citizens in one of the occupations have more political influence than citizens in the other, be it the majority or the minority group. Let us start out in a situation with the optimal size of the public sector (i.e. the optimal tax rate) and the corresponding optimal allocation of workers. Given that the envelope theorem applies, then at this tax rate and allocation of workers, those in the service sector will prefer a higher tax rate than the optimal one, while those in the goods sector will prefer a lower tax rate. The group with most political influence will thus shift the tax rate in their preferred direction. But then the standard arbitrage condition above, assuring allocative efficiency, is no longer fulfilled. The socially optimal size of the public sector, with the corresponding optimal occupational allocation, cannot constitute an equilibrium.[1] Government is either too big or too small. The above explains why (i) competitive markets, (ii) free entry, and (iii) allocative efficiency constitute an impossible trinity.[2]

The claim we make above is that the efficient allocation is *not* an equilibrium. The natural follow-up question then becomes: What *is* an equilibrium? In the Barro-model with entry and voting, we show how small differences in preferences over goods and services can translate into large differences in government size. In fact, even countries with identical fundamentals may end up with very different government sizes. If a majority enters into services, then taxes will be high, supporting services as the majority sector in equilibrium. Conversely, if a majority enters into the goods sector, taxes will be low, supporting goods as the equilibrium majority sector. Thus, our model provides a new perspective on a well-known observation: countries often have very different sizes of their public sector, despite being similar in many other dimensions such as income levels.

Compare France and the UK, for example. The two countries have approximately the same level of GDP per capita. But, based on OECD data from 2019 (i.e. before the pandemic), general government spending is 55,4% of GDP in France while it is 40,9% in the UK. Our impossibility result suggests a new, and in our view plausible, explanation for these differences. We suggest that the expected payoff from working in the public sector in France is high exactly because the public sector is big, making public employees powerful enough to support policies that bolster their own remuneration. In the UK, by contrast, it could be less tempting to aim for a public sector career exactly because the public sector is small, making public employees politically weak. In other words, it could be more attractive to enter the public sector in those countries where there is an abundance of public sector employees, not scarcity.

The novelty of our approach is that we study the role of free entry into economic activities in political equilibrium. This is a main and necessary driver of our mechanism, and also what distinguishes our contribution from existing models of inefficient policy. There are, however, a number of literatures that our paper relates to. In the example above, as well as in our general impossibility result, a crucial feature is that voters cannot commit to vote for a future policy which is against their own interest (when the future arrives). Thus our result relates to the more general literature on political economy, of which Acemoglu (2003, p. 622) asks

> "why do politicians and powerful social groups not make a deal with the rest of the society to choose the politics and institutions that maximize output or social welfare, and then redistribute parts of the gains to themselves?"

He goes on to argue that the problem with such a solution is (p. 622) "its applicability is limited because of inherent commitment problems associated with political power." In our model, if voters could commit to policy ahead of their entry into economic activities, then entry would ensure that also the (endogenous) factor allocation became optimal and the trinity would be feasible (as we show below). Our model is thus related to the large literature focusing on the lack of commitment and time inconsistency starting with Kydland and Prescott (1977). However, in contrast to this literature, it is not the politicians, but *the voters* who cannot commit to their own future political behavior. For this reason our mechanism can also reverse some of the results in the previous literature. To see why, consider the well-known example of capital taxation where a policymaker cannot commit to keeping taxes low once capital is in place. Thus investment suffers, and the capital stock becomes smaller than its socially optimal level. But what if those entering as investors become so politically powerful that they are able to tilt policy in their preferred direction? Then the endogenous entry has created a situation where low capital taxes may constitute a political equilibrium. Moreover, too low capital taxes may attract more investors, cementing this equilibrium. Allowing entry that renders voter characteristics endogenous may in this way turn the previous prediction of too high capital taxes on its head.[3]

Our emphasis on economic entry naturally relates our paper to theories of political entry, in particular the citizen-candidate model of Osborne and Slivinski (1996) and Besley and Coate (1997, 1998). As in their setting, endogenous entry may result in multiple equilibria where no-arbitrage conditions are satisfied, because entry decisions are strategic and depend on what others do. In our setting it is the endogenous entry of voters into economic activity, not of the politicians, that drives our results, and for this reason the welfare implications differ from those of the citizen-candidate models. In particular, in citizen-candidate models the equilibrium can be socially efficient, while in our approach it cannot.

As regards entry of voters, our paper is related to those of voter mobility originating from the work of Tiebout (1956). A main difference is that in this literature entry into some jurisdictions is driven by exogenous differences in voter interests, while in our model voters' political interests are endogenous to entry.[4] Models of social mobility, in particular such as those of Benabou and Ok (2001), Hassler et al.

[1] In our model, we assume workers in the majority occupation decide. But, as this example reveals, our impossibility result would equally well hold if policy is decided by the minority. All that is needed for the result to go through, is that one group has more political influence than the other. See also footnote 23.

[2] As we discuss in Sections 3.3 and 4 the impossibility result holds under more general assumptions than in this simple example. For instance, it is not necessary that agents choose their occupation once and for all, as we show below.

[3] Consider the classical study by Rosen and Rosen (1980) of how favorable tax treatment of owner-occupied housing stimulates homeownership. Our approach would imply that the extent of homeownership is not only a consequence of the tax system, but also a cause.

[4] For this reason the welfare implications are also very different. See chapter 8 in Drazen (2000) for a detailed discussion of welfare implications in the literature on fiscal federalism.

(2003), Benabou and Tirole (2006), and Acemoglu et al. (2018), also study voters' entry into different groups. Also in these models policy preferences shift when agents transition from one social group to another. However, in these models the emphasis is on how and why payoffs differ between groups.[5]

The contributions most closely related to our particular application are probably two papers with discussion of public bureaucracies by Tullock (1974) and Buchanan and Tullock (1977). In the first of these papers, Tullock notes that as the number of bureaucrats increases (p. 129).

> "it would be possible to use more and more of their power to directly increase wages. In a sense, the individual bureaucrat tries to increase his wages, but realizes that there are political gains from increasing the number of bureaucrats in that he will be able to have more political power to increase his wage in the next period. Expansion becomes a sort of investment."

In Buchanan and Tullock (1977) this view is developed further, linking it to the voting patterns of public sector employees under the heading "Wagner Squared". When the share of the public sector increases with economic growth, for instance because public goods and services have an income elasticity that exceeds unity (Wagners law), then, according to Buchanan and Tullock (1977, p. 148), as

> "[...] the bureaucracy members come to make up a larger and larger share of the total voting constituency, the possibility of the usage of civil servant voting power to expand salaries directly becomes real."

Hence, similar to us Buchanan and Tullock point out that the political interests of agents are shaped by their occupations,[6] and that these interests can be more forcefully represented the larger the occupational group. In contrast to us, however, Buchanan and Tullock analyze only one group of voters and do not observe that the mechanism they discuss might equally well give the result that the public sector can become too small. Moreover, Buchanan and Tullock do not consider the arbitrage condition as a source of inefficiency.[7]

The rest of our paper is organized as follows. In Section 2 we present our baseline model of occupational choice, voting, and government size, and characterize the social optimum. Section 3 derives the political equilibrium and provides our impossibility result. In Section 4 we generalize the impossibility result.

2. A model of government size

In this section we present our model on government size, and characterize the socially optimal allocation in this environment.

2.1. Preferences, technology, and institutions

We consider a society with a continuum of citizens of measure normalized to 1. Citizens choose to enter as workers in one out of two sectors, where they inelastically supply one unit of labor. We simply term the two sectors the goods sector and the services sector.[8]

Citizens have preferences over private consumption of goods and services, and the provision of public services which is supplied by the services sector. Citizen i derives utility according to

$$U_i = c_{N,i}^{\alpha} c_{T,i}^{\beta} g^{\gamma}, \quad \alpha + \beta + \gamma = 1, \tag{1}$$

where $c_{N,i} \geq 0$ denotes i's consumption of private services (N), $c_{T,i} \geq 0$ denotes the consumption of goods (T), where g is the flow of public services provided by government, and where all exponents are strictly positive. With a slight abuse of notation, we utilize that in equilibrium citizens' consumption will differ only as a result of the sector they supply labor to and let U_j represent the utility of a typical citizen in the workforce of sector $j \in \{N, T\}$. Similarly, $c_{N,T}$ will denote consumption of private services by a typical citizen working in the goods sector, and so on.

We let l_N and l_T denote the mass of citizens who constitute the services and the goods workforces, respectively. All citizens supply labor to one of the two sectors, hence $l_N + l_T = 1$. Goods is the numeraire. Denote the price of private services relative to goods by p, and wages in terms of goods in the two sectors by w_N and w_T, respectively. The budget constraint for each citizen in the workforce of sector j is

$$pc_{N,j} + c_{T,j} = (1 - \tau) w_j, \quad j \in \{N, T\}, \tag{2}$$

where τ is a proportional income tax rate. The fact that taxation is proportional to income does not cause distortions because citizens inelastically supply their one unit of labor.[9] One unit of services sector labor produces one unit of services, and one unit of goods sector labor produces one unit of goods. Public services g are purchased from the services sector.[10] Thus, services and goods available for private consumption, denoted by x_N and x_T respectively, are given by

$$x_N = l_N - g, \tag{3}$$

and

$$x_T = l_T. \tag{4}$$

[5] At an abstract level our idea resonates with Acemoglu et al. (2005), who study why England and the Netherlands diverged economically and politically from Spain and Portugal with the discovery of the New World. They emphasize different entry conditions, where in the two former countries entrepreneurs were allowed to take part in the new trade to a much larger extent than in the two latter, where these possibilities were monopolized and regulated by the crown and its allies. In turn, entry of new entrepreneurs in England and the Netherlands made this group more politically powerful, in turn allowing them to tilt institutions in favor of more secure property rights. Although their focus and results are very different from our approach, our model does share the property that entry affects factor allocations, and, more importantly, that factor allocations in turn affect the balance of political power.

[6] Lindbeck (1995) points out a similar mechanism in that (p. 14) "An unwinding of welfare-state spending could be expected to be particularly difficult in societies where a large share of the electorate is financed by the public sector (i.e. is tax financed rather than market-financed.)" Similarly, Christoffersen and Paldam (2003) develop the concept of "the welfare coalition" to describe such a situation. The conjecture that individual occupation causally affects policy preferences finds empirical support in e.g. Rattsøand Sørensen (2016).

[7] Determinants of the size of the public sector have been extensively studied in the political economy literature. In Meltzer and Richard (1981) the size of the public sector is determined by income inequality, where high income inequality produces a high tax rate. Persson et al. (1997, 2000a) show how separation of powers influences the size of the public sector, and how this can be interpreted as differences between characteristics of political institutions, such as if there is a presidential or parliamentary system, or if the election system is proportional or majoritarian (on this see also Lizzeri and Persico, 2001; Milesi-Ferretti et al., 2002). In these theories, in contrast to ours, voter characteristics are exogenous and there is no entry of voters, which is the driving mechanism in our model. For a more complete review of the political economy literature on government size, see several of the chapters in Persson and Tabellini (2000b) or Besley (2006), or chapter 14 in Drazen (2000), which is entirely devoted to this issue.

[8] These two sectors could also be the private and the public sector, traded and non-traded and so on. In the Appendix we consider a continuum of sectors.

[9] We assume proportional, rather than lump sum taxes, for analytical convenience.

[10] In the Appendix where we extend the model to a continuum of sectors, public services may be purchased from all sectors.

Each sector j clears when

$$x_j = l_N c_{j,N} + l_T c_{j,T}. \quad (5)$$

There is perfect (Bertrand) competition in both markets. Given linear production technologies profits are zero, and wages are simply determined to equal the value of the marginal productivity of labor. Since the marginal productivity of labor is unity in both types of production, wages in a sector are always equal to prices in the sector. Wages thus satisfy

$$w_N = p, \quad (6)$$

and

$$w_T = 1. \quad (7)$$

We return to the equilibrium determination of p below.

Turning next to the political decision, this simply regards the tax rate and, by implication, the amount of the public services. Each citizen votes for a preferred tax rate $\tau_j \in [0,1]$, where the subscript $j \in \{N,T\}$ indicates that the voting decision depends on the worker's sector. Any tax rate that receives a majority of votes is implemented, and we denote the implemented tax rate by τ.

The public-sector budget constraint reads

$$pg = \tau\left(l_N w_N + l_T\right), \quad (8)$$

where we have already incorporated Eq. (7). Note that the tax rate not only determines the provision of public services, but, as will become clear, also affects the supply and demand of private goods and services.

2.2. Social optimum

With linear utility, any distribution of consumption between different citizens is consistent with a social optimum. Hence, the distribution between citizens can be ignored here. Let g^o denote the first-best level of public services, c^o_N the first-best level of services, and so on. We then have:

Proposition 1. *The socially optimal allocation satisfies*

$$g^o = \gamma, c^o_N = \alpha, c^o_T = \beta, l^o_N = \alpha + \gamma, l^o_T = 1 - l^o_N = \beta.$$

Proof. Under inelastic supply of labor and positive marginal utility of consumption, a socially optimal labor allocation implies that all labor is used for production. Given the unit labor requirement in all production technologies, we can write the maximization problem of a social planner as

$$\max_{[g,l_N]} \left(l_N - g\right)^\alpha \left(1 - l_N\right)^\beta g^\gamma.$$

The two first-order conditions w.r.t. g and l_N, respectively, read

$$\frac{\alpha}{l_N - g} - \frac{\gamma}{g} = 0, \quad (9)$$

and

$$\frac{\alpha}{l_N - g} - \frac{\beta}{1 - l_N} = 0. \quad (10)$$

By solving (9) for l_N, and using the resulting expression to substitute for l_N in (10), we obtain $g(\alpha+\beta+\gamma) = \gamma$. Because $\alpha+\beta+\gamma = 1$, the first part of the proposition follows; $g^o = \gamma$. The labor allocation $l^o_N = \alpha + \gamma$ then follows by inserting $g^o = \gamma$ into (9), while (10) implies $1 - l^o_N = \beta$. Finally, given production, the consumption levels follow. ■

Note that the first-best allocation is a natural benchmark with which to compare the political equilibrium in this model. It coincides with that implemented by a Ramsey planner choosing the tax rate to maximize welfare subject to the competitive equilibrium, as we clarify in Section 3.5. The reason is that in this model there is no need to distort the allocation of resources to provide the public good.

3. Competitive and political equilibrium

In this section we characterize the possible political economic equilibria in our model on government size.[11] After outlining the timing of the stages in our model, we characterize equilibrium outcomes in three scenarios. First, we assume that after entry there is occupational immobility, as reflected in the timing of events specified in Section 3.1. Thereafter, we study the case where workers can switch occupation. Finally, we show what happens if we allow for once-and-for-all voting before entry.

3.1. Timing of events and equilibrium concept

This is a static model with four stages. To summarize, the timing of these stages is as follows:

1. Each citizen undertakes their occupational choice, i.e. decides in which sector to enter.
2. Each citizen votes for a tax rate. The tax rate that receives a majority of votes is implemented.
3. Each citizen supplies one unit of labor to their sector.
4. Production, prices and wages are determined. Each citizen gets their income, and derives utility from private consumption and public services.

A strategy for citizens simply determines their choice of sector, voting over the tax rate, and their consumption decisions. A subgame perfect equilibrium (SPE) is defined as a strategy profile in which all actions are best responses to other strategies at all stages of the game. Since we have many voters, the set of SPEs involves a large number of equilibria in which voters use weakly dominated strategies, such as voting for a tax rate that is not preferred because a majority of other voters are doing so. To rule out such unreasonable equilibria we focus on (pure-strategy) SPEs in undominated strategies. In our setting, where voters in each group will all have the same expected utility, and where there are only two groups of voters, this will imply that in equilibrium each citizen simply votes for their most preferred tax rate.[12]

3.2. Occupational immobility

We solve for the model's SPEs by backward induction. We start with a citizen in a given sector, facing a given tax rate, given prices, and a given net income, and characterize consumption choice. Thereafter, we characterize the voting decision, given the occupational choice of a citizen. After this characterization, we go to the first stage of the game, where we determine occupational choice, i.e. entry into the services and goods sectors. Finally, we contrast the possible SPEs with the social optimum.

Preliminaries

All citizens maximize (1) subject to (2), taking goods prices, wages, and the tax rate as given. The resulting consumption demands are

$$c_{N,j} = \frac{\alpha}{(\alpha+\beta)p}(1-\tau)w_j, \quad j \in \{N,T\}, \quad (11)$$

and

$$c_{T,j} = \frac{\beta}{\alpha+\beta}(1-\tau)w_j, \quad j \in \{N,T\}. \quad (12)$$

[11] Note that our impossible trinity, as a statement about what cannot be an equilibrium, is valid under fairly general conditions, as we show in Section 4. What actually can constitute an equilibrium, however, is more dependent on the precise model considered.

[12] We also adopt the convention that if two tax rates receive the same amount of votes, the tax rate is decided by the tax rate preferred by a majority of goods workers. This has no bearing on our results, and only works to simplify notation.

From (4), (12), (5) it follows that in the goods sector, supply equals demand when

$$l_T = \frac{\beta}{\alpha+\beta}(1-\tau)\left[l_N w_N + l_T w_T\right].$$

Utilizing $w_T = 1$ and $w_N/w_T = p$, we may conveniently express this market clearing condition as

$$p\frac{l_N}{l_T} = \frac{\alpha+\tau\beta}{\beta(1-\tau)}, \tag{13}$$

which will be central in what follows.

Next, after combining (11), (12) and (8) with (1), we observe that the utility citizens finally enjoy in any equilibrium is

$$U_j = \Phi(1-\tau)^{1-\gamma}\tau^\gamma\left(w_j\right)^{1-\gamma}\left(l_N + \frac{l_T}{w_N}\right)^\gamma p^{-\alpha}, \tag{14}$$

where $\Phi \equiv \frac{\alpha^\alpha\beta^\beta}{(\alpha+\beta)^{\alpha+\beta}} > 0$. The term $(1-\tau)^{1-\gamma}\tau^\gamma$ reflects the same trade-off as in Barro (1990), regarding the size of a public sector. On the one hand, public services directly increase utility. On the other hand, their financing is costly in terms of private goods foregone. Maximization of this term alone gives the Barro result that the optimal size of a public sector entails $\tau = \gamma$.

Expression (14) provides an important insight: the direct effect of taxes (τ) on a citizen's utility is independent of her sectoral attachment (j). Hence, the only sources of conflict regarding preferred government size, are the indirect effects of taxes through prices and quantities in the labor market (w_j, l_j).

Market Clearing for Given Taxes and Sectoral Labor Supplies

Under competitive markets the equilibrium involves wages determined by (6) and (7). Clearing of the goods market as expressed by (13) implies a relationship between the equilibrium relative price and the pre-determined tax rate and work-force composition:

$$p = \frac{l_T}{l_N}\frac{\alpha+\beta\tau}{\beta(1-\tau)}. \tag{15}$$

Intuitively, a higher tax rate shifts demand in the direction of service sector labor and appreciates the relative price of services for given l_N and l_T.

Voting over the Tax Rate

Having characterized equilibrium outcomes for a given workforce composition and tax rate, we now analyze the preceding stage of the game: the voting over taxes. For that purpose, it is useful to first find workers' indirect utility functions over taxes.

From (6), (7), and (14), it follows that services and goods workers obtain the utilities

$$U_N = \Phi(1-\tau)^{1-\gamma}\tau^\gamma\left(l_N + \frac{(1-l_N)}{p}\right)^\gamma p^\beta, \tag{16}$$

and

$$U_T = \Phi(1-\tau)^{1-\gamma}\tau^\gamma\left(l_N + \frac{(1-l_N)}{p}\right)^\gamma p^{-\alpha}. \tag{17}$$

Comparing the two expressions, we note that the indirect utility functions are nearly identical. The only difference arises in the last terms containing p on the right hand sides of (16) and (17). These terms reveal that there is a conflict of interest between workers in the two sectors: a relative price increase always benefits workers in the services sector more than workers in the goods sector. The reason is that a higher relative price shifts the income distribution toward services workers. Note that this conflict of interest carries over to taxation, as the relative price characterized by (15) is monotonically increasing in τ. Consequently, the optimal tax rate from the point of view of services workers, τ_N, is always higher than the optimal tax rate from the point of view of goods workers, τ_T.[13] Moreover, the preferred tax rates will lie on each side of the first-best tax rate, γ. This allows us to establish the following lemma:

Lemma 1. *The optimal tax rates for N-workers (τ_N) and T-workers (τ_T), exist, are unique, and satisfy* $1 > \tau_N > \gamma > \tau_T > 0$.

Proof. After inserting (15) into (16) and (17), the indirect utility functions over taxes may be compactly expressed as

$$U_N \equiv U_N(l_T, l_N, \tau) = \Lambda_N(1-\tau)^\alpha\tau^\gamma(\alpha+\beta\tau)^{\beta-\gamma}, \tag{18}$$

where $\Lambda_N = \Phi\frac{(l_T)^\beta}{(l_N)^{\beta-\gamma}}\frac{(\alpha+\beta)^\gamma}{\beta^\beta}$, and

$$U_T \equiv U_T(l_T, l_N, \tau) = \Lambda_T(1-\tau)^{2\alpha+\beta}\tau^\gamma(\alpha+\beta\tau)^{-\alpha-\gamma}, \tag{19}$$

where $\Lambda_T = \Phi\frac{(l_N)^{\alpha+\gamma}}{(l_T)^\alpha}(\alpha+\beta)^\gamma\beta^\alpha$. Differentiating (18) and (19) with respect to τ, yields:

$$\frac{dU_N}{d\tau} = U_N\left[\frac{\gamma}{\tau} + \frac{\beta(\beta-\gamma)}{(\alpha+\beta\tau)} - \frac{\alpha}{(1-\tau)}\right], \tag{20}$$

and

$$\frac{dU_T}{d\tau} = U_T\left[\frac{\gamma}{\tau} - \frac{\beta(\alpha+\gamma)}{(\alpha+\beta\tau)} - \frac{(2\alpha+\beta)}{(1-\tau)}\right]. \tag{21}$$

It immediately follows that $\lim_{\tau\to 0} dU_j/d\tau > 0$ and $\lim_{\tau\to 1} dU_j/d\tau < 0$, for $j = N, T$. Hence, the optimal tax rates τ_N and τ_T both lie in the interval $\langle 0,1\rangle$. Because both U_N and U_T are differentiable over $\tau \in \langle 0,1\rangle$, it follows from (20) and (21), as well as utilizing that $\gamma = 1-\alpha-\beta$, that τ_N and τ_T satisfy the first-order conditions:

$$\tau_N = \max\left\{\tau : -\beta\tau^2 - \left(\frac{\alpha}{\alpha+\beta} - \beta\right)\tau + \frac{(1-\alpha-\beta)\alpha}{\alpha+\beta} = 0\right\} \in \langle 0,1\rangle, \tag{22}$$

and

$$\tau_T = \max\left\{\tau : -\beta\tau^2 - \left(\frac{\alpha}{\alpha+\beta} + \alpha\right)\tau + \frac{(1-\alpha-\beta)\alpha}{\alpha+\beta} = 0\right\} \in \langle 0,1\rangle. \tag{23}$$

Note that (22) and (23) both contain parabolas which have two roots and since $\lim_{\tau\to 0} dU_j/d\tau > 0$ and $\lim_{\tau\to 1} dU_j/d\tau < 0$, one of the roots, which is also an optimum, is on the interval $\langle 0,1\rangle$. The parabola in (22) is symmetric around $\tau = -\alpha(1+\alpha+\beta)/(2(\alpha+\beta)\beta) < 0$ while the parabola in (23) is symmetric around $1/2 - \alpha/(2\beta(\alpha+\beta)) < 1/2$. Therefore, in both cases, only the larger of the two roots is in the interval. It therefore follows that both (22) and (23) have one, and only one, solution at the interval $\langle 0,1\rangle$. Thus the optimal tax rates τ_T and τ_N are unique.

Evaluation in $\tau = \gamma$ implies that $\frac{dU_N}{d\tau} > 0$ while $\frac{dU_T}{d\tau} < 0$. As $\frac{dU_N}{d\tau} > 0$ evaluated in γ, the optimal τ_N must be larger than γ. As $\frac{dU_N}{d\tau} < 0$ evaluated in γ, the optimal τ_T must be smaller than γ. Hence, $0 < \tau_T < \gamma < \tau_N < 1$, which completes the proof. ■

The intuition for this result is that a higher tax rate brings a greater provision of public services. This shifts demand (at given wages and prices) for services up. To re-establish an equilibrium with less services available for private consumption, the price of services, and the wage for services labor, has to increase. In the new equilibrium, a higher tax rate is therefore associated with a relative price increase. A relative price increase is, viewed in isolation, advantageous for services workers because their (pre-tax) consumer real wage increases, while it hurts the

[13] Our model is one of direct democracy in that citizens vote over the tax rate directly. An alternative interpretation would be that there are two political parties each representing one of the two occupational groups, where the two parties run on platforms with τ_N and τ_T as their tax rates.

goods workers since their (pre-tax) consumer real wage decreases. For this reason, the services workers always prefer a higher tax rate, and a larger public services provision, than the goods workers.

Note also that from (22) and (23) the preferred tax rates τ_N and τ_T are independent of how citizens are allocated across sectors, given by l_N and l_T (yet to be determined). This property has less generality than the result stated in Lemma 1, as it rests on the utility functions assumed, but it still provides useful intuition. Behind it lie two countervailing forces that cancel out exactly in the Cobb–Douglas case. On the one hand, a higher share of services workers pulls toward a higher preferred tax rate, as relatively more resources are available to produce public services. On the other hand, a higher share of services workers allows for more private consumption of services, which pulls the preferred tax rate down.

We can now determine which tax rate that ultimately is implemented. Due to our restriction to weakly undominated strategies, in equilibrium voters simply vote for the tax rate they prefer. Thus, the political equilibrium tax rate, τ, is decided by the majority:

$$\tau = \begin{cases} \tau_N \text{ if } & l_N > 1/2 \\ \tau_T \text{ if } & l_N \leq 1/2. \end{cases} \tag{24}$$

Occupational Choice

We now turn to the first stage of the game where l_N and l_T are determined by citizens' occupational choice. Any equilibrium must imply that no citizen regrets his or her occupational choice, given the policy that will eventually be decided. Hence, absent corner solutions the occupational decision must imply $U_N = U_T$, where utilities follow from Eqs. (16) and (17).[14] As seen from these equations, the indifference condition boils down to $p = 1$. From (6) and (7) this condition in turn means that $w_N = w_T$. Inserting $p = 1$ in (15), and solving with respect to l_N, we obtain

$$l_N = \frac{\alpha + \beta\tau}{\alpha + \beta}. \tag{25}$$

Thus, the fraction of citizens entering the services sector increases with the equilibrium tax rate. A high tax rate implies high demand for services relative to goods workers, which (all else equal) makes it relatively attractive to enter the service sector.

Equilibrium

Taking into account that the choice of taxes and occupations must satisfy Eqs. (24) and (25) in equilibrium, we can now characterize the possible SPEs.

First, from Eq. (25) we directly observe that $l_N > 1/2$ if τ is sufficiently high. Moreover, the tax choice (24) implies that if $l_N > 1/2$, then $\tau = \tau_N$. Hence, there will exist a threshold tax level $\overline{\tau}$ such that if all citizens prefer taxes above this rate, then $l_N > 1/2$ and N-workers decide the tax rate. From (25), this threshold tax level is $\overline{\tau} = (\beta - \alpha)/2\beta$. Moreover, because we have established that $\tau_N > \tau_T$, a sufficient condition for $l_N > 1/2$, is that $\tau_T > \overline{\tau}$. Hence, if $\tau_T > \overline{\tau}$, the equilibrium is unique with $\tau = \tau_N$ and $l_N > 1/2$. From (23) and the constraints on α and β it follows that this is satisfied if and only if $\beta < \sqrt{5\alpha^2 + 2\alpha} - 2\alpha$

Second, if the threshold tax level is between τ_T and τ_N, $\tau_T < (\beta - \alpha)/2\beta < \tau_N$, there are two possible equilibria. Assume that when choosing their occupation, citizens expect $\tau = \tau_N$. Then, according to (25), an equilibrium must entail $l_N > 1/2$. Naturally, when taxes later are voted over and set according to (24), the initial expectation is confirmed. Hence, $l_N > 1/2$ and $\tau = \tau_N$ is one possible equilibrium. Now assume citizens expect $\tau = \tau_T$. Then, according to (25), an equilibrium must entail $l_N < 1/2$. When taxes later are voted over and set according to (24), the initial expectation is confirmed. Hence, $l_N < 1/2$ and $\tau = \tau_T$ is another possible equilibrium. From (22) and (23) and the constraints on α and β it follows that this is satisfied if and only if $\beta > \sqrt{5\alpha^2 + 2\alpha} - 2\alpha$.

The two cases above preclude other cases when we recall that α and β are positive and sum to less than one. Hence, the alternative where $\tau = \tau_T$ is the unique equilibrium is not possible.

Moreover, note that in any SPE citizens have the same utility across sectors (since citizens in one sector all have the same utility, and since in any SPE the occupational decision implies that the no-arbitrage condition $U_N = U_T$ is fulfilled). Therefore, when comparing two situations, the one that is more socially efficient Pareto dominates the other situation.

The following proposition summarizes these insights:

Proposition 2. *The possible SPEs are as follows:*

1. *If $\beta < \sqrt{5\alpha^2 + 2\alpha} - 2\alpha$ then $\tau = \tau_N$ with $l_N > 1/2$ is the unique SPE.*
2. *If $\beta > \sqrt{5\alpha^2 + 2\alpha} - 2\alpha$, there are two SPEs:*
   - *(a) $\tau = \tau_N$ with $l_N > 1/2$,*
   - *(b) $\tau = \tau_T$ with $l_N \leq 1/2$,*

with τ_N determined by (22) and τ_T by (23).

The size of the service sector will never be socially optimal. If $\tau = \tau_N$, the service sector is larger than is socially optimal. If $\tau = \tau_T$, the service sector is smaller than is socially optimal. All SPEs are Pareto dominated by the socially efficient situation $\tau = \gamma$.

Proposition 2 contains the impossibility result: free entry and perfectly competitive markets are not consistent with efficiency. The reason is simple: entry affects not only equilibrium factor prices, entry also affects political power. The majority group will tilt policy in its own favor. But since agents realize this at the point of entry, a no-arbitrage equilibrium must, necessarily, involve too much entry into the political majority and too little entry into the political minority. Scarcity on the one hand invites entry since it is economically attractive to supply the scarce factor, but on the other hand scarcity deters entry because it is politically unattractive to be part of the minority. From the point of view of society, however, only the former incentive enhances efficiency, the second does not. The impossibility result holds under more general, and weaker, conditions than in this particular model, as we show in Section 4.

Fig. 1 further clarifies the intuition underlying Proposition 2. Here we have depicted case 2 in the proposition, where there exist two possible SPEs. One possible equilibrium is point A where $\tau = \tau_T$ and $l_N \leq 1/2$, while the other possible equilibrium is point B where $\tau = \tau_N$ and $l_N > 1/2$.[15]

The figure displays both worker types' indifference curves in the (l_N, τ) plane, as dictated by Eqs. (18) and (19). The two curves to the right in the figure represent two indifference curves for goods workers. Their preference direction is rightward. Intuitively, for a given tax rate the utility of a goods worker increases in l_N, as that makes his labor scarcer; fewer goods workers implies a shortage of goods, which increases the wage in the goods relative to the services sector (depreciates the relative price p), and hence increases the purchasing power of goods workers. The D-shaped indifference curves to the left represent two indifference curves for a services sector worker. The preference direction for them is leftward; lowering the number of services workers for a given tax rate creates a shortage of services that increases their wage relative to the goods sector wage (appreciates the relative price p).

For both goods and services workers, utility is first increasing, then decreasing, in the tax rate. One mechanism is identical for the two

[14] In our model, corner solutions can never be part of an SPE because our assumed utility function has the property $\lim_{c_j \to 0} \frac{dU_i}{dc_j} = \infty$.

[15] For the illustration we have used parameters $\alpha = 1/4$, $\beta = 1/2$ and $\gamma = 1/4$. This leads to $l_N = 0.419 < 1/2$ when $\tau = \tau_T = 0.129$ and $l_N = 0.738 > 1/2$ when $\tau = \tau_N = 0.608$.

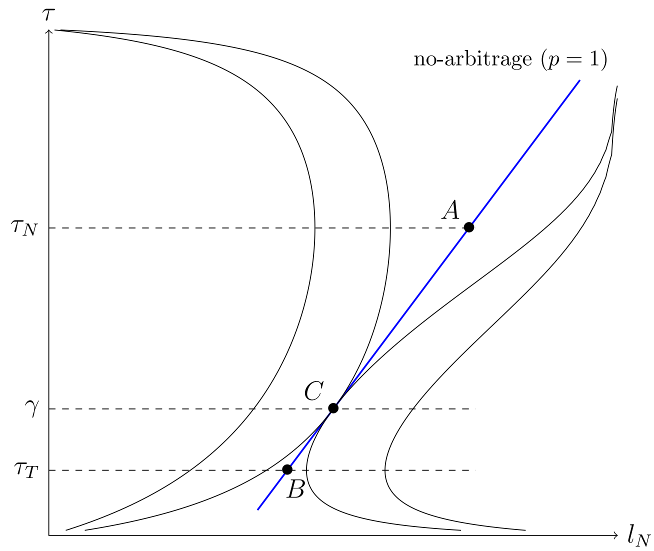


Fig. 1. Preferred tax rate for N and T-workers and candidates for equilibrium
Notes: The two curves to the right depict indifference curves for goods workers. The two curves to the left depict indifference curves for services workers. The blue upward-sloping line in the middle represents combinations of the workforce composition (l_N) and the tax rate (τ) such that relative price (p), is unity.

groups. Raising the tax from zero first supports the provision of essential public services, hence utility increases. As τ (and hence g) increases, the marginal gain from additional public services declines. Eventually, the gain is less than the opportunity cost and utility declines. For services workers, however, as we have seen, higher taxation comes with an additional positive effect. Keeping l_N fixed, a higher tax rate creates a shortage of services workers, which increases the price of services goods and increases the real wage of services workers. For goods-sector workers this same effect leads to a reduction in their real wage. Consequently, as we established in Lemma 1, for a given l_N, the goods workers have a preferred tax rate below the preferred tax rate for services workers. Moreover, as we also established in Lemma 1, the socially optimal size of the public sector, γ, exceeds that financed by the tax rate τ_T, and falls short of that financed by τ_N. The reason is that when l_N is predetermined, taxation redistributes purchasing power from workers in the goods sector to workers in the services sector. Starting out with a tax rate equal to γ, therefore, goods workers would like to see the tax rate reduced while services workers would like to see it increased.

As explained above, free entry implies that the no-arbitrage condition $p = 1$ holds so that no workers regret their choice of sector. The straight line in the figure gives the combinations of l_N and τ that are consistent with $p = 1$ from Eq. (25). Hence, we see that the two points A and B are possible SPEs because they each maximize the utility of the majority and satisfy $p = 1$. The efficient point C however, is not an SPE. No majority would ever vote for $\tau = \gamma$. In the current example, B is the better of the two possible equilibria. Hence, if the economy ends up in A rather than in B, this is to be considered a coordination failure.

Whether there are one or two SPEs depends on the location of the $p = 1$ line, and in particular for what two values of l_N it intersects τ_N and τ_T. If services are in high demand, i.e. if α is high, $l_N > 1/2$ even when $\tau = \tau_T$ and the only equilibrium will then be A, where the median voter is in the service sector. This is case 1 of Proposition 2. Conversely, if goods are in high demand, i.e. β is sufficiently high, then there will be an additional equilibrium B where the goods sector is too big and government spending is too low. This is case 2 of Proposition 2. Thus, small changes in preferences may translate into large differences in government size. Moreover, even with identical preferences, the size of government across countries may be very different.

A promise to vote for $\tau = \gamma$ would not be credible for anyone, and thus cannot support the socially efficient equilibrium C, even though C Pareto dominates both A and B.

3.3. Occupational mobility

An important question is if our impossibility result is subject to the Wittman-critique that "behind every model of government failure is an assumption of extreme voter stupidity, serious lack of competition, or excessively high negotiation/transfer costs," (Wittman, 1989, p. 1422). Clearly, there is no voter stupidity nor lack of competition behind our result. However, a remaining question is if the result rests on high costs of transferring from one occupation to another. Indeed, in the model as we have presented it so far, the occupational choice is once and for all, with no possibility to switch occupation at later stages of the game. We now relax this assumption and allow citizens to switch occupation at a strictly positive, possibly infinitely small, transfer cost ϵ.

Occupational mobility allows citizens to switch occupation after observing policy. We will first establish that no other SPEs than those in Proposition 2 can exist. Thereafter we turn to the question of when the SPEs in Proposition 2 remain.

To establish that no additional SPEs exist, we start out at the last stage of the game. Without loss of generality, assume the N sector is in majority. A citizen will not switch occupation if the utility gain from doing so falls short of the transfer cost ϵ. Therefore, for any factor allocation l_N there is an interval of tax rates such that there is no occupational switching. Within this interval there is one and only one tax rate, given by Eq. (25), that fulfills the arbitrage condition exactly. Term this tax rate τ_U. Thus, at any factor allocation l_N the median voter's choice must be τ_U in order for this factor allocation to be part of an SPE. Consider any factor allocation l_N that differs from the allocations in Proposition 2. Then τ_U also differs from the tax rates in Proposition 2. But majority citizens can always do at least marginally better than voting for τ_U (again, given that $\tau_U \neq \tau_N$). To see this, note that when $\tau_U \neq \tau_N$, majority citizens can vote for a tax rate that gives them marginally higher utility without making minority agents shift into the majority group (i.e. by keeping the utility differential between sectors less than ϵ). Such a tax rate gives the majority citizens higher utility than voting for τ_U. Thus, SPEs other than those in Proposition 2 do not exist.[16]

It follows that the only candidates for SPEs are those in Proposition 2. We now turn to the question of when they actually remain equilibria under occupational mobility.

The first order conditions for SPEs in Proposition 2 were derived under the restriction of no occupational mobility. Recall that we focus on SPEs in undominated strategies. The question is whether there exists a deviation in tax choice, which by causing occupational mobility makes the majority better off, and thereby also rules out the SPEs from Proposition 2.[17]

Assume that N is the majority and denote the factor allocation and tax rate in the SPE from Proposition 2 by $\left(l_N^*, \tau_N\right)$. Majority agents then prefer a tax rate and a subsequent sector movement of workers if there exists a τ such that

$$U_N(1 - l_N, l_N, \tau) = U_T(1 - l_N, l_N, \tau) - \epsilon, \tag{26}$$

$$U_N(1 - l_N, l_N, \tau) > U_N(1 - l_N^*, l_N^*, \tau_N). \tag{27}$$

[16] Note that although this argument rules out other possible SPEs than those in Proposition 2, we do not claim that majority citizens will never vote in a way that implies occupational shifting. This becomes clear in the next paragraphs. Nevertheless, the argument in this paragraph is sufficient to rule out other SPEs than those in Proposition 2.

[17] Recall that marginal deviations, i.e. deviations that do not cause agents to shift occupation, do not exist in the SPEs in Proposition 2.

The first condition states that some workers should be willing to move from sector N to sector T. The second condition states that the majority should be better off. Combining the two conditions implies

$$U_T(1-l_N,l_N,\tau) > U_N(1-l_N,l_N,\tau) > U_N(1-l_N^*,l_N^*,\tau_N).$$

Hence, a new candidate for (l_N,τ) must Pareto dominate the candidate(s) in Proposition 2 (and in order to generate moving, the T sector workers in this new candidate must obtain ϵ higher utility than the N sector workers). Because the candidates for SPEs from Proposition 2 are inefficient, a tax rate closer to the first-best might generate efficiency enhancement for both sector's workers also when moving costs are taken into account. Hence, when ϵ is sufficiently small, an N majority will never choose the tax rate τ_N (and conversely a T majority will never choose τ_T). Then there are no SPEs in the model.

We can now summarize:

Proposition 3. *Suppose citizens (at any stage) can switch occupation at a cost $\epsilon > 0$. Then, the only candidates for equilibria are those in Proposition 2. If ϵ is sufficiently small, there are no equilibria.*

It follows that the trinity is impossible even if the cost of switching occupation is infinitely small, while the inefficient equilibria in Proposition 2 exist only if the moving cost is sufficiently high.

3.4. Entry with a predetermined tax rate

We now study a different timing in the model, where we assume that the tax rate is voted over before entry (and where it is not possible to vote for a different tax rate after entry). As we will see, is such a case competitive markets and free entry is compatible with efficiency. Indeed, in our setting, the efficient equilibrium is the unique equilibrium. As such, this case also clarifies how our impossibility result rests crucially on the assumption about commitment. If voters could, ahead of their occupational choice, commit to stick to a particular future tax rate, the trinity of competitive markets, free entry and allocative efficiency is possible.

To study this case we may simply switch the order of stage 1 and 2 in the timeline described in Section 3.1, so that stage 2 in that timing arises ahead of stage 1. We have the same definition of equilibrium as in the main model, and find candidates for SPEs with backward induction given this new timing of events.

Stages 3 and 4 in the game are as before. Consider the new stage 2 of occupational choice. Given the predetermined tax rate, any equilibrium in occupational choice requires that the arbitrage condition is fulfilled. Thus, the equilibrium entry decision, for a given tax rate, is determined according to the line $p = 1$ in Fig. 1.

Moving back to the new first stage in the game, where there is voting over the tax rate, our restriction to SPEs in undominated strategies implies that given $p = 1$, any citizen votes for the first-best policy $\tau = \gamma$. Thus the unique equilibrium involves the optimal tax rate and allocative efficiency.[18]

3.5. Outcomes under a Ramsey planner

We have so far compared the political equilibrium to the first-best allocation. As stated in Section 2.2, in our model the first-best allocation would be implemented when there is a Ramsey planner. We can now infer this statement as a consequence of the analysis above.

A Ramsey planner sets the tax rate subject to the constraint that the other choices the citizens make are respected. In our case these choices regard the determination of production and prices, and in particular the decision about which sector to enter into. We assume the planner's objective is to maximizes the sum of all citizens' utility, with equal weight on each individual.

We first consider the case where the Ramsey planner can commit to a tax rate ahead of the entry decision of citizens. We are then back in a similar case to the one just discussed with the tax rate voted over before entry, except that now the tax rate is decided by the planner rather than by voters. The Ramsey planner has in this case no incentive to deviate from the optimal tax rate. Thus this tax is implemented, and then free entry of citizens will ensure that the first-best allocation follows.

We next consider the case where the Ramsey planner cannot commit to a tax rate ahead of the entry decision of citizens. We then have the following:

Proposition 4. *Without commitment the equilibrium outcome with a Ramsey planner will be the first-best outcome where citizens enter sectors so that $l_n = (1-\beta)$ and the planner sets the tax rate $\tau = \gamma$.*

Proof. In a no commitment equilibrium with free entry and a Ramsey planner, any candidate for equilibrium must have a wage that is identical in both sectors. Hence, if an equilibrium exists, we can restrict attention to cases where $p = 1$ after labor allocation has happened and after the Ramsey planner has set the tax rate. When setting the tax rate, the Ramsey planner maximizes the sum of utility, W, across all agents

$$W = l_N U_N + l_T U_T. \tag{28}$$

Differentiating W with respect to τ, inserting from (15) in (16) and (17) and evaluating where $p = 1$, using (15) again yields

$$\frac{\partial W}{\partial \tau}\frac{\tau}{W} = \frac{(\gamma-\tau)\alpha}{(1-\tau)(\beta\tau+\alpha)}. \tag{29}$$

It follows that the Ramsey planner will want to deviate from a tax rate in accordance with $p = 1$ for all τ except for the case where $\tau = \gamma$. Moreover, it is only when citizens allocate according to $l_N = 1-\beta$ that $\tau = \gamma$ is in accordance with $p = 1$. The allocation $l_N = 1-\beta$ in combination with tax $\tau = \gamma$ coincides with the social optimum. $\tau = \gamma$ is therefore optimal for the planner and it is thus the unique equilibrium. ■

Hence, a Ramsey planner would implement the first-best allocation, both with and without commitment.

4. Generalization

In this section, we show that the impossible trinity result holds under quite general conditions.

Consider a continuum of identical agents $i \in [0,1]$ who choose between two different activities N and T.[19] Let $\mathbb{I}_j$ denote the set of agents i who are in activity j. Then, the factor allocations $\{l_N, l_T\}$ are $l_j = \int_{i\in\mathbb{I}_j} 1\mathrm{d}i$, $j = N,T$. After choosing sector, a majority policy choice is made. Let the real number P denote the policy. The payoff, measured in transferable consumption units, U_i to an agent i depends on P and the agent's sector of occupation j, and is continuous in factor allocations:

$$U = U(P, l_N, l_T, j), \quad j \in \{N,T\}. \tag{30}$$

where U is continuously differentiable in the first three arguments. Let $P = P^*$ denote the efficient policy where (i) competitive markets and (ii) free entry yield the socially efficient allocation $\{l_N^*, l_T^*\}$ where the first order condition for efficiency yields

$$l_N^* U_P'(P^*, l_N^*, l_T^*, N) + l_T^* U_P'(P^*, l_N^*, l_T^*, T) = 0. \tag{31}$$

[18] Without our restriction to SPEs in undominated strategies, any tax rate along $p = 1$ in Fig. 1 is an SPE.

[19] In the Appendix, we further consider a model with a continuum of activities.

In order not to have a degenerate problem we restrict attention to settings where neither sector is superfluous:[20]

Assumption 1. $l_N^* > 0$ and $l_T^* > 0$.

We also assume that agents in each of the two activities have conflicting interests over policies.

Assumption 2. $U_P'(P, l_N, l_T, N) \neq U_P'(P, l_N, l_T, T)$.

As before, we restrict attention to SPEs in undominated strategies. A general proposition follows:[21]

Proposition 5. *Consider a majority policy choice where markets are competitive and there is free entry, and Assumptions 1 and 2 hold. Then, the socially efficient allocation $\{l_N^*, l_T^*\}$ is not an SPE.*

Proof. Consider the allocation $l_N = l_N^*$ and $l_T = l_T^*$. If $P = P^*$, then by definition of this policy, free entry implies that no agent would regret their choice of activity. However, from Assumption 2, agents in the majority activity j strictly prefer a different policy. Combined with the condition in (31) it follows that either of the two is true

$$U_P'(P^*, l_N^*, l_T^*, N) > 0 \wedge U_P'(P^*, l_N^*, l_T^*, T) < 0, \tag{32}$$

$$U_P'(P^*, l_N^*, l_T^*, N) < 0 \wedge U_P'(P^*, l_N^*, l_T^*, T) > 0. \tag{33}$$

Hence, under our restriction to undominated strategies, the policy choice P^* cannot be part of an SPE. If the majority implements their preferred policy, any agent in the minority activity will regret their choice of activity. Thus, $\{l_N^*, l_T^*\}$ is not an SPE. ■

Agents anticipate that if the allocation is $\{l_N^*, l_T^*\}$, no agent will vote for the efficient policy P^*. As there is free entry and perfect competition, the efficient allocation $\{l_N^*, l_T^*\}$ can never be an SPE. Proposition 5 shows that Assumptions 1 and 2 are sufficient[22] for the impossibility result and which therefore extends well beyond our baseline model.[23]

Heterogeneous productivities. The impossibility result is also robust to letting agents differ by individual productivity. Let the agents have productivity $a_{ji} > 0$ in activity j, with income in each activity proportional to productivity. Assuming that utility U is measured in consumption equivalents, payoffs to agent i in each activity j now also depend on productivity:[24]

$$U_i = a_{ji} U(P, l_N, l_T, j), \quad j \in \{N, T\}, \tag{34}$$

where the factor allocations are given by: $l_j = \int_{i \in \mathbb{I}_j} a_{ji} \mathrm{d}i$, $j = N, T$. Here again $\mathbb{I}_j$ denotes the subset of agents i who are in activity j. Without loss of generality, let i be ordered such that $\frac{\partial \frac{a_{Ni}}{a_{Ti}}}{\partial i} \geq 0$, implying that agent $i = 0$ has the highest comparative advantage in activity T, while agent $i = 1$ has the highest comparative advantage in activity N. In order not to introduce discontinuities, we also assume that the set of $\frac{a_{Ni}}{a_{Ti}}$ is connected.

Under the assumed ordering of i, we can denote the efficient allocation of workers by i^* and define efficient factor allocations $\{l_N^*, l_T^*\}$, where

$$l_T^* = \int_{i<i^*} a_{Ti} \mathrm{d}i, \quad l_N^* = \int_{i \geq i^*} a_{Ni} \mathrm{d}i.$$

With heterogeneous agents we can now be sure that if $P = P^*$, agents on *both* sides of i^* prefer to be in the majority activity, which implies that some minority activity agent regrets his or her choice. Hence the impossibility result still applies.

5. Conclusion

In this paper we have proved that under fairly general conditions, the combination of (i) competitive markets and (ii) free entry is inconsistent with (iii) allocative efficiency. Key to this impossibility result is that, in general equilibrium, allocations affect not only prices, but also policies. The distinguishing feature of our analysis is that agents are free to choose in which economic activity to enter, which in turn renders political preferences endogenous to entry. Agents must then take into account that entry determines payoffs not only through the standard economic returns to scarcity, but also through the political power of different groups, and thus, equilibrium policy. The requirement that arbitrage conditions must be fulfilled guarantees that equilibria cannot be socially optimal, conditional on the requirement that policy responds to political power. In our specific model, it follows that the optimal size of government never constitutes an equilibrium.

Our particular application offers a new perspective on the well-known observation that seemingly similar countries differ greatly by their size of government. In our model, the equilibrium size of the public sector will be either smaller or larger than the social optimum. Moreover, small differences in preferences, even identical ones, can result in very different government sizes.

Declaration of competing interest

We, the authors, declare that we have no undeclared funding sources, no conflicts of interest, nor any other ethical issues related to our submission.

Data availability

No data was used for the research described in the article.

Appendix. Model with a continuum of sectors

In Section 2 we developed a two-sector model where the government purchases goods or services from one of the sectors only. We now generalize the analysis by studying a model with a continuum of sectors as in Dornbusch et al. (1980) and where the government potentially purchases the output from all sectors.

The sectors are distributed on the unit interval $i \in [0, 1]$. The mass of workers is normalized to unity. Let $l(i)$ denote the mass of workers in sector i. Full employment then implies

$$\int_{i=0}^{1} l(i) \mathrm{d}i = 1.$$

Productivity in each sector is constant and equal to one. Output then equals employment and is used for private and government consumption:

$$l(i) = c(i) + G \cdot g(i), \tag{A.1}$$

[20] Note that $P = P^*$ could be the absence of policy, as would be the case if the well-known conditions for perfect competition hold. Alternatively, in the presence of market failures $P = P^*$ would be the optimal policy that corrects for these. In our baseline model, for instance, $P = P^*$ is the policy of $\tau = \gamma$.

[21] Again, should the mass of agents in each activity be identical, voters in one pre-specified activity j will be decisive.

[22] Trivially, if Assumption 1 is violated for example by $l_N^* = 0$, then the socially efficient allocation is a corner solution with all agents in the T-activity, and the efficient policy $P = P^*$ would be an SPE. Also trivially, if $P = P^*$ is the majority groups preferred policy, in violation of Assumption 2, the efficient allocation would be an SPE.

[23] The impossibility result applies also in circumstances where there is free entry and where the minority is most influential, as argued by Olson (1965) (and further analyzed by Esteban and Ray, 2001). The only modification is that now the policy will be shifted in the direction of the minority sector. Hence, also in the final possible scenario where the decisive group is drawn stochastically, as in e.g. a probabilistic voting model, the impossibility result holds.

[24] All results go through if the true utility is a monotone transformation of $a_{ji} U(P, l_N, l_T, j)$.

where $c(i)$ is private purchases and $G \cdot g(i)$ is government purchases from sector i.

To allow for endogenous heterogeneity in preferences, we let the government purchase a share $g(i)$ of its total consumption G from each sector i, hence $\int_{i=0}^{1} g(i)\mathrm{d}i = 1$. The sectors are indexed such that $g'(i) > 0$. For later purposes, note that the average g equals unity (because the sectors are distributed on the unit interval). The resulting budget constraint for the government is

$$\tau Y = G \int_{i=0}^{1} p(i)\, g(i)\, \mathrm{d}i, \tag{A.2}$$

where $p(i)$ is the price of goods from sector i, τ is the tax rate and Y is aggregate private income:

$$Y = \int_{i=0}^{1} l(i)\, p(i)\, \mathrm{d}i. \tag{A.3}$$

As the mass of workers is unity, Y is also average income.

A citizen can only be employed in one sector. The income of a citizen working in sector j, later referred to as worker j, is $p(j)$. Let $c_j(i)$ denote worker j's consumption of goods from sector i. The utility of worker j is

$$U_j = (1-\gamma)\int_{i=0}^{1} \ln(c_j(i))\, \mathrm{d}i + \gamma \ln G, \tag{A.4}$$

which implies that the budget share spent on each private good is independent of j's income. Since citizens only differ by the sector in which they work (and thus potentially differ by income), it follows that all citizens have equal expenditure shares. Hence,

$$c_j(i) = \frac{p(j)}{Y} c(i), \tag{A.5}$$

where $\frac{p(j)}{Y}$ is the relative income of worker j. It also follows that total spending on each sector's output is

$$c(i)\, p(i) = Y(1-\tau). \tag{A.6}$$

We now turn to characterizing each citizen's preferred size of government, G, for any given occupational choices. For any given allocation of workers, the utility of a worker in sector j may be written as

$$\begin{aligned} U_j &= (1-\gamma)\left(\int_{i=0}^{1} \ln\left(c(i)\frac{p(j)}{Y}\right)\mathrm{d}i\right) + \gamma \ln G \\ &= (1-\gamma)\int_{i=0}^{1} \ln(c(i))\,\mathrm{d}i + \gamma \ln G - (1-\gamma)\ln\frac{Y}{p(j)}, \end{aligned}$$

where we in the first line have applied Eq. (A.5). Eq. (A.3) and (A.6) imply that $Y = \int_{i=0}^{1} l(i)\, Y(1-\tau)/c(i)\,\mathrm{d}i$, while (A.6) implies $p(j) = \frac{Y(1-\tau)}{c(j)}$. Hence, $\frac{Y}{p(j)} = c(j)\int_{i=0}^{1} \frac{l(i)}{c(i)}\mathrm{d}i$. Utility is therefore given by

$$U_j = (1-\gamma)\int_{i=0}^{1} \ln(c(i))\,\mathrm{d}i + \gamma \ln G - (1-\gamma)\ln\left(c(j)\int_{i=0}^{1}\frac{l(i)}{c(i)}\mathrm{d}i\right).$$

When the allocation of workers is fixed, it follows from (A.1) that the effect of government size on each $c(i)$ is $\mathrm{d}c(i) = -g(i)\,\mathrm{d}G$. The effect of G on worker j's utility is therefore given by

$$\begin{aligned} \mathrm{d}U_j =& \left(-(1-\gamma)\int_{i=0}^{1}\frac{g(i)}{c(i)}\mathrm{d}i + \frac{\gamma}{G}\right)\mathrm{d}G \\ &+ (1-\gamma)\left[\frac{g(j)}{c(j)} - \frac{\int_{i=0}^{1}\frac{l(i)g(i)}{c(i)^2}\mathrm{d}i}{\int_{i=0}^{1}\frac{l(i)}{c(i)}\mathrm{d}i}\right]\mathrm{d}G. \end{aligned} \tag{A.7}$$

We now impose that in the first stage of the game, where citizens choose their occupation, there is free entry. As in the main text, an equilibrium with free entry must have the property that $p(i) = 1$ for all i. It follows from (A.3) that $Y = 1$, and thus from (A.2) that $\tau = G$. Eq. (A.6) then implies that $c(i) = 1 - G$ for all i, which we can insert into Eq. (A.7) to obtain

$$\mathrm{d}U_j = \frac{1}{1-G}\left[\frac{\gamma}{G} - 1 + (1-\gamma)\left(g(j) - \int_{i=0}^{1} l(i)\, g(i)\,\mathrm{d}i\right)\right]\mathrm{d}G.$$

Next, we note that $c(i) = 1 - G$ and (A.1) imply

$$l(i) = 1 - G + g(i)\, G. \tag{A.8}$$

It follows that

$$\mathrm{d}U_j = \frac{1}{1-G}\left[\frac{\gamma}{G} - 1 + (1-\gamma)\left(g(j) - 1 + G - G\int_{i=0}^{1} g(i)^2\,\mathrm{d}i\right)\right]\mathrm{d}G. \tag{A.9}$$

The square bracket in (A.9) is continuously decreasing in G. Therefore, for each worker j, there is only one G that satisfies the first-order condition $dU_j/dG = 0$. This G constitutes a global optimum for worker j because any intermediate value of G is superior to both border values 0 and 1. Hence, if we let j^* denote the sector of a citizen who considers a given $G \in (0,1)$ as optimal, the following must hold:

$$0 = \frac{\gamma}{G} - 1 + (1-\gamma)\left(g(j^*) - 1 + G - G\int_{i=0}^{1} g(i)^2\,\mathrm{d}i\right). \tag{A.10}$$

Note that (A.9) implies $\mathrm{d}U_j$ is monotonically increasing in $g(j)$. We can therefore define a function $h(G)$ that traces out how j^* in (A.10) depends on G:

$$j^* = h(G) \equiv g^{-1}\left(\left(1 - \frac{\gamma}{G}\right)\frac{1}{1-\gamma} + 1 + G\mathrm{V}_g\right), \; h'(G) > 0. \tag{A.11}$$

Here V_g is the variance of g

$$\mathrm{V}_g = \int_{i=0}^{1} g(i)^2\,\mathrm{d}i - 1.$$

To be clear, $h'(G) > 0$ means that worker j's preferred G is increasing in j. Thus, the preferences among workers are heterogeneous, where the higher is a worker's sector's share $g(j)$ in government consumption, the higher G does the worker prefer.

From Eq. (A.10), we note that no worker wants zero government consumption:

$$h(G) = 0 \iff G = \underline{G} > 0. \tag{A.12}$$

A political equilibrium requires that the median voter does not want to alter the size of government. Thus $m = j^*$, where m denotes the sector of the median voter.

We next turn to characterizing the sector of the median voter. By definition, this sector follows from $\int_{i=0}^{m} l(i)\mathrm{d}i = 1/2$. Eq. (A.8), which followed from free entry and market clearing, implies that whenever $G > 0$, $l'(i) > 0$. Hence, because there is a unit mass of workers, it follows that $m > 1/2$ for positive G. Moreover, the higher is G, the higher is $l'(i) > 0$, and thus the higher is the index m of the median voter. The intuition is simply that when G is high, sectors that produce a high share of government consumption demand more labor. With free entry, more citizens will then choose to enter those sectors.

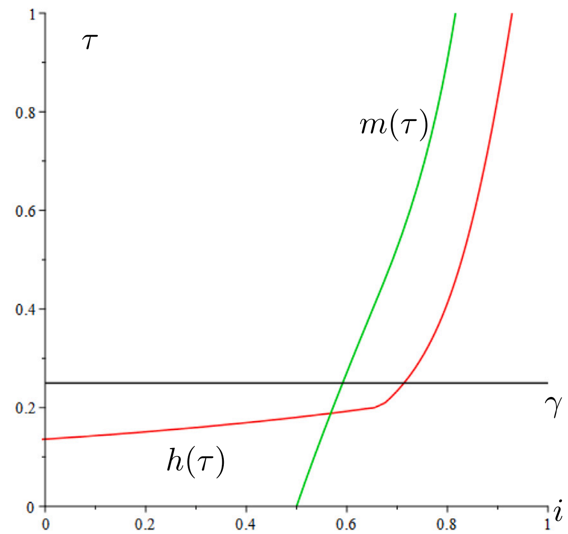
We can now define how m depends on G. By combining $\int_{i=0}^{m} l(i)\mathrm{d}i = 1/2$ with (A.8), we obtain the relationship

$$m(G) = \mathrm{argsolve}_m\left[\int_{i=0}^{m}(1-G) + g(i)\,G\mathrm{d}i = 1/2\right],$$

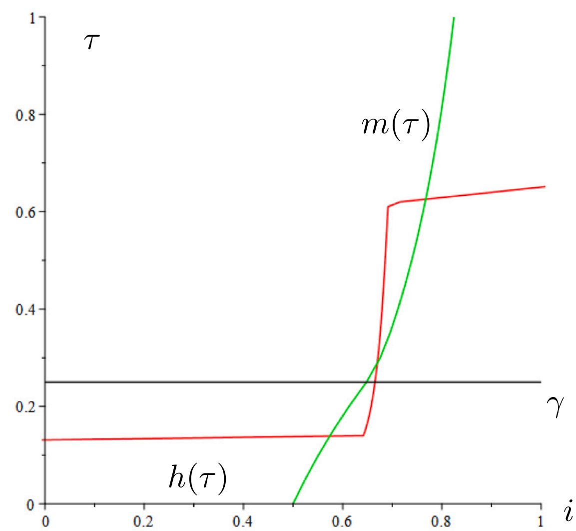
which is equivalent to

$$m(G) = \mathrm{argsolve}_m\left[m(1-G) + G\int_{i=0}^{m} g(i)\mathrm{d}i = 1/2\right]. \tag{A.13}$$

We note immediately that m is less than one even if $G = 1$. Hence, $m \in [1/2, \bar{m}]$, where $\bar{m}$ is the median sector i in the distribution $g(i)$. Because i is continuous, $\bar{m} < 1$.



(a) Unique equilibrium, subotimal G, $\tau < \gamma$.



(b) Multiple equilibria.

Fig. 2. Political economy equilibria with continuum of sectors.

Implicit derivation of (A.13) implies that

$$m'(G) = \frac{m - 1/2}{Gg(m) - G + 1} \frac{1}{G} > 0 \text{ since } m > 1/2, G > 0.$$

The two functions h from Eq. (A.11) and m from Eq. (A.13) define the two conditions that have to be satisfied in equilibrium. An internal solution G has to be at a level so the median voter corresponds to an agent that do not want to alter G. As in the main text, in equilibrium prices are $p(i) = 1$ for all i and hence $\tau = G$. Formally, we then have

$$h(\tau) = m(\tau).$$

As both $h(\tau)$ and $m(\tau)$ are upward sloping, there will be cumulative forces at play as in the simpler model in the main text. Expectations about a high τ will attract many workers to the high g sectors and the median voter is one who prefers a high τ. Conversely, expectations about a low G will attract many workers to low g sectors and the median voter is one who prefers a low τ. Equilibrium requires that the expected τ is preferred by the median voter. Depending on the exact distribution of g, these cumulative forces may generate multiple equilibria.

We may now conclude. Our impossibility result in the main text, namely that the combination of competitive markets, free entry and efficiency is unfeasible also holds when we extend our model with a continuum of sectors. When τ is determined by voter preferences $h(\tau)$, with competitive markets it is impossible to jointly satisfy the free entry condition, i.e. the identity of the median voter is determined by the allocation of workers $m(\tau)$, and efficiency, i.e. $\tau = \gamma$.

Proposition A.1. *The combination of competitive markets and free entry will generally not deliver an efficient outcome.*

Proof. The political condition $h(\tau) = m(\tau)$ and the condition for efficiency $\tau = \gamma$ yields an overdetermined system. It would only be by coincidence if $h(\gamma) = m(\gamma)$. ■

In Fig. 2 we have illustrated the impossibility result with two examples. In panel (a) there is a unique equilibrium with $\tau = G < \gamma$. In panel (b) there are multiple equilibria, similar to the one in our main text, but now with a continuum of sectors.

References

Acemoglu, Daron, 2003. Why not a political coase theorem? Social conflict, commitment, and politics. J. Comp. Econ. 31, 620–652.

Acemoglu, Daron, Egorov, Georgy, Sonin, Konstantin, 2018. Social mobility and stability of democracy: Reevaluating De Tocqueville. Q. J. Econ. 113, 1041–1104.

Acemoglu, Daron, Johnson, Simon, Robinson, James A., 2005. The rise of Europe: Atlantic trade, institutional change and economic growth. Amer. Econ. Rev. 95 (3), 546–579.

Barro, Robert J., 1990. Government spending in a simple model of endogenous growth. J. Polit. Econ. 98, 103–125.

Benabou, Roland, Ok, Efe, 2001. Social mobility and the demand for redistribution: The POUM hypothesis. Q. J. Econ. 116, 447–487.

Benabou, Roland, Tirole, Jean, 2006. Belief in a just world and redistributive politics. Q. J. Econ. 121, 699–746.

Besley, Timothy, 2006. Principled Agents? the Political Economy of Good Government. Oxford University Press, Oxford.

Besley, Timothy, Coate, Stephen, 1997. An economic model of representative democracy. Q. J. Econ. 112, 85–114.

Besley, Timothy, Coate, Stephen, 1998. Sources of inefficiency in a representative democracy: A dynamic analysis. Amer. Econ. Rev. 88, 139–156.

Buchanan, James M., Tullock, Gordon, 1977. The expanding public sector: Wagner squared. Public Choice 31 (1), 147–150.

Christoffersen, Henrik, Paldam, Martin, 2003. Markets and municipalities: A study of the danish municipalities. Public Choice 114, 79–102.

Dornbusch, Rudiger, Fischer, Stanley, Samuelson, Paul A., 1980. Heckscher-ohlin trade theory with a continuum of goods. Q. J. Econ. 95 (2), 203–224.

Drazen, Allan, 2000. Political Economy in Macroeconomics. Princeton University Press, Princeton.

Esteban, Joan, Ray, Debraj, 2001. Collective action and the group size paradox. Am. Polit. Sci. Rev. 95, 663–672.

Hassler, John, Mora, José V. Rodríguez, Storesletten, Kjetil, Zilibotti, Fabrizio, 2003. The survival of the welfare state. Amer. Econ. Rev. 93 (1), 87–112.

Kydland, Finn E., Prescott, Edward C., 1977. Rules rather than discretion: The inconsistency of optimal plans. J. Polit. Econ. 85, 473–492.

Lindbeck, Assar, 1995. Hazardous welfare-state dynamics. Amer. Econ. Rev. 85 (2), 9–15.

Lizzeri, Alessandro, Persico, Nicola, 2001. The provision of public services under alternative electoral incentives. Amer. Econ. Rev. 91 (1), 225–239.

Meltzer, Allan H., Richard, Scott F., 1981. A rational theory of the size of government. J. Polit. Econ. 89, 914–927.

Milesi-Ferretti, Gian Maria, Perotti, Roberto, Rostagno, Massimo, 2002. Electoral systems and public spending. Q. J. Econ. 117, 609–657.

Olson, Mancur, 1965. The Logic of Collective Action. Harvard University Press, Cambridge.

Osborne, Martin J., Slivinski, Al, 1996. A model of political competition with citizen-candidates. Q. J. Econ. 111, 65–96.

Persson, Torsten, Roland, Gerard, Tabellini, Guido, 1997. Separation of powers and political accountability. Q. J. Econ. 112, 1163–1202.

Persson, Torsten, Roland, Gerard, Tabellini, Guido, 2000a. Comparative politics and public finance. J. Polit. Econ. 108, 1121–1161.

Persson, Torsten, Tabellini, Gudio, 2000b. Political Economics. MIT Press, Cambridge.

Rattsø, Jørn, Sørensen, Rune J., 2016. Public–private political cleavage: what happens after retirement. Public Choice 166, 315–339.

Rosen, Harvey S., Rosen, Kenneth T., 1980. Federal taxes and homeowership: Evidence from time series. J. Polit. Econ. 88, 59–75.

Tiebout, Charles M., 1956. A pure theory of local expenditures. J. Polit. Econ. 64, 416–424.

Tullock, Gordon, 1974. Dynamic hypothesis on bureaucracy. Public Choice 19 (1), 127–131.

Wittman, Donald, 1989. Why democracies produce efficient results. J. Polit. Econ. 64, 416–424.