Divided We Stad: a Fiscal Bargaining Model for Divided Countries

Vanschoonbeek, Jakob

KU Leuven - University of Leuven, Research Centre for Regional Economics (VIVES)

15 July 2020

DIVIDED WE STAND:
A FISCAL BARGAINING MODEL FOR DIVIDED COUNTRIES

Jakob Vanschoonbeek*, †

VIVES Discussion Paper 88. This version:
July 15, 2020

Abstract

I propose a multi-region fiscal bargaining model, where country border stability is determined by a trade-off between preference heterogeneity, income inequality and scale economies in the provision of public policy. I demonstrate how increasing preference heterogeneity can actually *increase* border stability in this framework; by increasing political disagreements over government expenditures, it functions as a tax-reducing mechanism that brings equilibrium tax rates more in line with the fiscal preferences of the discontented and the wealthy. I provide empirical evidence for the model's main predictions in a large panel of countries over the past half-century. I also show its capacity to match some of the main fiscal and political evolutions in Belgium since its inception. The model provides a rationale for the observed border stability in highly politically heterogeneous and economically unequal countries, such as Belgium, which is difficult to explain in existing models.

Keywords: Secession and unification, political economy, fiscal policy, fiscal federalism, decentralization, transfer schemes, quantitative models, Belgium

JEL Classification: C70, D71, H41, H71, H77

*KU Leuven - University of Leuven, Research Centre for Regional Economics (VIVES)
†Ph.D. Fellow of the Research Foundation - Flanders (FWO)

The author would like to thank Christophe Crombez, Klaus Desmet, Joep Konings, Jo Reynaerts and Marvin Suesse for valuable comments and suggestions on earlier versions of this paper.

1 Introduction

Separatism is on the rise in many parts of the developed world, raising the importance of understanding its root causes. In sharp contrast to the first wave of separatism, which dates back to the decolonization movements that spawned after World War II and exclusively originated in the poorest parts of the world, today's separatist tensions have primarily sprung up in the richest regions. Ranging from Flanders in Belgium, Catalonia and Basque Country in Spain and Veneto, Lombardy and South Tyrol in northern Italy, independence movements all over the world have increasingly *"swapped out the 'identity card' for the 'economic card"'* (Rodríguez-pose, 2017). The task of explaining this development is often identified as one of the main challenges of economic geography and regional economics because, as long as orthodox economics provides no answers to the frictions caused by divergent regional development, *"populist insurgents will"* (The Economist, 2016).

Although political economists have recently made significant progress in analyzing the determinants of country border stability, the fact that some regions witnessed a surge in separatist vote shares while others with comparable degrees of preference heterogeneity and economic development did not is difficult to explain in the existing literature (Gehring & Schneider, 2020). This relates to an important shortcoming in the two workhorse theories, which jointly consider scale economies in the provision of public policy as the most important centripetal force that may keep countries together, but identify either preference heterogeneity (Alesina & Spolaore, 1997, 2003) or differing redistributive preferences (Bolton & Roland, 1997) as the single most potent centrifugal force. Up until now, almost no attention has been given to how both centrifugal forces interact to jointly affect country border stability. This is both due to Alesina and Spolaore's (1997, 2003) explicit assumption that government size is fixed and exogenously given, which effectively shuts down the role of fiscal policy, but also to the absence of preference heterogeneity in the model developed by Bolton and Roland (1997), which obscures its potential fiscal effects.[1]

In this paper, I address this shortcoming by extending a model which encompasses all three centripetal and centrifugal forces. In this multi-region model, regional populations majority vote both over the optimal location and amount of public spending in the union, taking into account increasing returns to scale in the provision of public policy, the distinctiveness of their policy preferences as well as their individual income level. Crucially, both richer voters and voters with more distinct policy preferences have a preference for lower income tax rates and smaller governments. Assuming that discontented regions may threaten to majority vote over secession if the welfare costs of the union become too large, this allows me to derive how secession proneness depends on the interaction of preference heterogeneity, income inequality and scale economies in the provision of the public good.

[1] To be fair, Alesina and Spolaore (1997, p. 1046) explicitly acknowledge that income differences *"may be crucial determinants of the degree of heterogeneity* [...] *that determines the equilibrium size and number of countries"*, while Bolton and Roland (1997, p. 1085) also recognize that, in their model, *"the effect of differences in language and culture on incentives to separate should also be explored further"*.

The central finding is that, in this model, preference heterogeneity may actually have either centripetal or centrifugal effects, depending on how it affects equilibrium tax rates. While the direct effect of preference heterogeneity is to reduce country border stability, by increasing the gap between government and the people and hence the welfare cost of preserving the union, its indirect effect is to reduce the willingness to pay taxes, driving down equilibrium tax rates by decreasing the utility derived from public consumption. Consequently, the tax-reducing effect of preference heterogeneity serves as an instrument that better aligns the union's equilibrium tax rates with the fiscal preferences of the discontented and the wealthy and may therefore reduce the fiscal gains of secession for the richest regions within the union, who are net-contributors to the tax system. As the outside option is generally more attractive for wealthy regions, due to their tax base advantage, preference heterogeneity may in fact increase country border stability and a certain level of political disagreement may actually be required to cement borders in highly economically unequal countries. The model therefore provides a rationale for the observed border stability in highly politically divided and economically unequal countries, such as Belgium, which is more difficult to explain in either of the existing canonical models.

One intriguing corollary is that low levels of preference heterogeneity may produce inefficient separatism in richer regions. That is, when political consensus on public policy in the union sufficiently drives up the median voter's willingness to pay taxes, their lack of majority control over fiscal policy may drive wealthy regions to secession, even if this reduces aggregate welfare. This naturally begs the question of which instruments the union could use to stabilize efficient country borders. Comparing three such instruments, I find that both a tax ceiling and a system of compensating transfers to the most discontented regions are capable of fully stabilizing efficient country borders, while fiscal decentralization can only stabilize efficient country borders to a certain extent. The reason for this slightly counter-intuitive finding is that discontented regions actually face a trade-off under fiscal decentralization, which on the one hand lowers the heterogeneity costs of the union by reducing the gap between government and the people, but for the same reason also increases the willingness to pay taxes, and may thereby magnify the discrepancy between regionally preferred and centrally decided government size, especially if preference heterogeneity in the union is low. Although fiscal decentralization nonetheless appears to be the most politically feasible instrument, granting every inhabitant a larger say in fiscal matters, it thus may also leave countries vulnerable to inefficient separatist threats.

I confront the predictions of the model with historical data in two ways. First, I employ large scale cross-country comparisons to show that national tax revenue shares are inversely related to electoral fractionalization in a large sample of countries over the past half-century, even when controlling for other variables that may plausibly affect fiscal policy. I also test for interaction effects, finding empirical evidence that fiscally centralized countries are more sensitive to the tax-reducing effect of preference heterogeneity, as predicted by the model. To indirectly test the model's central prediction, that political heterogeneity

could have centripetal or centrifugal effects, I regress a proxy for secession proneness, the number of independence movements, on a proxy for preference heterogeneity, an index of electoral fractionalization, in a model that either excludes or includes tax revenue shares as a control. I find empirical evidence that the centrifugal effects of preference heterogeneity increase once I include tax revenue shares to control for its - potentially stabilizing - fiscal effects, consistent with the idea that the centrifugal effects of preference heterogeneity may be at least partially counteracted by its potentially centripetal fiscal effects.

As a cross-country comparison can only go so far in providing evidence for the model's validity, I complement this multivariate regression analysis with a detailed numerical application to Belgium. Belgium is an interesting case study, both known for longstanding political disputes along linguistic lines and the scene of a dramatic reversal of fortune, where Flanders for a long time was Belgium's poorest region but economically surpassed Wallonia in the postwar period. Most interestingly, separatism shone in its absence throughout the 19^{th} and early 20^{th} centuries, despite transfers moving from then-rich Wallonia to then-poor Flanders, while Flemish separatism would briefly emerge in the interwar period to return in full force after World War II. To determine the extent to which the model can account for the timing and location of secession proneness in Belgium, I calibrate a two-region dynamic model based on two exogenous 'laws of motion' that are directly taken from the data: the democratization process, which gradually extended voting rights to poor, less-educated and female voters; and income growth, which gradually increased living standards. Not only does this allow for an assessment of model predicted equilibrium taxes, political heterogeneity and separatist vote shares in Belgium for the past 200 years against actual historical data, it also allows for a better understanding of how the model relates to a number of historical forces that are often thought to be important drivers of the historical evolution in the number and sizes of nations.

The calibrated model turns out to be well capable of predicting the timing, if not the location, of all the main fiscal and political developments in Belgium since its inception in 1830. According to the model, the absence of separatism in the 19^{th} century is mainly related to the electoral system of census voting, which ensured that median voters favored small government both under union and under separation. As a result, low levels of public consumption effectively minimized the importance of maintaining influence over the location of public policy and scale economies stacked the odds in favor of existing country borders. The emergence of separatism is then related to the unprecedent growth in democracy, with the Belgian electoral reforms of 1893 and 1918, which shifted electoral power to the poor and fueled the first separatist tensions between the then-richer south and the then-poorer north through the associated popular pressures to extend the welfare state. Even after universal suffrage, increasing economic security kept expanding the number of politically relevant policy issues, leading to ever larger levels of political disagreement. This reduced the willingness to pay income taxes in the union and gradually increased the attractiveness of the outside option in the fastest growing region of Flanders, whose tax

base advantage would allow it to declare independence at the lowest economic cost.

The emergence of secession proneness in the model is thus intimately related to the democratization process, but less because it increased preference heterogeneity as suggested by Alesina and Spolaore (1997, 2003), than because it opened the way for the welfare state and drastically increased the relative importance of public to private consumption. As alluded to by Gerard (2014), the emergence of the welfare state thus constitutes the biggest distinguishing feature between 19^{th}- and 20^{th}-century Belgium, where the absence of government explains the absence of separatism throughout the 19^{th} century, while its growth in a time of regional divergence explains the perpetual Flemish-Walloon fiscal disputes in the postwar period. In turn, this implies that preference heterogeneity may actually stabilize Belgian country borders, by keeping down tax rates to mutually agreeable levels.

To verify this, a comparative static exercise suggests that Belgian tax revenue shares would roughly lie 30% higher if policy preferences were fully congruent. More interestingly, the model implies that preference heterogeneity had different effects on the intensive and extensive margins of separatism: with respect to the former, its tax-reducing effect significantly lowered the net independence gains for the most secession-prone inhabitants in the top of the income distribution, who are primarily located in Flanders; with respect to the latter, its increase of the gap between government and the people increased the number of Flemish voters experiencing positive independence gains. Although preference heterogeneity thus turns out to have ambiguous net results, to the extent that the independence gain must exceed a certain threshold for the Flemish electorate to risk voting for separatist parties, it may well have had centripetal effects in Belgium. Finally, I also find suggestive evidence that federalism has had strong centripetal effects on the extensive margin and mainly so for the most secession-prone Flemish region, as the ability to spend a fixed share of own tax contributions subnationally reduced the common tax pool problem. By increasing the willingness to pay taxes, however, federalism also exerted strong centrifugal effects on the intensive margin for the wealthy, consistent with the recent rise in Flemish minority separatism even after Belgium became a fully-fledged federal state.

This research is closest in spirit to the small but growing empirical literature on the political economy of country borders. Lake and O'Mahony (2004) relate the long-term trend in average state size to the general trends in a number of potential determinants but fail to find any clear relationship capable of explaining the incredible rise and subsequent fall in average state size witnessed in modern history. Desmet, Le Breton, Ortuño-Ortín, and Weber (2011) calibrate a static model of border formation that connects genetic distances among populations to country border stability and is able to successfully predict the order of state breakup in former Yugoslavia. Suesse (2018) finds evidence that the trade-off between scale economies and population heterogeneity partially explains regional variation in separatist demonstrations in the former Soviet Union. Vanschoonbeek (2020) relies on a similar model to predict regional secession proneness for a large number of regions in Europe, finding that the model is well capable of explaining heterogeneity in

regional instability in a European context. None of these papers look in any detail how preference heterogeneity and fiscal policy interact in determining country border stability.

Of particular interest is the work of Gancia, Ponzetto, and Ventura (2017), who recently develop the first dynamic model of the sizes of nations to study a more primitive technological driver of border formation, relating shifting country borders to a gradual decline in trade costs. Departing from the premise that falling trade costs raise the gains from trade, they derive a non-monotonic effect of globalization on country border formation, as trade borders were initially removed by expanding country borders, at the increasing cost of political heterogeneity and colonial wars, but once the gains from trade became sufficiently large, trade borders were altogether eliminated by setting up a world trade union, at the fixed cost of diseconomies of scope, decreasing optimal country size. Nevertheless, their model abstracts from several other relevant forces that are emphasized in this analysis, namely the global democratization process and the simultaneous increase in government size. The numerical application to Belgium abstracts from globalization, yet supplements their analysis with another potential explanation, namely that the exogenous increase in democracy and income growth could also reproduce the initial rise and subsequent fall in average country size as it was observed over the past two centuries.[2]

On the theoretical front, this work also builds on a number of extensions to Alesina and Spolaore (1997). Staal (2010) confirms their results in a model where public spending and taxation exogenously depend on country size, while Etro (2006) is among the first to study country border stability when endogenizing the provision of public goods, subjecting the level of public spending to majority voting at the country level. Perhaps most closely related is the fiscal model of Desmet, Ortuño-ortín, and Weber (2017), where a country's overall diversity reduces tax compliance and governments either substitute public goods for transfers when a society's *overall* diversity increases, to compensate for the tax revenues lost due to declining tax compliance, or transfers for public goods when *peripheral* diversity increases, to reduce the heterogeneity costs stemming from political disagreements over the content of public policy. Nevertheless, none of these studies analyze the role of fiscal policy as a crucial mediator between political heterogeneity and country border stability. Also relevant is the literature on asymmetric federalism, which typically finds side payments to discontented regions to sometimes be necessary to avert inefficient state breakup in heterogeneous countries (Le Breton & Weber, 2003; Haimanko, Le Breton, & Weber, 2005; Claeys & Martire, 2015), as well as the fiscal federalism literature, which similarly argues that a direct transfer of fiscal power from the center to the regions may sometimes be sufficient to keep efficient but heterogeneous countries together (Spolaore, 2009, 2010, 2016). The results here nevertheless show that compensating trans-

[2]Interestingly, Adserà and Boix (2002) even suggest a potential link between both models in the form of a relation between trade openness and government size. They contend that, as globalization and trade creates winners and losers, this may enable the losers to extract rents from the winners in the form of an enlarged welfare state, in return for their subsequent democratic approval of increased trade openness.

fers and fiscal decentralization may also be necessary to stabilize efficient country borders if preference heterogeneity becomes *too low*.

Finally, this analysis is also related to a recent literature that connects ethnic diversity to a worsening of several political economy outcomes, including public goods provision. Alesina, Baqir, and Easterly (1999) started this literature by establishing that ethnically diverse jurisdictions in the United States also tend to spend less on public goods, a finding that was later replicated for India, Kenya and a cross-section of countries (Banerjee, Iyer, & Somanathan, 2005; Miguel & Gugerty, 2005; La Porta, Lopez-de Silanes, Shleifer, & Vishny, 1999). Desmet, Ortuño-ortín, and Wacziarg (2017) find cross-country evidence that ethnic diversity only hampers public goods provision when overlapping with cultural diversity, but not when there is little overlap between ethnicity and culture. Several theoretical channels have been proposed to rationalize these findings, including that people from different ethnic groups may have different preferences for public policy, complicating policy formulation; may be reluctant to interact with other ethnic groups; or may be hampered to act collectively to promote group interests, especially if their group is small and there are many competing groups (Alesina, Baqir, & Hoxby, 2004; Banerjee, Iyer, & Somanathan, 2007). This literature has also spawned a methodological discussion on how to best measure the social antagonisms induced by ethnic diversity, which essentially revolves around the question whether it is the number of groups that matters or their relative size (Montalvo & Reynal-Querol, 2005). While in the former case, a fractionalization index is the most appropriate quantitative measure, in the latter case, a polarization measure should be employed instead. The model presented here moves beyond the ethnic dimension and analyzes how public goods provision depends on preference heterogeneity more generally. Moreover, it links the fiscal effects of preference heterogeneity to the stability of country borders. It also sidetracks the methodological debate by not conditioning the analysis on a particular measurement of preference heterogeneity. In the empirical analysis, I nevertheless follow Desmet, Ortuño-ortín, and Weber (2017) in assuming that it is *overall* diversity that matters for the willingness to pay taxes, and micro-found a fractionalization index to measure preference heterogeneity using electoral data.

The rest of this paper is structured as follows. Section 2 presents the baseline model and analyzes under what circumstances a tax ceiling, a compensating transfer scheme or fiscal decentralization may be necessary to stabilize efficient country borders. Section 3 provides cross-country empirical evidence for two central predictions of the model, namely that preference heterogeneity has a tax-reducing effect, and that it may both have centrifugal or centripetal effects. Section 4 calibrates the model to Belgium and explores how well its numerical predictions for equilibrium taxes, political heterogeneity and separatist vote shares correspond to observed data since its inception. Section 5 concludes.

2 A fiscal bargaining model for divided countries

This section develops a highly stylized fiscal bargaining model of collective choice in divided countries, where diversity involves both income inequality and political heterogeneity. The most important feature is that both richer voters and voters with more distinct policy preferences have a preference for lower income tax rates and smaller governments. This allows an analysis of how political heterogeneity affects tax rates, welfare and border stability, when taxes are decided by majority voting and voters in discontented regions within the country may threaten secession if not granted stronger fiscal autonomy.

To lay out my case, I consider the specific example of a rich region with distinct policy preferences contemplating secession to acquire more influence on the size and content of public policy - either within the union or as a newly formed country. This seems relevant, as the outside option generally is more attractive for wealthy regions due to their tax base advantage. I derive how the rich region's separatist threat depends on the degree of political heterogeneity in the rest of the country, through its effect on equilibrium tax rates. I subsequently discuss a number of fiscal policies that the central government may want to consider to stabilize efficient country borders and analyze their effectiveness in avoiding credible separatist threats when breakup would lower the sum of everybody's utilities. The model is straightforwardly generalized to study alternative settings, such as a poor region wanting to secede or fiscal bargaining under multiple secessionist threats.

2.1 The basic model

I consider a simplified version of the model developed in Vanschoonbeek (2020). In this model, I consider the determination of public expenditures in a country consisting of $R > 2$ regions, indexed by $r \in \{1, \ldots, R\}$, each inhabited by an identical population of I geographically immobile agents, indexed by $i \in \{1, \ldots, I\}$. Incomes are uniform within regions but are regionally stratified into one rich region, R, where per capita income equals y_R, and the remaining $(R-1)$ poor regions, where income equals $y_P < y_R$.

Agents derive linear utility from private consumption and isoelastic utility from the consumption of a single, non-rival bundle of public goods, g_c, provided by the central government and located on a particular point in the interval $[0, 1]$. Public services are financed by a non-distortionary proportional income tax, t_c, decided by majority voting. Regions have distinct policy preferences such that any deviation, $d_r = |g_r^* - g_c|$, from their preferred bundle linearly reduces the utility from public consumption. As a result, resident i of region r attains the following utility under the unified country:

$$U_{i,r} = (1 - t_c)\, y_{i,r} + \frac{(1 - d_r)\,(t_c Y_c)^\theta}{\theta} \tag{1}$$

where $\theta \in [0, 1]$ captures the common elasticity of marginal utility from public good consumption and $Y_c = \sum_{r=1}^{R} \sum_{i=1}^{I} y_{i,r}$ denotes total national income.

For simplicity, assume that agents form accurate expectations over the location of the public goods bundle before the start of any election, such that d_r is exogenously given for every agent, i.[3] I also make the extremely simplifying assumption that poor regions share an identical preference distance to the public good, such that $d_{r_1} = d_{r_2} = d_P \; \forall r_1, r_2 \neq R$, which is also assumed lower than the preference distance experienced in the rich region, implying that $d_R > d_P$. Although these assumptions are not necessary to derive the basic results, it greatly simplifies the analysis by ensuring that national tax rates will never be determined in the rich region, R, and that agents in the rich region will always favor lower tax rates. This allows us to derive the most preferred national tax rate of any agent i, $t^*_{i,r,c}$, conditional on their preference distance to the public goods bundle, d_r, as

$$t^*_{i,r,c} = \begin{cases} \left(\frac{1-d_P}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} & \text{if } r \neq R \\ \left(\frac{1-d_R}{y_R}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} & \text{if } r = R \end{cases} \tag{2}$$

Note that preference heterogeneity induces a redistributive conflict between the rich region, R, where richer and more politically distinct agents favor lower tax rates, and the rest of the country, $r \neq R$, where poorer agents with policy preferences more closely aligned to those of the majority of the electorate prefer higher taxes. As the equilibrium tax rate equals the one most preferred by the national median tax voter, poor regions can nevertheless exploit their population numbers advantage to set the prevailing tax rate at

$$t^*_c = \left(\frac{1-d_P}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} \tag{3}$$

As the only net contributor to the central budget facing the highest preference distance to public policy, the rich region may be tempted to contemplate secession, which would allow it to exercise greater influence on the content and size of government. To investigate the circumstances under which the rich region may credibly threaten to secede, I follow Bolton and Roland (1997) in assuming that a declaration of independence is decided by majority voting at the regional level and does not involve extra costs. Note that separation would allow the median tax voter in the rich region to set the equilibrium tax at

$$t^*_R = \left(\frac{1}{y_R}\right)^{\frac{1}{1-\theta}} Y_R^{\frac{\theta}{1-\theta}} \tag{4}$$

[3] Without this assumption, I would have to replace d_r in equation (1) with its expected value, $E[d_r]$, unnecessarily complicating the analysis.

implying that the utility of rich agents under union and secession is respectively given by

$$U_{i,R} = \begin{cases} (1 - t_c^*)\, y_R + \dfrac{(1 - d_R)\,(t_c^* Y_c)^\theta}{\theta} & \text{if } R \in c \\ (1 - t_R^*)\, y_R + \dfrac{(t_R^* Y_R)^\theta}{\theta} & \text{if } R \notin c \end{cases} \tag{5}$$

Consequently, separation will be preferred whenever $\Delta_{i,R} = U_{i,R}^{R \notin c} - U_{i,R}^{R \in c} > 0$, or

$$\Delta_{i,R} = (t_c^* - t_R^*)\, y_R + \frac{(t_R^* Y_R)^\theta - (1 - d_R)\,(t_c^* Y_c)^\theta}{\theta} > 0 \tag{6}$$

As noted in Vanschoonbeek (2020), the attractiveness of the outside option depends on three channels: scale economies in public goods provision, which decrease in relative regional income, $\frac{Y_r}{Y_c}$, and favor the preservation of the union; differing political views on the *content* of public policy, quantified by d_R and favoring separation by causing the political welfare costs emphasized by Alesina and Spolaore (1997); and differing fiscal views on the optimal *size* of government, captured by the differences between t_c^* and t_R^* and determining interregional redistributive conflicts over optimal country borders that were highlighted by Bolton and Roland (1997). Crucially, the latter implies that the rich region's welfare gain of separation *also* depends on the heterogeneity of the political views in the rest of the country, d_P, through its effect on equilibrium national tax rates, t_c^*, see equation (3). That is, since preference heterogeneity decreases utility from public consumption, it serves as a mechanism that reduces equilibrium tax rates. This implies that increasing preference heterogeneity may actually *reduce* the incentives to secede in the rich region, by bringing national income tax rates more in line with its fiscal preferences.

To see this, it is instructive to rewrite equation (6) to reflect that for a fixed preference distance in the rich region, $d_R = \tilde{d_R}$, preference heterogeneity in the rest of the country only affects the utility rich agents derive from the *status quo* scenario of preserving the union:

$$\Delta_{i,R} = \underbrace{U_{i,r}^{R \notin c}}_{\text{outside option}} + \underbrace{t_c^* y_R - \frac{\left(1 - \tilde{d_R}\right)(t_c^* Y_c)^\theta}{\theta}}_{\text{status quo}} \tag{7}$$

Appendix A.1 then establishes the following proposition,

Proposition 1

1. $\frac{\partial \Delta_{i,R}}{\partial d_P}$ *reaches a critical point at* $d_P = d_P^* = 1 - \left(1 - \tilde{d_R}\right) \frac{y_P}{y_R}$.
2. $\frac{\partial \Delta_{i,R}^2}{\partial^2 d_P} > 0$ *if* $\theta < \frac{1}{2}$.
3. $\left(\frac{y_R}{(1-\tilde{d_R})y_P}\right)^{\frac{\theta}{1-\theta}} \left(1 - \frac{\theta y_R}{(1-\tilde{d_R})y_P}\right) + \theta < 1 \Rightarrow \begin{cases} \Delta_{i,R} > 0 & if\ d_P = 0 \\ \Delta_{i,R} \leq 0 & if\ d_P = d_P^* \end{cases}$

The first result states that the rich region's net welfare gain of secession, $\Delta_{i,R}$, is a parabolic function of preference heterogeneity in the rest of the country and reaches a critical point at $d_P = 1 - \left(1 - \tilde{d_R}\right) \frac{y_P}{y_R}$. This establishes that an increase in preference heterogeneity can either increase or decrease the incentives for the rich region to secede, depending on the existing degree of preference heterogeneity in the rest of the country, d_P. The second result further specifies that as long as the intensity of preference for the public good, θ, is sufficiently low, $\Delta_{i,R}$ is a convex function of d_P. In such a case, increasing preference heterogeneity initially decreases the welfare gain of secession in the rich region until d_P reaches its critical value of d_P^*, after which further increases in political heterogeneity once again increase the risk of separatism. As Vanschoonbeek (2020) calibrates values of θ that all lie well below $\frac{1}{2}$ for a large number of European countries, in what follows, I will assume that in effect $\theta < \frac{1}{2}$. Finally, the third result confirms that intensifying preference heterogeneity in the rest of the country can actually reduce the incentives to secede in the rich region when a sufficient degree of heterogeneity is in fact necessary to lower national tax rates enough to make its outside option unattractive.

A direct corollary of proposition 1, which is also proven in appendix A.1, is that existing country borders can be stable only if there exists a range of values for d_P, $[d^*, d^{**}]$ for which neither rich nor poor regions would prefer secession over union. The following proposition formalizes this result and identifies these implicit threshold values.

Proposition 2

Existing country borders are stable if and only if $d^* \leq d_P \leq d^{**}$, *where*

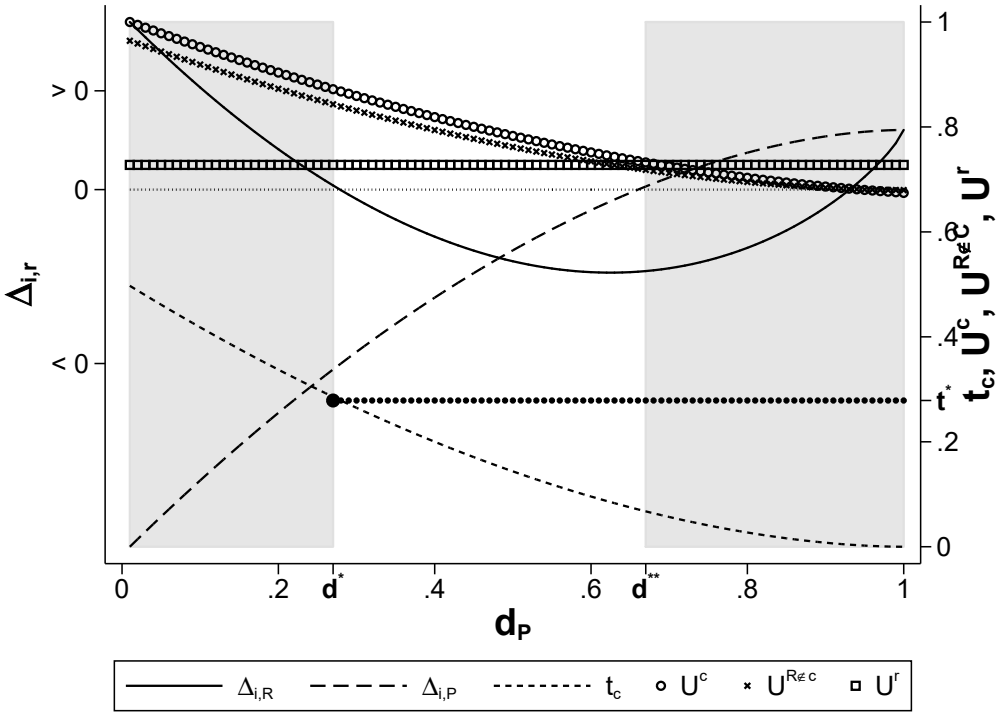
$$d^* \approx 1 - \left(\left[1 - \tilde{d_R}\right] - \left[\frac{\sum_{i=1}^{I} y_R \frac{y_P}{y_R}}{Y_c} \right]^{\frac{\theta}{1-\theta}} [1 - \theta] \right) \frac{y_P}{\theta y_R}$$

$$d^{**} = 1 - \left(\frac{\sum_{i=1}^{I} y_P}{Y_c} \right)^{\theta} \leq 1.$$

The implications of propositions 1 and 2 can be graphically summarized in figure 1. This figure illustrates how equilibrium income taxes and welfare gains of independence evolve as preference heterogeneity rises, for a calibrated model with parameter values

satisfying proposition 1.[4] A first straightforward result is that increasing preference heterogeneity reduces the utility derived from public policy under the union and, hence, the willingness to pay national income taxes, t_c. Consequently, total welfare under the *status quo* scenario, $U^c = \sum_{r=1}^{R} \sum_{i=1}^{I} U_{i,r}^{r \in c}$, steadily declines with preference heterogeneity. As preference heterogeneity is only assumed to occur between but not within regions, total welfare under complete state fragmentation, $U^r = \sum_{r=1}^{R} \sum_{i=1}^{I} U_{i,r}^{r \notin c}$, is not affected by preference heterogeneity and remains constant across the entire support of d_P.

Figure 1: **Fiscal bargaining, political heterogeneity and country border stability**



Note: This figure plots how the net gain of independence for the rich and poor regions, Δ_R and Δ_P, and the equilibrium national tax rate, t_c, evolve as poor regions face increased levels of preference heterogeneity, d_P. The hollow dots, crosses and hollow squares show the evolution of aggregate welfare under the preservation of the union (U^c), the secession of the rich region ($U^{R \notin c}$) and complete state fragmentation (U^r) as defined in equations (1) and (5), normalized to reach a maximal value of 1 on the right axis. t^* captures the maximal feasible tax rate poor regions can levy to keep the rich region in the union for low values of $d_P < d^*$. The grey areas depict levels of preference heterogeneity driving at least one region to declare independence according to equation (6), with d^* and d^{**} the corresponding critical values of d_P.

The shaded area in the right of the figure reproduces the standard finding that state breakup is unavoidable if preference heterogeneity becomes too large, $d_P > d^{**}$, as this leads to positive welfare gains of independence in both poor and (eventually) rich regions. Note that intense levels of preference heterogeneity imply large preference distances to public policy, reducing the utility derived from public consumption to such an extent that equilibrium national tax rates approach zero. In sharp contrast to the existing literature,

[4]More specifically, the numerical illustration uses the set of parameter values $(R, I, \theta, y_P, y_R, d_R) = (10, 1000, .45, 4000, 8000, .25)$, which suffices to derive all other necessary parameter values. It is easy to verify that these parameter values satisfy the last inequality of proposition 1.

however, the shaded area in the left of the figure shows that state break-up may *also* become unavoidable if preference heterogeneity is *too low*, $d_P < d^*$. The reason is that too large a consensus on public policy sufficiently increases poor agents' willingness to pay taxes to far outpace the maximal tax burden rich agents are willing to accept to remain in the union. Low preference heterogeneity may therefore leave the country vulnerable to separatism by inducing political disagreements over mutually acceptable tax rates.

This leads to the central argument of this paper, namely that under some circumstances *increasing political heterogeneity may actually increase country border stability*. More specifically, in our stylized example, the tax-reducing effect of increased preference heterogeneity in the rest of the country acts as a mechanism that brings (national) tax rates more in line with the fiscal preferences of the richest strata, thereby decreasing the welfare gains of secession for this particular group if initial heterogeneity is too low. Note that the existing literature is silent on potential *stabilizing* effects of preference heterogeneity.

One final important takeaway from figure 1 is that when preference heterogeneity is too low, total welfare is in fact higher under the preservation of the union, as the secession of the rich region would result in an aggregate welfare loss. Economic efficiency considerations then unambiguously favor the preservation of the union, raising the question whether the central government can enhance the efficiency and stability of country borders through an interregional redistribution of resources. The following subsections therefore analyze three potential interventions. A first strategy is to implement a tax ceiling, $\bar{t}^*$, that is sufficiently low to eliminate the attractiveness of the outside option for the rich region. In addition, the central government may also contemplate the introduction of fiscal transfers targeted at discontented regions to fully compensate the opportunity costs of their outside option. Finally, it may also aim to give its constituent regions a greater say in fiscal matters by a direct transfer of fiscal powers from the center to the regions.[5]

2.2 Stabilizing efficient country borders through a tax ceiling

The previous section may have painted an overly pessimistic picture of the risk of separatism in divided countries, as state breakup due to preference heterogeneity being too low turned out to reduce *aggregate* welfare. It is straightforward to see that the prerequisites for a Pareto-optimal transfer scheme are satisfied in such a scenario, as all agents could stand to gain from a redistribution of resources from poor to rich that succeeds in suppressing the rich region's separatist threat. One way of achieving this would be for poor agents to anticipate the rich region's secessionist threat if preference heterogeneity becomes too low and to commit to a tax ceiling, $\bar{t}_c^*$, that leaves rich agents indifferent between the outside option and the *status quo* whenever $d_P < d^*$. This corresponds to Buchanan and Faith's (1987) classical finding that the existence of a secession-option implicitly imposes

[5]While the latter measures are usually analyzed under the general heading of fiscal decentralization, compensating transfers are often considered a special case of so-called asymmetrical fiscal decentralization.

an upper limit on the tax burden a ruling elite can impose on a minority.

This intuition is confirmed in the following proposition, proven in appendix A.2.

Proposition 3

A tax ceiling of $\bar{t_c^*} = \left(\frac{1-d^*}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}$ *fully stabilizes efficient country borders.*

This result is easiest to understand by reconsidering considering the numerical example in figure 2. The grey lines in the figure reproduce the baseline findings as reference point, while the black lines show how these baseline results change under a tax ceiling, t_c^*. As can be seen, this tax ceiling is only binding when preference heterogeneity is too low such that the only substantive difference with the baseline scenario is that income taxes t_c^* will now be lower and constant whenever $d_P < d^*$. Compared to the baseline scenario, the tax ceiling induces a modest welfare cost for poor agents when it becomes binding, $\Delta_{i,P}^{t^*} < \Delta_{i,P}$, by preventing them to set their preferred income taxes and reducing government size. Importantly, these welfare costs are small enough to ensure that the net gain of secession for each poor region firmly remains negative. The tax ceiling also eliminates the welfare gains of independence for the rich region, $\Delta_{i,R}^{t^*}$, and effectively renders its inhabitants indifferent between the *status quo* and the outside option as intended. Finally, for low values of $d_P \approx 0$, aggregate welfare, U^{c,t^*}, is only slightly lower than it would be in the welfare-optimal disequilibrium of voluntary union, U^c. The tax ceiling thus effectively offers a Pareto-improvement over the baseline outcome ($U^{R \notin c}$ for $d_P \leq d^*$).

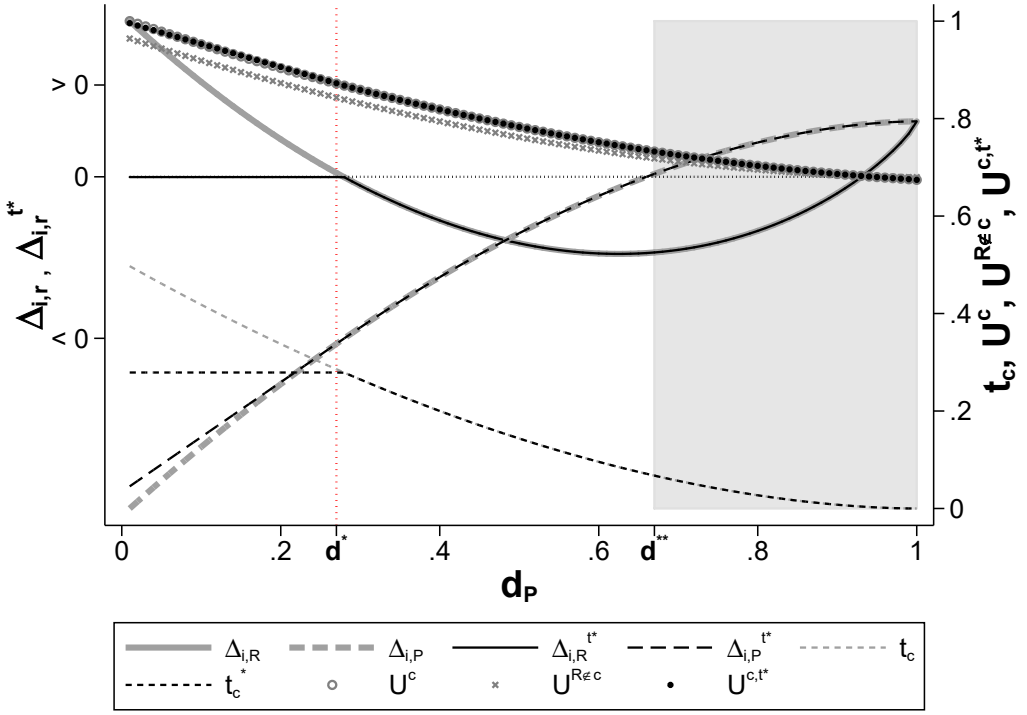
In principle, divided countries can thus adequately contain welfare-reducing separatist tendencies by committing to a suitable tax ceiling. Interestingly, this suggests that the threat of separatism may affect public finances even in the absence of explicit side payments or preferential treatments for secession-prone regions. In practice, however, this solution may be difficult to implement, for instance because political parties mainly garner electoral support in their respective regions and are therefore not inclined to accept regionally suboptimal tax rates.[6] In these scenarios, the central government may contemplate an alternative route of maintaining the majority voter's prerogative of setting income taxes but providing sufficient fiscal transfers to deter regions which would otherwise secede.

2.3 Stabilizing efficient country borders through fiscal transfers

Another way for the central government to stabilize efficient country borders would be to 'buy' the rich region's loyalty through an appropriate transfer from the poor regions that compensates their opportunity costs of secession. Indeed, poor regions will be collectively willing to pay the rich region a fiscal transfer, T_R, whenever their gain of keeping the rich region in is larger than the rich region's net gain of independence. This resonates with a large literature on assymetric federalism, which typically finds that side payments to

[6] In Belgium, for instance, the electorate can *only* vote for politicians belonging to their own language group during national legislative elections.

Figure 2: **Fiscal bargaining, political heterogeneity and a tax ceiling**



Note: This figure tracks how the net gain of independence for the rich and poor regions, $\Delta_{i,R}^{t^*}$ and $\Delta_{i,P}^{t^*}$, and the equilibrium national tax rate, t_c^*, evolve if the central government uses a tax ceiling to stabilize efficient country borders. The small black dots show the evolution of aggregate welfare under the tax ceiling, U^{c,t^*}, as defined in equation (11A), normalized to reach a maximal value of 1 on the right axis.

discontented regions may sometimes be necessary to avert inefficient state breakup.

The following proposition, proven in appendix A.3, confirms that this is the case:

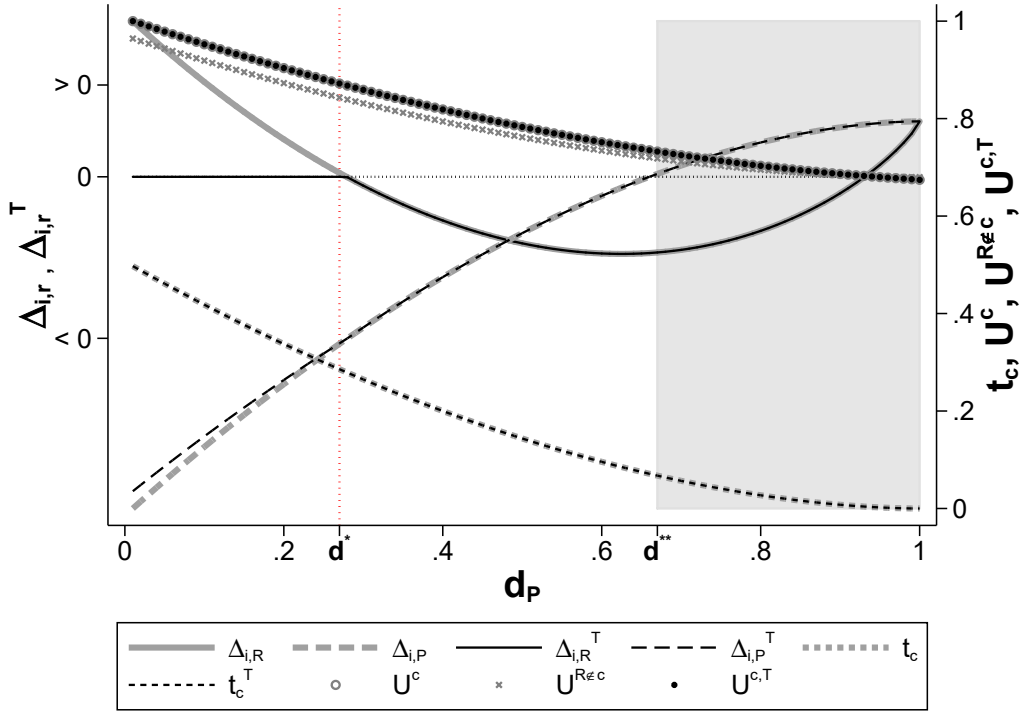
Proposition 4

A transfer of $T_R = \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}(R-1)} \left(\left[\frac{\sum_{i=1}^{I} y_R \frac{y_P}{y_R}}{Y_c} \right]^{\frac{\theta}{1-\theta}} [1-\theta] + [1-d_P]^{\frac{1}{1-\theta}} \left[\theta \frac{y_R}{y_P} - \frac{1-d_R}{1-d_P} \right] \right)$ *from each poor region to the rich region fully stabilizes efficient country borders.*

Graphically, figure 3 shows that the fiscal transfer scheme only kicks in when political heterogeneity is sufficiently low, $d_P < d^*$, and transfers are necessary to deter a credible separatist threat. The transfers leave the equilibrium tax rate unaffected, but decrease the utility of the poor regions paying for them, which slightly increases their net gain of independence, $\Delta_{i,P}^T > \Delta_{i,P}$. Nevertheless, poor regions still clearly benefit from the existing country borders and these transfers succeed at eliminating the gains of the inefficient outside option for the rich region, $\Delta_{i,R}^T$. As income taxes do not differ with those of the baseline model, aggregate utility is identical in both scenario's, in spite of the redistribution from poor to rich under the fiscal transfer scheme. This once again confirms that the transfer scheme effectively offers a Pareto-improvement over *laissez faire*.

The central government can thus also deter welfare-reducing separatist threats by suit-

Figure 3: **Fiscal bargaining, political heterogeneity and fiscal transfers**



Note: This figure plots how the net gain of independence for the rich and poor regions, $\Delta_{i,R}^T$ and $\Delta_{i,P}^T$, and the equilibrium national tax rate, t_c^T, evolve if the central government uses fiscal transfers to stabilize efficient country borders. The small black dots show the evolution of aggregate welfare under the transfer scheme, $U^{c,T}$, as defined in equation (18A), normalized to reach a maximal value of 1 on the right axis.

able use of fiscal transfers. The figure shows that this may even be necessary if countries as a whole grow *less*, as opposed to more, heterogeneous. Nevertheless, the example also clarifies why this solution may be difficult to implement in practice, as the required transfers are regressive and redistribute from *poor to rich*. In this light, among other reasons, Spolaore (2009, 2010, 2016) explains the rarity of so-called preference-based transfers in the real world by their potential incompatibility with other social goals of the government.[7]

Finally, the following proposition analyzes which policy option is more likely to occur in practice, by comparing the poor regions' median voter utility under both scenarios. More specifically, the proposition implies that poor regions will be more likely to favor a fiscal transfer scheme over a tax ceiling if their preference heterogeneity approaches zero. One potential explanation for this finding is that political consensus on the optimal location of the public goods bundle simultaneously increases the utility of public consumption in the poor regions as well as the attractiveness of maintaining control over government size, even if this requires considerable transfers to keep discontented regions in. On the other hand, whenever they also differ in their optimal location of public policy, poor regions may prefer the hard constraint of a tax ceiling to prevent fiscal spending that comes at

[7] Nevertheless, Haimanko et al. (2005, p.1278) argue that *"many federal countries have explicit transfers schemes to assist their disadvantaged regions"*, citing the examples of *"Russia, China, France, Italy, Belgium, Germany, Canada, Australia and many other countries"*.

an extra costs of increasing the necessary stabilizing transfers. Nevertheless, I also find the utility discrepancy between both policy options to be negligible, such that they are substantively interchangeable from an individual welfare perspective and their real-world attractiveness may be primarily driven by their practical political feasibility.

Proposition 5

The median voter prefers fiscal transfers over a tax ceiling whenever $d_P < d_P^*$*, where*

$$d_P^* \approx 1 - \left[\frac{(1-\theta)\left(\frac{\sum_{i=1}^{I} y_R}{Y_c}\right)^{\frac{\theta}{1-\theta}} - \left(\frac{y_R}{y_P}\right)^{\frac{\theta}{1-\theta}} \left[1 - d_R + \theta(1-d^*)^{\frac{1}{1-\theta}}\right]}{\left(\frac{y_R}{y_P}\right)^{\frac{\theta}{1-\theta}} \left(1 - \left[\frac{y_R}{y_P} + 1\right]\theta - [1-d^*]^{\frac{\theta}{1-\theta}}\right)} \right]^{1-\theta}.$$

2.4 Stabilizing efficient country borders through fiscal decentralization

Finally, as mentioned by Spolaore (2009, 2010, 2016), the central government can also try to enhance the stability and efficiency of country borders by a direct transfer of fiscal power from the center to the regions. The main idea here is that giving all regions a larger say in fiscal spending decisions may sufficiently reduce the heterogeneity costs the rich region faces when staying in the union to eliminate the attractiveness of its outside option. To formally analyze to what extent fiscal federalism can effectively stabilize efficient country borders, I assume that fiscal decentralization allows each region to spend a fixed share, D, of its tax contribution on the provision of a regional public goods bundle. For simplicity, I also assume that the utility derived from regional and national public consumption is additively separable and that the amount of decentralization is determined by majority voting. I also abstract from economies of scope and other externalities.

Appendix A.4 then clarifies that, under these assumptions, individual utility, equilibrium income taxes and the equilibrium level of fiscal decentralization respectively equal

$$U_{i,r}^D = \left(1 - t_c^D\right) y_{i,r} + \frac{\left(t_c^D Y_c\right)^\theta \left(\left[1 - d_r\right]\left[1 - D\right]^\theta + \alpha_r^\theta D^\theta\right)}{\theta} \tag{8}$$

$$t_c^* = t_c^D = \frac{Y_c^{\frac{\theta}{1-\theta}} \left(\left[1 - d_P\right]\left[1 - D\right]^\theta + \alpha_P^\theta D^\theta\right)^{\frac{1}{1-\theta}}}{y_P^{\frac{1}{1-\theta}}} \tag{9}$$

$$D^* = D_P^* = \frac{\alpha_P^{\frac{\theta}{1-\theta}}}{\alpha_P^{\frac{\theta}{1-\theta}} + (1 - d_P)^{\frac{1}{1-\theta}}} \tag{10}$$

with $\alpha_r = \frac{\sum_{i=1}^{I} y_r}{Y_c}$ region r's GDP share.

Using these results, appendix A.4 then also proves the following proposition:

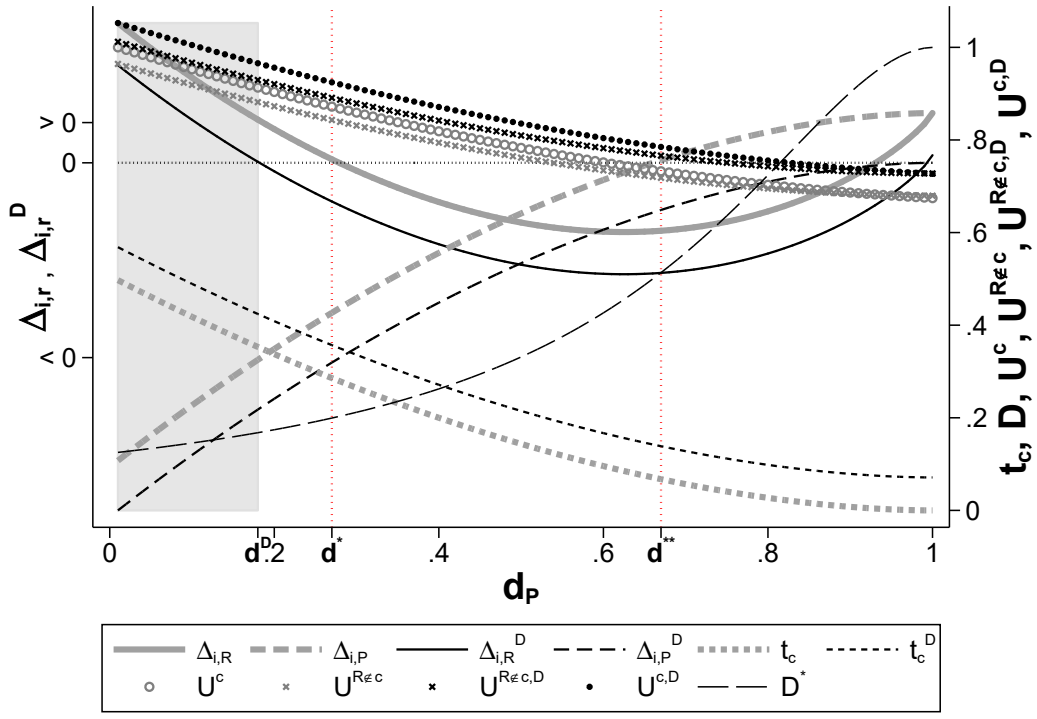
Proposition 6

Fiscal decentralization only stabilizes efficient country borders for $d_P \geq \underline{d^D}$*, where*

$$\underline{d^D} \approx 1 - \frac{1 - d_R}{\theta \frac{y_R}{y_P} + (1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}} + \alpha_P^{\frac{\theta}{1-\theta}}\left(\theta \frac{y_R}{y_P} - \frac{\alpha_R^\theta}{\alpha_P^\theta}\right)}$$

Once again, the intuition behind this finding can best be understood through the numerical example in figure 4. Several notes are in order. First of all, the figure shows that the poor region's median voter will always favor a positive degree of fiscal decentralization, D^*, in order to avoid the welfare costs of political heterogeneity at the country level. This also explains why the degree of decentralization unambiguously increases in preference heterogeneity, d_P, and reaches full decentralization when preference heterogeneity reaches its maximum, see equation (10). More importantly, by increasing the utility derived from public consumption, fiscal decentralization also clearly increases each region's willingness to pay taxes, t_c^D, driving up equilibrium income tax rates. As a result, fiscal decentralization both increases the efficiency of the union, $U^{c,D} > U^c$, as well as the welfare costs associated with the rich region's secession, $U^{c,D} - U^{R \notin c,D} > U^c - U^{R \notin c}$.

Figure 4: **Fiscal bargaining, political heterogeneity and fiscal decentralization**



Note: This figure tracks how the net gain of independence for the rich and poor regions, $\Delta_{i,R}^D$ and $\Delta_{i,P}^D$, and the equilibrium national tax rate, t_c^D, evolve under fiscal decentralization, D^*. The small black dots show the evolution of aggregate welfare in the federal union, $U^{c,D}$, as defined in equation (8), normalized to reach a maximal value of 1 on the right axis. The grey area depicts levels of preference heterogeneity driving at least one region to declare independence, with d^D the corresponding critical value of d_P.

Moreover, the welfare gain of decentralization is sufficiently large to prevent state breakup when political heterogeneity is large. Indeed, under fiscal federalism, no single region prefers secession over a federal union for any level of heterogeneity exceeding the *laissez-faire* upper threshold value of d^{**}. In line with much of the existing literature, I thus find clear evidence that fiscal federalism can not only enhance the efficiency of the union, but may also increase border stability in heterogeneous countries. In sharp contrast to the existing literature, however, I find that, though fiscal decentralization does also increase the border stability for intermediate levels of heterogeneity $d_P \in \left[d^D, d^*\right]$, it fails to stabilize efficient country borders if the country becomes too homogeneous. The reason for this slightly counter-intuitive finding is that discontented regions actually face a trade-off under fiscal decentralization. On the one hand,decentralization lowers the heterogeneity costs of the union and increases the utility per tax dollar spent, reducing the attractiveness of the outside option. On the other hand, equation (9) implies that decentralization may also increase the willingness to pay taxes and may thereby fuel existing conflicts over optimal taxation, potentially increasing the attractiveness of the outside option in discontented regions. In other words, fiscal decentralization may simultaneously reduce the discontented region's heterogeneity costs while increasing the discrepancy between preferred and levied taxes, with ambiguous net effects on country border stability.

Proposition 6 implies that if the rich region sufficiently differs from the rest of the country in terms of policy preferences and income distribution, this latter effect will tend to dominate for low levels of political heterogeneity, d_P , and no degree of fiscal decentralization fully stabilizes efficient country borders over the entire support of d_P. Thus, even under fiscal decentralization, inefficient breakups may still occur when political heterogeneity becomes too low. This point is graphically illustrated in figure A2, which shows that the rich region will still prefer secession over union for sufficiently low values of heterogeneity, d_P, *even if it can unilaterally decide on the degree of decentralization,* D^*.

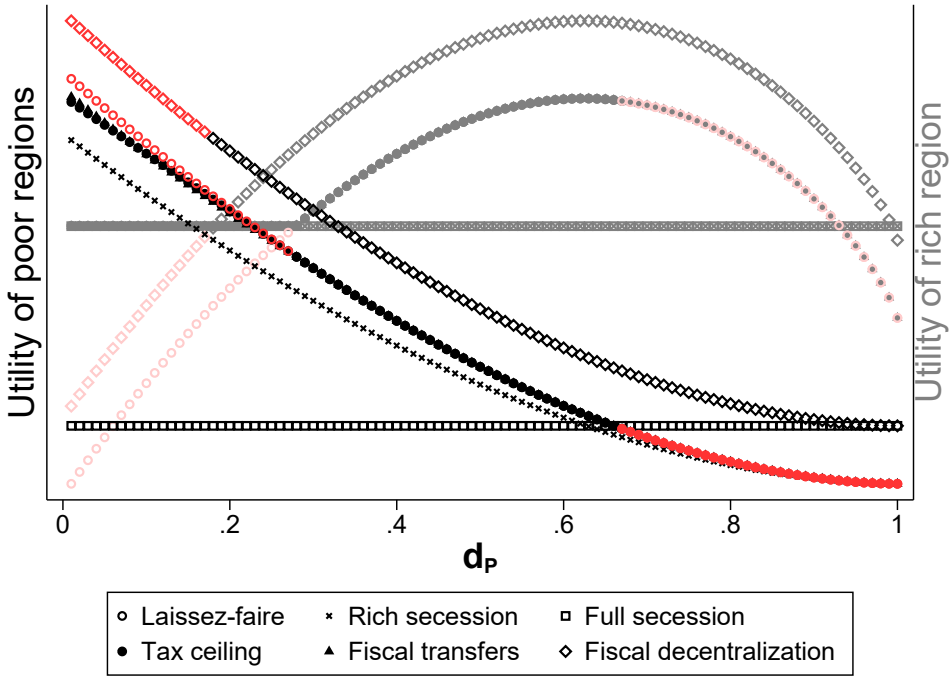
This leads to the paradoxical result that fiscal decentralization may simultaneously increase aggregate welfare while failing to fully stabilize efficient country borders. Consequently, it may be incapable of fully immunizing the country for welfare-reducing separatism. This implies that even decentralized countries may have to complement fiscal decentralization with a tax ceiling or fiscal transfers, to help keeping the genie of separatism in the bottle in times where too large a political consensus risks driving up equilibrium income taxes to such an extent that discontented regions are no longer willing to pay them.

2.5 Regional welfare implications

Figure 5 provides a bird's-eye view of the regional welfare implications associated with these various policy options for the particular numerical example considered here. Focusing first on the country's representative poor agent, fiscal decentralization clearly welfare-dominates any other scenario for them. Nevertheless, if political agreement in the poor

regions increases their willingness to pay taxes too much, state breakup becomes unavoidable. In that case, poor agents will recognize the value of keeping the rich region in and either agree to pay compensating transfers or to respect a tax ceiling for this purpose. In this particular example, the *laissez-faire* scenario would thus never materialize as long as country borders and income taxes are decided by majority voting. Finally, poor agents reach the highest possible utility when their policy preferences are perfectly congruent, or $d_P = 0$, and when they stabilize the union through voluntarily imposing a tax ceiling.

Figure 5: **Welfare implications**



Note: This figure plots how individual utility in poor (dark shades, left axis) and rich (light shades, right axis) regions evolves as the country becomes more and more divided, d_P. Utility under *laissez-faire*, the rich region's secession, full secession, a tax ceiling, fiscal transfers and fiscal decentralization are respectively computed through equations (1), (5), (5), (11A), (15A), (21A). Country border stability is identified by a black color while red colors identify situations where at least one region prefers secession over union.

Interestingly, the analysis differs for inhabitants of the rich region, which would also experience the highest utility under a federal union, but for intermediate levels of preference heterogeneity. The reason is that, all else equal, rich agents favor low income taxes and that preference heterogeneity, by reducing the utility from public consumption, induces the poor region's median tax voter to lower income taxes, moving them closer to the fiscal preferences of the wealthy. Their lack of majority control over fiscal policy also makes them most dependent on the availability of a credible separatist threat to ensure that they are compensated for any excessive heterogeneity costs from existing country borders. Absent such a threat, the rich region would be vulnerable to a tyranny of the majority that would leave it unable to prevent excessive taxation when political consensus on the content of public policy grows sufficiently large in the rest of the country.

3 Empirical evidence

The model described in the previous section suggests that fiscal policy is a crucial mediator between preference heterogeneity and country border stability. More specifically, increasing preference heterogeneity may have either centripetal or centrifugal effects, depending on whether it brings equilibrium tax rates more or less in line with the fiscal preferences of the most discontented regions. This section sets out to evaluate the real-world plausibility and relevance of this finding. To do so, it first micro-founds a quantitative measure of preference heterogeneity by developing a voting model that clarifies how a democratic country's political diversity can be inferred from electoral data. Using the resulting fractionalization index, this section subsequently complements a bivariate research design with multivariate analysis to explore the interrelation between preference heterogeneity, fiscal policy and secession proneness in the available data. The main goal of this exercise is not to identify causal relations between these channels, but merely to present reduced-form evidence that the basic correlations in the data conform to theoretical expectations.

3.1 Measuring preference heterogeneity using electoral data

Although preference heterogeneity plays an important role in determining both optimal and stable country borders in most of the existing literature, how to best measure it remains a topic of ongoing debate. Vanschoonbeek (2020) recently advocates the use of electoral data, as these are widely available and voting patterns during general elections should be indicative of societal (dis)agreement on politically salient topics in democratic countries. Lacking information on the preference distances *between* party programs, however, aggregating preference distances to public policy using party vote shares in a given election year becomes a non-trivial task. In line with the methodological discussion outlined in section 1, the crucial question is whether electoral data are best aggregated using indexes of fractionalization or polarization (Montalvo & Reynal-Querol, 2005). The former implicitly assume that political disagreement is at its most intense when many political parties are able to attract a small number of votes, while the latter implies that the most intense political disagreements arise when a large minority party faces a single majority party. Note that the model described in section 2 does not formally describe how the location of the public goods bundle, g_c, is determined and thus does not offer any guidance on how to select an appropriate electoral measure for preference heterogeneity.

The current subsection therefore relies on the collective action framework of Banerjee et al. (2007) to put some more structure on this model and to specify how the location of the public goods bundle is determined in an initial stage. In general, this extended model can thus be thought of as a two stage process. In a first stage, political parties contest elections during election years based on a party platform defined over the preferred location of the public goods bundle, before forming a government that credibly commits to a specific location of the public good during a subsequent governing period. In a sec-

ond stage, the government tailors fiscal policy as described in section 2, conditional on the equilibrium location of public policy. The main advantage of the additional assumptions in this extended model is that they unambiguously favor a fractionalization over a polarization index as the most appropriate electoral measure of preference heterogeneity.

To fix ideas, assume that political parties, indexed by p, can freely contest general elections based on a party platform defined over their preferred location of the public goods bundle, g_p^*. In deciding whether to compete elections, aspirant parties trade off the benefits of competing - the share of votes they expect to attract - against the cost of entry. Agent's can decide whether or not to vote but, conditional on voting, each agent votes for the party platform closest to its preferred policy bundle such that for any two parties, p_1 and p_2, agent i will vote for party p_1 whenever $\left|g_i^* - g_{p_1}^*\right| \leq \left|g_i^* - g_{p_2}^*\right|$. Conditional on the other party platforms, this implies that the electoral base of potential voters for party p_1, π_{p_1}, can be written as $\sum_{i=1}^{I} \mathbf{M}_{|g_i^* - g_{p_1}^*|}$, with $\mathbf{M}$ an indicator equal to 1 if agent i's preferred location of public policy is closest to that of party p_1 and 0 otherwise.[8]

Political parties compete to influence the location of the public goods bundle, g_c, by becoming pivotal in the formation of the government after the election took place. The probability that a political party becomes pivotal is assumed to crucially depend on its electoral influence, or the total number of votes, V_p, it is able to attract during the election. If a political party becomes pivotal in forming the government, the influence it has on the location of the public goods bundle is implicitly defined by $f_p\left(\pi_G, g_G^*\right)$, with π_G and g_G^* the $(1 \times G)$ vectors containing the voter bases and preferred locations of all G pivotal parties in forming the government. For simplicity, assume that pivotality is always welfare-increasing for the potential voters of a party, because political parties will only influence government policy if doing so moves government policy closer to their preferred position in comparison to the outside option. Conditional on pivotality, the positive welfare gain of party p's influence on the location of the public goods bundle for one of its potential voters, i, is denoted by $G_i = f\left(g_i^*, \pi_{p_p}, \pi_G, g_G^*\right) > 0$.

Finally, agents have to expend effort to vote resulting from the opportunity costs of having to stay politically informed and to express their votes during the election. Denote the effort put in by agent i by e_i and assume that voting only occurs if this effort level exceeds a certain threshold value, $e_i \geq \bar{e}$. Assuming quadratic effort costs, e_i^2, the net benefit of voting to agent i with preferred party p_1 takes the following form

$$e_i \pi_{p_1} G_i - e_i^2 \tag{11}$$

where the inclusion of π_{p_1} captures that it is easier for a party with a larger voter base to obtain a given electoral influence, as each member has to expend less effort to do so.

[8] Note that the primary purpose of this model is to provide a theoretical justification for the way to measure preference heterogeneity *conditional on observed voting patterns*. Therefore, there is no need to formally model the political party entry decision or, in the next paragraph, government coalition negotiations.

It follows that the equilibrium level of effort expended by this agent amounts to

$$e_i^* = \frac{\pi_{p_1} G_i}{2} \tag{12}$$

where it is clear that agents will be incentivized to put in more effort if their preferred party has a large voter base, π_{p_1}, and has a high probability to become pivotal after the election; *or* if the welfare gain of pivotality, G_i, is large because the agent's preferred party has very distinct policy preferences in comparison to the other parties contesting elections.

As such, given the threshold value of effort $\bar{e}$ required to vote, potential voters for minority parties with small voter bases will not bother to express their vote *unless* their preferred party has a very distinct preferred location for public policy. Potential voters of large parties will be tempted to express their vote *even* if their opinions do not differ much from the majority in the rest of the country, because the probability of pivotality is large. Thus, the main takeaway from this highly stylized model is that if voting is costly and if the benefits of voting in terms of policy influence increase in a party's voter base, votes for smaller political parties are in fact more informative of the existence of preference heterogeneity. This model thus provides a theoretical justification for using a fractionalization index to infer preference heterogeneity from electoral data.[9]

In what follows, preference heterogeneity is therefore measured as the probability that two randomly chosen individuals voted for a different political party during general elections. I include only democratic countries, as electoral variation is meaningful only for those countries.[10] More specifically, when a total of P parties competed in the national election that took place at time t, I define electoral fractionalization of democratic country c at time t, $EF_{c,t}$ as the normalized Herfindahl index of party vote concentration[11]:

$$EF_{c,t} = 1 - \frac{\sum_{p=1}^{P} v_{c,t,p} - \frac{1}{P}}{1 - \frac{1}{P}} \tag{13}$$

where $v_{c,t,p}$ denotes the vote percentage for party p in country c in election year t and higher values correspond to election years with stronger electoral disagreements over which political party should be in power. Implicitly assuming that political disagreements evolve gradually over time, I linearly interpolate non-election years at the country-level. All electoral data is taken from Kollman, Caramani, Backer, and Lublin (2019).

3.2 Bivariate analysis

I first bring central equations (3) and (9) to the data, which posit that preference heterogeneity has a tax-reducing effect by decreasing the utility from public consumption. Taxes

[9] In fact, if anything, this voting model would predict the willingness to pay taxes to *increase* in electoral *polarization* and this is also borne out in the data, see appendix B.

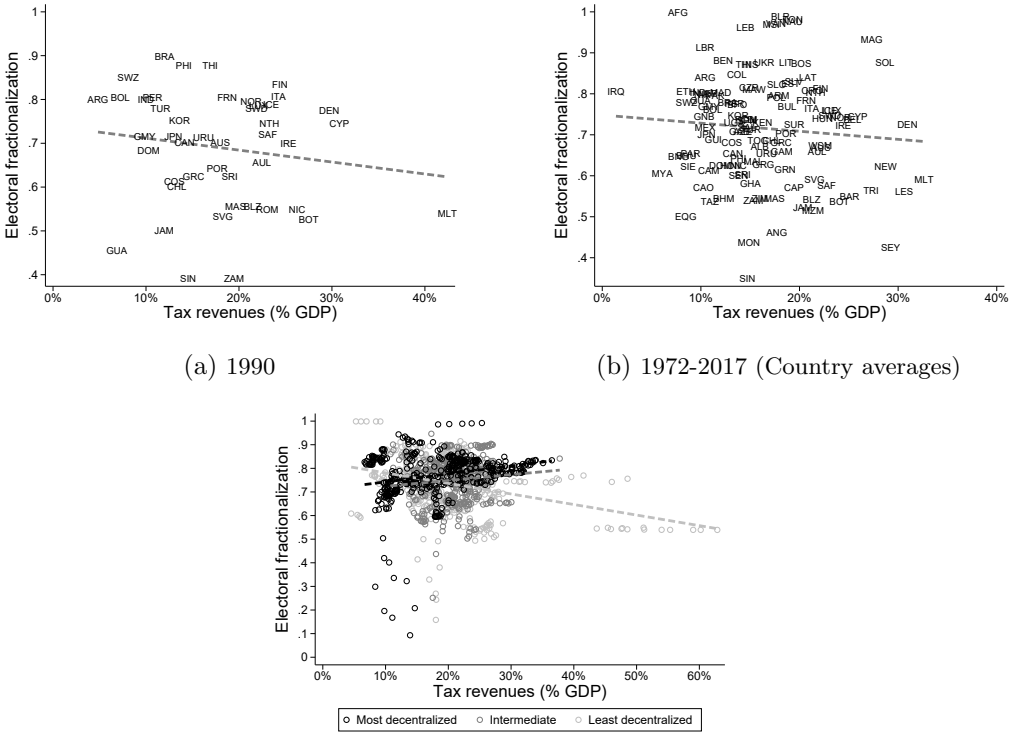
[10] In practice, this amounted to the removal of reported electoral results for four theocracies and semi-democracies: Bahrain, Iran, Kuwait and the United Arab Emirates.

[11] The Herfindahl index is normalized to accommodate that the number of parties differ across elections.

are measured as the tax revenue share of GDP, as reported by the World Bank (2019).

Figures 6a and 6b reveal that tax revenue shares were negatively correlated with electoral fractionalization in the previous half-century. Ignoring potential confounders, these figures offer suggestive evidence that political disagreement on public policy may partially explain fiscal policy choices, where more intense disagreements lead to lower equilibrium taxes. Note that equations (3) and (9) also suggest that the tax-reducing effect of political heterogeneity should be less pronounced in fiscally decentralized countries. To verify to what extent this prediction is also borne out in the data, figure 6c shows the relation between government size and electoral fractionalization separately for three tertiles of tax revenue decentralization, when fiscal decentralization is measured as the subnational tax revenue ratio reported by the IMF (2019).[12] Interestingly, the results in this smaller sample also conform to theoretical expectations, as political fractionalization is only negatively related to government size in the most centralized countries. In federal countries, the relation even turns slightly positive, consistent with the idea that fiscal decentralization increases the willingness to pay taxes by bringing government closer to the people.[13]

Figure 6: **Electoral fractionalization and government size (1972-2017)**



(a) 1990

(b) 1972-2017 (Country averages)

(c) 1972-2017

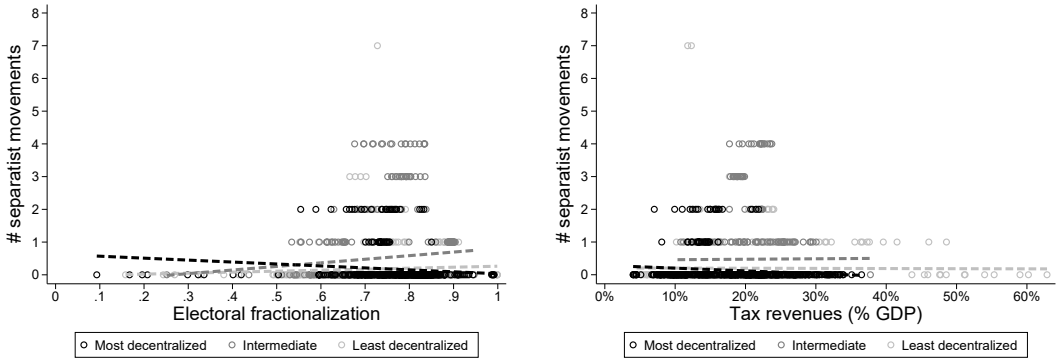
Note: This figure shows the bivariate relation between preference heterogeneity, quantified through equation (13), and the share of tax revenues in GDP for a sample of 125 countries. Figure 6a shows the correlation in the year 1990; figure 6b shows the correlation when both variables are averaged over 1972-2017 for each available country; figure 6c shows the correlation separately for each tertile group of fiscal decentralization.

[12] The tax revenue ratio is computed as the share of local and state to general government tax revenues.

[13] Another plausible explanation lies in the duplication of government agencies, e.g. Germany has 17 Ministries of Education. Nevertheless, column 4 of table 1 finds no evidence that subnational government payroll shares meaningfully affect the relation between taxes and centralization in divided countries.

Subsequently, note that proposition 1 implies that secession proneness should become especially salient at high levels of preference heterogeneity. This is confirmed in figure 7a, which plots electoral fractionalization against the number of independence movements in each country, finding that separatist movements almost solely emerge in highly divided countries. However, I also find ample evidence of politically fragmented countries that nevertheless lack separatist threats, pointing to the potentially stabilizing effects of political heterogeneity. Additionally, figure 7b shows that, conditional on having independence movements, their numbers increase in the tax revenue share. Finally, the finding that separatism is primarily prevalent in countries with low tax revenues also resonates well with the model, which posits that electoral fractionalization simultaneously exerts a tax-reducing effect and functions as an important driver of separatist tendencies.

Figure 7: **Separatism, electoral fractionalization and taxes (1972-2017)**



(a) 1990 (b) 1972-2017 (Country averages)

Note: This figure shows the relation between the number of independence movements, the tax revenue share in GDP and electoral fractionalization, separately for each tertile of fiscal decentralization. The sample consists of 60 countries and covers the period 1972-2017. The number of independence movements are taken from Griffiths (2015).

3.3 Multivariate analysis

The previous subsection provided suggestive evidence that preference heterogeneity has a tax-reducing effect and is meaningfully related to separatist tendencies in present-day countries. The fiscal bargaining model of section 2 implies that electoral fractionalization therefore may both increase and decrease secession proneness, depending on its fiscal effects. To more carefully evaluate this claim, I now turn to multivariate analysis.

In addition to the variables already discussed, data on population size and GDP come from The Maddison Project (2017), trade openness is obtained from Feenstra, Inklaar, and Timmer (2015), democracy scores are taken from Vanhanen (2000), income inequality is measured through the gini indices reported by Van Zanden, Baten, Foldvari, and Van Leeuwen (2014) and the World Bank (2019), and subnational payroll shares are computed from IMF (2019).[14] Finally, I measure fiscal centralization through three dummy variables identifying each tertile of tax revenue centralization, where tax revenue centralization is

[14] All missing country-year observations are linearly interpolated whenever possible.

computed as one minus the subnational tax revenue ratio.[15]

Table 1 then presents a first set of multivariate regression results.[16] The first column clearly confirms that tax revenues are inversely related to electoral fractionalization, even when controlling for other variables that may plausibly affect fiscal policy. The second column confirms that the effect of preference heterogeneity seems to be mediated by the institutional structure and that fiscally decentralized countries are less sensitive to the tax-reducing effect of preference heterogeneity. The third column suggests, for a slightly smaller sample, that these findings are not produced by cross-country differences in income inequality. The last column suggests that the increased willingness to pay taxes in decentralized countries is not an artifact of the duplication of government agencies, as it remains visible when controlling for the subnational share in the government payroll.

Table 1: **The tax-reducing effect of electoral fractionalization**

Channel	*(1)*	*(2)*	*(3)*	*(4)*
Electoral fractionalization	-.045***	-.198***	-.11***	-.093**
	(.012)	*(.04)*	*(.036)*	*(.043)*
Intermediate decentralized		-.09***	-.043	.009
		(.034)	*(.03)*	*(.037)*
Most decentralized		-.271***	-.191***	-.218***
		(.034)	*(.033)*	*(.068)*
Electoral fractionalization × Intermediate decentralized		.112***	.06	-.012
		(.043)	*(.038)*	*(.047)*
Electoral fractionalization × Most decentralized		.302***	.204***	.252***
		(.045)	*(.043)*	*(.084)*
Population size (log)	-.026***	-.027***	-.03***	-.041***
	(.002)	*(.004)*	*(.003)*	*(.004)*
GDP (log)	.01***	.013***	.016***	.031***
	(.002)	*(.003)*	*(.003)*	*(.004)*
Trade openness	.005	.006	-.006	-.003
	(.004)	*(.005)*	*(.005)*	*(.006)*
Democracy	.001	.002	.001	.003
	(0)	*(0)*	*(0)*	*(0)*
Income inequality			0	.002
			(0)	*(0)*
Subnational payroll share				-.019***
				(.006)
Observations [# countries]	2165 [94]	1243 [55]	1132 [53]	461 [30]
Adjusted R-squared	.324	.365	.286	.441
Year fixed effects	Yes	Yes	Yes	Yes

Note: Bootstrapped standard errors, based on 500 iterations, between brackets.

Before turning attention to the empirical relation between electoral fractionalization and secession proneness, it is useful to point out that the available data is insufficient to fully validate the model, as doing so would require counterfactual information on how

[15] Note that total tax revenues, T, appear both in the dependent variable, the tax revenue share or $\frac{T_{c,t}}{GDP_{c,t}}$, and the fiscal centralization index, the central government tax revenue share or $\frac{T^c_{c,t}}{T_{c,t}}$, with T^c total central government tax revenues. I therefore use a categorical rather than a continuous fiscal centralization index, to avoid capturing the mechanical (and hence uninformative) correlation that may arise from small changes in total tax revenues, T, simultaneously increasing the measured willingness to pay taxes, $\frac{T_{i,t}}{Y_{i,t}}$, and decreasing measured fiscal centralization, $\frac{T^c_{i,t}}{T_{i,t}}$. Nevertheless, appendix table A2 shows that the results are qualitatively the same when using a continuous fiscal centralization index instead.

[16] As several variables of interest tend to be rigid and path-dependent, e.g. due to mandatory spending in fiscal policy, I refrain from including country fixed effects and focus on cross-country comparisons.

secession proneness would have evolved in our sample under different degrees of electoral fractionalization. Nevertheless, I can indirectly validate the model by leveraging its prediction that if electoral fractionalization also has stabilizing effects, a simple regression of secession proneness on electoral fractionalization would understate its centrifugal effects. The reason is that *only* when electoral fractionalization does *not* affect tax rates does its potential for stabilizing fiscal effects disappear, rendering the net welfare gain of secession in equation (6) a strictly increasing function of electoral fractionalization. In other words, only when holding the tax rate constant does the model unambiguously predict that electoral fractionalization increases secession proneness.

Table 2 reports the results of this exercise when secession proneness is measured as the number of independence movements in each country. The first column already finds evidence that electoral fractionalization is positively related to the incidence of separatism in our sample. More interestingly, the second column confirms that the centrifugal effect of electoral fractionalization becomes larger once we control for its potential fiscal effects by holding tax revenue shares constant. Moreover, independence movements also turn out to be more prevalent in high-tax countries, suggesting that the tax-reducing effect of fractionalization may render countries less susceptible to separatism. The last column finds empirical evidence that the tertile of fiscally most decentralized countries are less sensitive to the destabilizing effects of political fragmentation, in line with model predictions.

Table 2: **Electoral fractionalization, government size and secession proneness**

Channel	*(1)*	*(2)*	*(3)*
Electoral fractionalization	.309**	.405***	.239
	(.145)	*(.132)*	*(.215)*
Tax revenue share		2.662***	1.276***
		(.294)	*(.298)*
Intermediate decentralized			-1.23***
			(.337)
Most decentralized			.305
			(.291)
Electoral fractionalization × Intermediate decentralized			1.944***
			(.458)
Electoral fractionalization × Most decentralized			-.621*
			(.353)
Population size (log)	.785***	.852***	.135**
	(.065)	*(.07)*	*(.056)*
GDP (log)	-.327***	-.359***	.148***
	(.046)	*(.048)*	*(.052)*
Trade openness	.444***	.422***	.096**
	(.05)	*(.052)*	*(.048)*
Democracy	.009***	.006**	-.006
	(.003)	*(.003)*	*(.005)*
Income inequality	-.026***	-.026***	-.004
	(.003)	*(.003)*	*(.004)*
Observations [# countries]	1895 [90]	1895 [90]	1132 [53]
Adjusted R-squared	.379	.395	.244
Year fixed effects	Yes	Yes	Yes

Note: Bootstrapped standard errors, based on 500 iterations, between brackets.

4 Numerical illustration for Belgium

The previous section finds empirical evidence that fiscal policy may be a crucial factor that determines the relation between country border stability and political heterogeneity. Nevertheless, the available data only allow for an indirect validation of some of the model's most important predictions, yet fall short of a conclusive verification of its accuracy in particular case studies. Therefore, this section supplements the large-scale cross-country statistical analysis with a numerical application to the case study of Belgium.

This section has four aims. First, it wants to better understand how the model relates to a number of historical forces that are often thought to be important drivers of the historical evolution in the number and sizes of nations. To this end, I adapt the model to account for a number of stylized facts that are related to its different subcomponents as they played out over the past two centuries: the worldwide democratization process; the gradual increase in government size; the recent decentralization wave; and the most recent intensification in political fragmentation. Second, in addition to tracking how the interplay of these global forces is predicted to affect country border stability, I also want to assess the accuracy of the model's predictions using real-world data. I therefore compare the model's predictions for equilibrium taxes, political heterogeneity and separatist vote shares in Belgium for the past 200 years with actual historical data. Third, I want to use the model to analyze whether political heterogeneity on balance had centrifugal or centripetal effects in the Belgian case. I develop a comparative static exercise for this purpose and contrast benchmark results with alternative predictions obtained by resolving the model without preference heterogeneity. Finally, I also want to illustrate how the general model summarized in section 1 can be taken to the data and can be suitably adapted to reflect the idiosyncratic characteristics of particular case studies.

Belgium is an interesting case for a number of reasons. Since its inception in 1830, it is known as a hotbed of longstanding linguistic bickering between the northern Dutch-speaking region of Flanders and the southern French-speaking region of Wallonia. Moreover, in the past two centuries, the country was the scene of a dramatic reversal of fortune, where Flanders for a long time was Belgium's poorest region but economically surpassed Wallonia in the immediate postwar period (Buyst, 2010).[17] Interestingly, separatism shone in its absence throughout most of the country's first hundred years of existence, when the economic cards still favored Wallonia, yet emerged as an increasingly important political factor ever since the interwar period, when Flanders began economically gaining ground. The primary reaction to the northern separatist threat then followed in 1970, when Belgium embarked on a journey towards federalism that progressively moved it from a unitary to a federal state (Gerard, 2014). In spite of these centripetal efforts, after the most recent 2016 election, the biggest national political party has included the strive for an independent Flanders as the first point in its party statutes (Nieuw-Vlaamse Alliantie, 2015).

[17] See also figure 8f.

Finally, in terms of stylized facts, the Belgian experience seems representative for the West more generally, such that the model's predictions for Belgium also shed some light on the broader question of how these factors shaped global patterns of state formation.

In what follows, I first motivate five historical stylized facts, relate them to the model and show how they played out in Belgium between 1830 and 2016. I then explain how I take model to the data to account for these stylized facts, solve it numerically for the case of Belgium and contrast its predictions with observed historical data.

4.1 Stylized facts

The model described in section 2 fundamentally describes the collective choice of equilibrium tax rates and country borders in democratic societies. As such, it is intrinsically linked to the democratic regime, which determines which actors are politically salient at any given point in time. In this regard, it is well-known that the suffrage was massively extended in the West between 1800 and 1945, progressively giving more voting rights to the poor. Acemoglu and Robinson (2000) moreover argue that this process was intimately related to fiscal preferences, tracing the economic roots of the 19^{th} century European democratization process to the contemporary increase in income inequality, which pressured elites to extend the suffrage to avoid social unrest and revolution. Figure 8a shows that democracy came to Belgium in four waves: from 1830 to 1893, census suffrage implied that only a small number of wealthy people paying a minimum amount of taxes could vote; from 1893 to 1918, there was a system of plural male voting allowing only men to vote and giving supplementary votes to wealthy and well-educated voters; equal male voting between 1918 and 1949 literally implemented the 'one man one vote' principle; and, finally, universal suffrage from 1949 onwards extended electoral rights to women. Conventional wisdom dictates that democratization increases secession proneness by 'activating' more and more diverse policy preferences, making it more difficult for the central government to satisfy a majority of the electorate with any given public goods bundle (Alesina & Spolaore, 1997, 2003). One potential explanation of the emergence of separatism in postwar Belgium that is consistent with the model, is that democracy magnified political heterogeneity and hence the welfare costs of centralized public policy provision.

As is also well-known, this democratization process has been accompanied by a steady increase in government size. This too is related to Acemoglu and Robinson's (2000) argument that democratization opened the way for redistribution, mass education and other social policies. On the empirical front, Tanzi and Schuknecht (2000) collect historical data on public expenditure shares for 17 industrialized countries and find that they gradually increased from 10.7% of GDP in 1870 to 45.6% in 1996. Looking at the tax revenue share, figure 8b shows that this trend also manifested itself clearly in Belgium, where the data plotted corresponds to historical estimates from Flora, Kraus, and Pfenning (1983) and more recent estimates from the OECD (2019). In terms of the model, democratization

may thus not only have directly contributed to Belgium's increasing secession proneness, by increasing political heterogeneity, but also indirectly, by increasing equilibrium tax rates and hence the relative importance of public to private consumption.

Another variable of interest are policy preferences, which appear to have grown more diverse after World War II. Inglehart (1971, 1981) argues that Western governments are increasingly confronted with a new arsenal of so-called post-materialist policy preferences, ranging from environmental concerns to women's rights, which have risen in importance due to drastic postwar improvements in economic security. In similar vein, Dalton (1996) and Inglehart and Norris (2016) fear that the postwar proliferation of political parties increasingly complicate policy formulation in Western democracies. Similar worries fuel contemporary discussions on polarization and populism, see for instance Rodrik (2017), Gennaioli and Tabellini (2018) and Rodríguez-pose (2018). Figures 8c and 8d show that Belgium grew politically more diverse and saw both a drastic increase in the number of political parties, including the environmental party in 1979, as well as in the heterogeneity of spatial voting patterns.[18] Increasing economic security may thus have widened the bandwidth of policy preferences in Belgium, magnifying the preference distance to centrally provided public policy and decreasing the willingness to pay taxes.[19]

Finally, it is relevant to note that separatism only really entered the world stage after World War II, following a brief peak during the 1918 Wilsonian moment, and started with the postwar decolonization wave. Griffiths (2014, 2016) argues that increased trade openness, military integration and a more favorable recognition regime all reduced the disadvantages of being small, ushering in an 'age of secession'. Similarly, Gancia et al. (2017) contend that the removal of trade barriers after World War II reduced optimal country size by making economic performance less dependent on the size of the domestic market. Figure 8e shows that this trend is also apparent in Belgium, where separatism remained politically irrelevant throughout the 19^{th} century in spite of the economic inequality between the south and the north, while separatist parties accrued increasingly large voting shares in Flanders throughout the 20^{th} century.[20] This begs the question whether the model can explain the timing and location of separatism in Belgium, and especially the absence of Walloon separatism in the 19^{th} century, when transfers moved from rich Wallonia to poor Flanders. Gerard (2014) offers a partial explanation that resonates

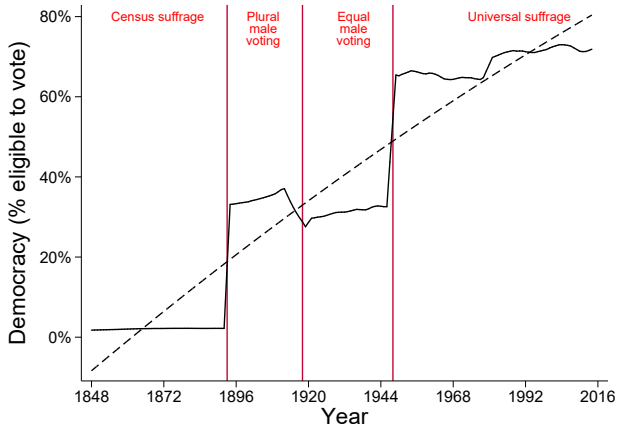
[18] Spatial electoral heterogeneity in figure 8d is computed from equation (13). Following Vanschoonbeek (2020, p. 19), votes are aggregated over party families after 1968 to account for the linguistic split of most Belgian political parties in Flemish and Walloon wings, see also Caramani (2004, p. 153).

[19] Lake and O'Mahony (2004, p.713) summarize that *"as the role of the government in the economy and society has expanded, more policies affect more individuals. Even if latent preferences have remained unchanged over time, more and more policy preferences are now activated and brought into play politically. This suggests that 'effective' preference heterogeneity may have increased over the 20th century"*.

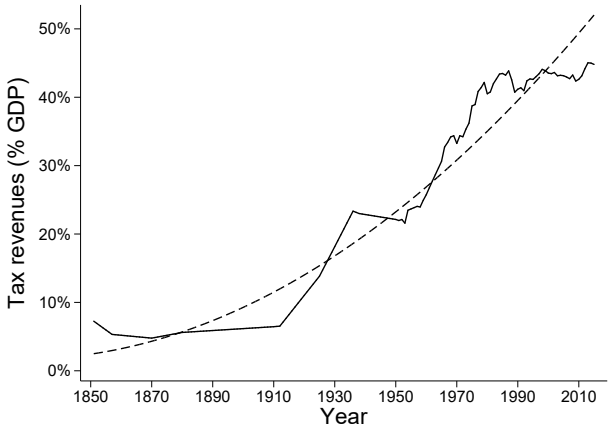
[20] Although it is inherently difficult to define nationalist and separatist political parties in practice, for the purpose of this exercise, I consider the following political parties as separatist (including their predecessors whenever applicable): the Christian Flemish People's Union (*'Christelijke Vlaamse Volksunie'*), Flemish Interest (*'Vlaams Belang'*), Flemish Concentration (*'Vlaamse Concentratie'*), Flemish National Union (*'Vlaams Nationaal Verbond'*), the Front Party (*'Frontpartij'*), the New Flemish Alliance (*'Nieuw-Vlaamse Alliantie'*) and the People's Union (*'Volksunie'*).

well with our central finding, namely that the transfers from Wallonia to Flanders were comparatively small due to the less developed social security system, see also figure 8b.

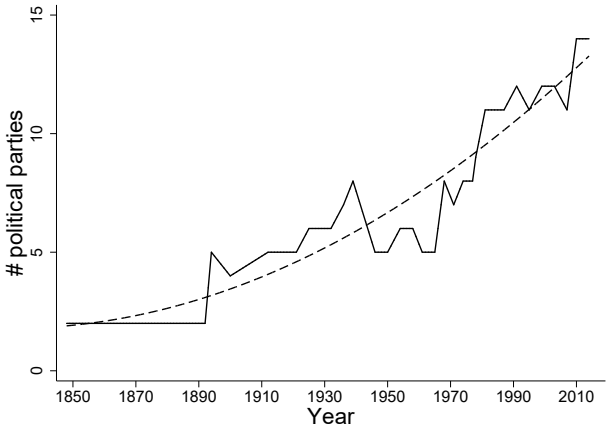
Figure 8: **Stylized facts for Belgium**



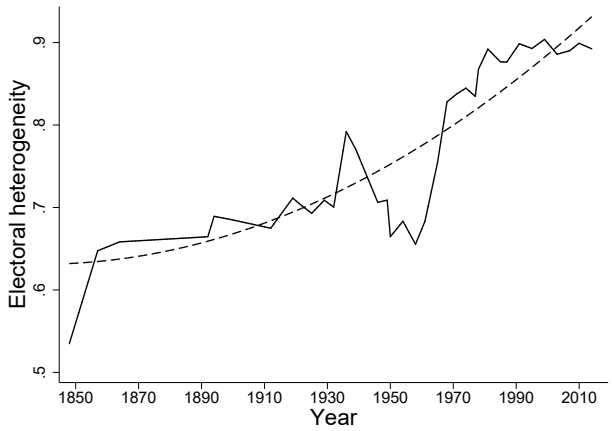
(a) Democratization waves in Belgium (1848-2019)



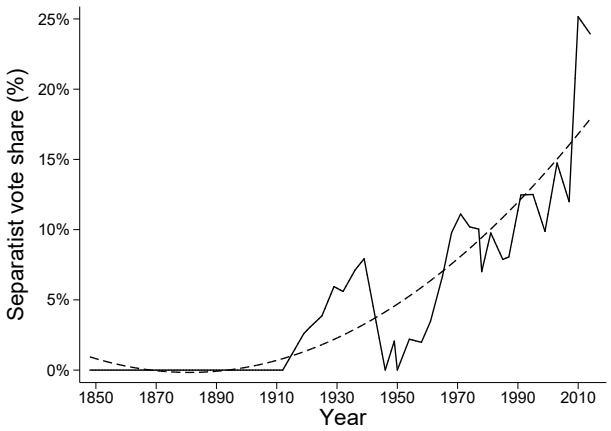
(b) Tax revenue shares in Belgium (1850-2016)



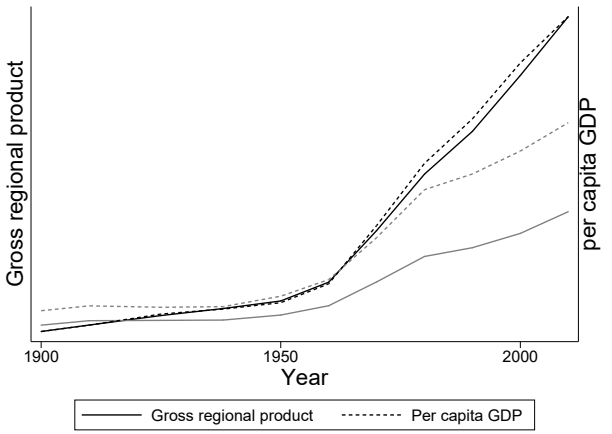
(c) Political parties in Belgium (1848-2016)



(d) Political heterogeneity in Belgium (1848-2016)



(e) Separatist vote shares in Belgium (1848-2016)



(f) Regional and per capita GDP in Belgium (1900-2016)

Note: Figure 8a shows the population share of eligible voters from the first known Belgian election in 1848 until the last one in 2016, based on the electoral data from Kollman et al. (2019). Figure 8b shows the tax revenue share in Belgium between 1870 and 2016, combining the information reported by Flora et al. (1983) and OECD (2019). Figure 8c tracks the evolution in the number of politically relevant political parties in Belgium between 1848 and 2016, based on the electoral data from Kollman et al. (2019). Figure 8d shows estimates of Belgian electoral heterogeneity during parliamentary elections between 1848 and 2016, as defined in equation (13). Figure 8e uses the electoral data of Kollman et al. (2019) to plot separatist vote shares in Belgian elections between 1848 and 2016. Figure 8f tracks regional and per capita GDP (in constant prices and at purchasing parity) of Wallonia (grey lines) and Flanders (black lines) between 1900 and 2016, using provincial data from Rosés and Wolf (2019). As data for Walloon Brabant and Flemish Brabant are not separately reported, but instead only given for the entire historical province of Brabant, these are excluded in the computations.

4.2 Model calibration

The central question is how well the model developed in section 2 can explain the observed evolutions in political heterogeneity, government size and secession proneness in Belgium, when calibrating its key parameters to match the main historical stylized facts discussed in the previous subsection. To this end, I calibrate a two-region dynamic model based on two exogenous 'laws of motion' that are directly taken from the data: the Belgian democratization process, which gradually extended voting rights to women and the poor; and income growth, which gradually increased living standards. The model is calibrated for 220 periods, where each period, t, equals a calendar year and $t = 1$ corresponds to 1800. The two regions correspond to Flanders and Wallonia respectively, where region 1 ('Flanders') has a 60% population share that is close to its actual population share, which remained quite stable over time. Crucially, the model is calibrated in such a way that the first region is initially poorer, yet experiences faster income growth, such that its gross regional product (GRP) surpasses that of region 2 after 145 periods. This is to account for the reversal of fortunes that was actually observed in Belgium, see figure 8f.[21]

To track the implications of these exogenous laws of motion on fiscal policy and regional instability in Belgium, I develop a simple calibration of the model with sufficient detail to capture the main stylized facts laid out in the previous section, yet sufficient parsimony to ensure traceable predictions that can be related to the underlying channels. First, I calibrate the marginal utility from public good consumption at $\theta = .35$, arbitrarily close to the value of .31 estimated for the year 2014 by Vanschoonbeek (2020). Total initial population size equals 2000, implying that the first region has a total population of $I_1 = 1200$, and $I_2 = 800$. I assume that half of the population in each region is male. This is relevant to track the political implications of Belgium's democratization process, which first extended suffrage to male voters before giving voting rights to women, see figure 8a.

I generate the initial income distribution at $t = 1$ by assuming that 10% of the agents in each region are rich, all the rich agents are male, rich men are 20 times richer than poor men, and women's incomes lie 50% below those of poor men. Note that these three agent types correspond to the various democratization waves in Belgium, where initially only rich agents were allowed to vote, the suffrage was later extended to poor male agents and eventually also female agents received the right to vote. Crucially, initial incomes for each of these three agent types lie three times higher in the second region, which consequently also attains a higher GRP in period $t = 1$. In practice, initial incomes for each agent type are drawn from a uniform distribution on $[0, \bar{y}]$ in such a way that average initial income in the whole country equals 100. Figure 9a plots the initial income distribution, separately for region 1 and region 2, where black dots identify male agents and grey dots identify female agents. It suffices to compare the incomes of rich agents in both regions to see that incomes for each agent type lie three times higher in the richest region 2.

[21] Note that I thus abstract, among many other things, from the economic ravages of both world wars.

With respect to the first exogenous law of motion, individual income growth is assumed to be common across agent types and is calibrated in such a way that the first region surpasses the second region's GRP at around $t = 145$, leading to a reversal of fortune. More specifically, I calibrate incomes in the second region to grow at an expected rate of 3% in each period, while per-period income growth in the first region averages 3.14%. To allow for sporadic negative income shocks, individual income growth is drawn from a uniform distribution on $[-.02, .05]$ for region 2 and $[-.02, .0514]$ for region 1. Figure 9b shows the resulting income distribution in period 200, when region 1 clearly outperforms region 2 in per capita GDP terms. Figure 9c shows the simulated per capita GDP trajectories for two particular rich agents in region 1 and region 2, showing that higher income growth allowed the first region's rich agent to attain a higher income around period $t = 200$. Finally, figure 9d shows the primary implication of the income growth differential, namely the reversal of fortune in which the first region surpassed the second region's GRP at around $t = 145$.

Note that the model described in section 2 does not explain how political heterogeneity emerges and evolves over time. Therefore, to generate agent-specific policy preferences that match the observed historical pattern of heterogeneity in Belgium, I capitalize on Inglehart's (1971) reinterpretation of Maslow's (1954) hierarchy of needs pyramid, with fundamental physiological needs such as food and physical safety at the bottom and the need for self-actualization at the top. Inglehart (1971) contends that in economically insecure times, policy preferences will be primarily restricted to narrow materialist values related to income and safety, while increasing prosperity widens the spectrum of policy preferences to also include postmaterialist values related to self-expression, gender equality and environmentalism. As a result, income growth should be associated with an extension of government tasks and increasing political disagreements, consistent with the observed increases in government size and political heterogeneity depicted in figures 8b and 8d. In spite of other potential explanations, this framework thus provides an elegant way to model the historical increase in preference heterogeneity in terms of the existing model.

To implement this, I draw agent's policy preferences from a distribution whose support is increasing in income. This implies that as income grows in the model, agent's increasingly take their material security for granted and instead place greater importance on non-material goals, mechanically increasing political heterogeneity, consistent with the empirical evidence in figure 8d. Simultaneously, de Smaele (2011) finds that there exist stable political disagreements between 'conservative' Flanders and 'progressive' Wallonia which go all the way back to the 19^{th} century. To also account for regional differences in policy preferences, I let the central tendency of the policy preference distribution differ between both regions. In the parlance of Montalvo and Reynal-Querol (2005), this implies that policy preferences are polarized *between* regions and fragmented *within* regions.

In practice, denoting the shape parameter determining the support of the policy preference distribution conditional on individual income, y_i, by δ, and the location parameter determining the regional preference distance between the optimal location of the public

goods bundle by μ, policy preferences of each agent, i, in each region, r, are drawn from

$$g_{i,r}^{*} = \min_{g_{i,r}^{*}} \left(\max_{g_{i,r}^{*}} \left(\frac{1+\mu\left(1.5-r\right)}{2} \pm \frac{\mathcal{U}\left(0,1\right)}{2} \left(\frac{y_{i,r}}{\bar{y}} \right)^{\delta}, 1 \right), 0 \right) \tag{14}$$

where $\mathcal{U}$ denotes the uniform distribution function, $\bar{y}$ captures the maximal income observation in our 220-period sample and the plus or minus sign, $\pm$, is randomly assigned as positive or negative.[22] It is easy to see that larger values of μ imply that regional preferences over the optimal location of the public goods bundle lie further apart, while smaller values of δ increase the income-induced variation of preferences around their central tendency of $\frac{1}{2} + \frac{\mu(1.5-r)}{2}$. In this particular application, I set $\mu = .6$ and $\delta = .2$.

In spite of the somewhat complicated functional form of the policy preference distribution function, its implications can be straightforwardly summarized in figures 9e and 9f. First of all, figure 9e shows that regional preferences over the optimal location of the public goods bundle differ between both regions, as inhabitants of region 1 generally prefer to locate g_c at higher interval values. Also visible is that the richest agents have more heterogeneous policy preferences in both regions, captured by the higher variability around the central tendencies of the distributions in both regions, as their affluence leads them to place more weight on non-material public policy goals. Figure 9f shows that, although the policy preference distribution remains qualitatively similar in period $t = 200$, the variation in policy preferences has dramatically increased as a result of income growth, which widened the relevant policy space to also include postmaterialist values.

Finally, this brings me to the second exogenous law of motion, which is the Belgian democratization process. Following figure 8a, collective decisions on the location of public policy, g_c^{*}, and income tax rates, t^{*}, are calibrated to be taken by majority voting by agents eligible to vote as follows: in periods 1 to 93, only rich agents can vote and get three votes per person; in periods 94 to 118, all men can vote, but rich male voters get two supplementary votes; in periods 119 to 149, all men can vote, receiving one vote per head; and from period 149 onwards, universal suffrage implies that women can vote too.

With respect to the timing in our model, I follow section 2. Thus, in each period t,

- The national median eligible voter decides the location of g_c in the union through equation (14).
- This determines the preference distance to public policy under the union, $d_{i,c} = g_{i,r}^{*} - g_c^{*}$.
- The national median eligible voter determines the income tax t_c in the union through equation (3).
- The regional median eligible voter decides the location of g_r under secession via equation (14).
- This determines the preference distance to public policy under separation, $d_{i,r} = g_{i,r}^{*} - g_r^{*}$.
- The regional median eligible voter determines the income tax t_r under secession via equation (15).
- The net gain of independence for each agent is computed through equation (16).
- The separatist vote share is computed as $\frac{\sum_{r=1}^{2}\sum_{i=1}^{I_r}\mathbf{1}_{\{\Delta_{i,r}>0\}}}{I_1+I_2} \geq 0$.

Note that the only difference with the general model outlined in section 2 is that I now also assume preference heterogeneity *within* regions, such that equilibrium taxes under separation and the net independence gain are respectively given by

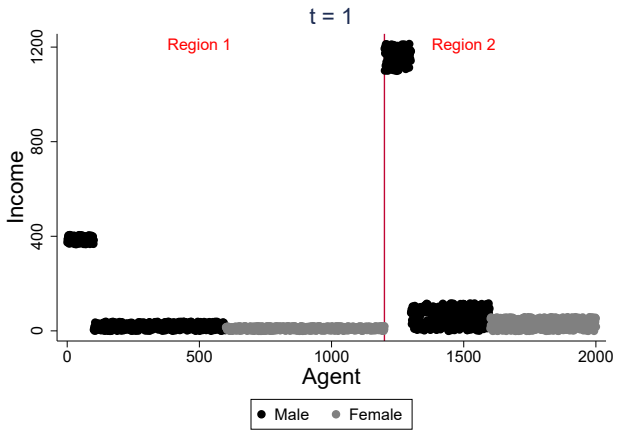
[22] The specific functional form was chosen to ensure that policy preferences generally lie between 0 and 1; policy preferences falling outside this interval are truncated whenever necessary.

$$t_r^* = t_{m,r} = \left(\frac{1 - d_{m,r}}{y_{m,r}} \right)^{\frac{1}{1-\theta}} Y_r^{\frac{\theta}{1-\theta}} \tag{15}$$

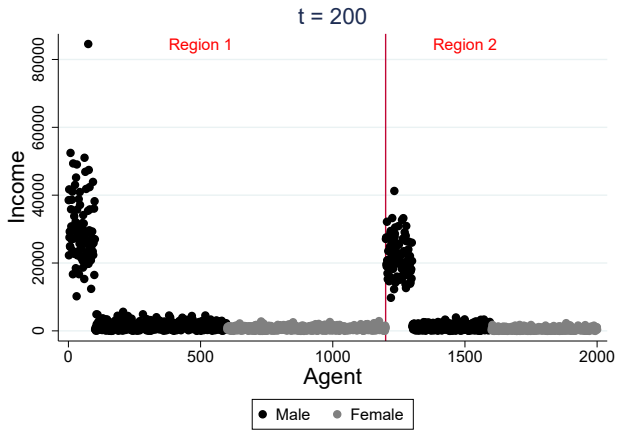
$$\Delta_{i,r} = (t_c^* - t_r^*)\, y_{i,r} + \frac{(1 - d_{i,r})\,(t_r^* Y_r)^\theta - (1 - d_{i,c})\,(t_c^* Y_c)^\theta}{\theta} \tag{16}$$

with m the median eligible voter in region r and $d_{i,r}$ the preference distance to public policy of agent i in region r under separation.

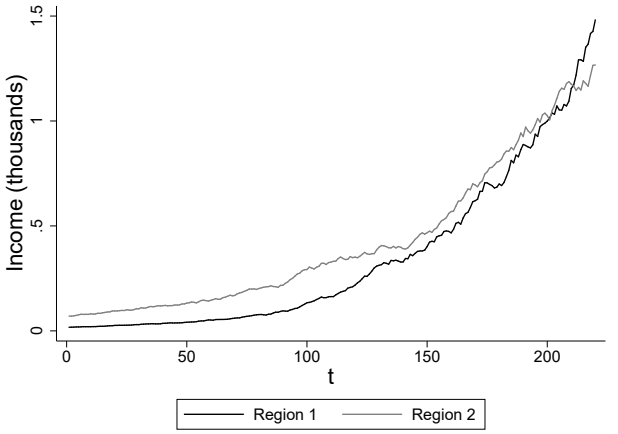
Figure 9: **Model calibration**



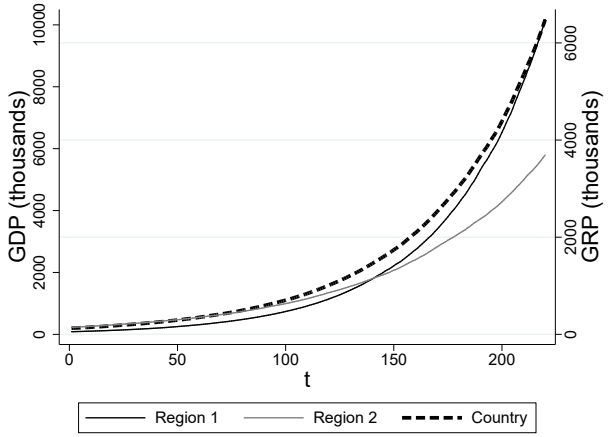
(a) Income distribution at $t = 1$



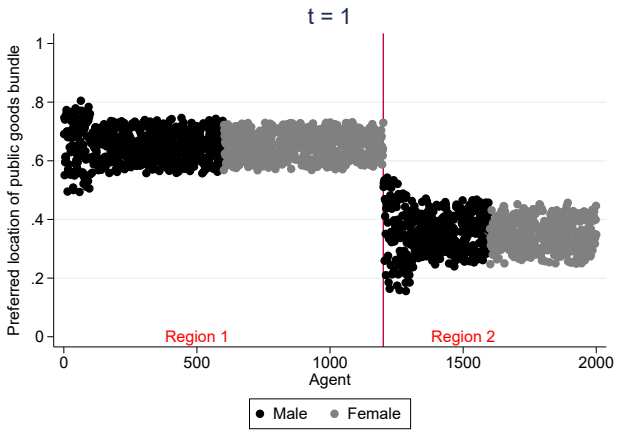
(b) Income distribution at $t = 200$



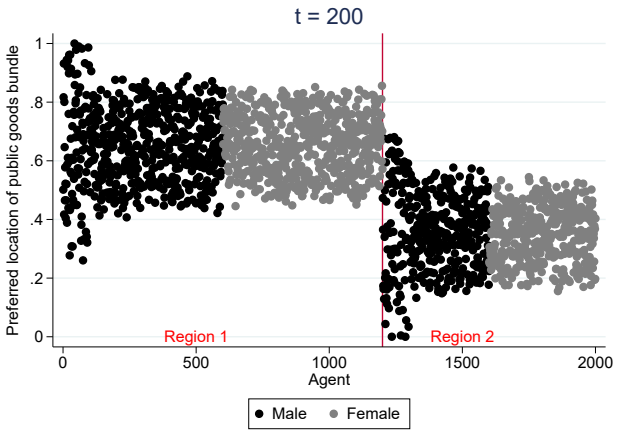
(c) Simulated per capita GDP trajectories for rich agents



(d) Simulated GRP trajectories



(e) Policy preference distribution at $t = 1$



(f) Policy preference distribution at $t = 200$

Note: Figure 9a shows the simulated initial income distribution, separately for region 1, $I \in [1, 1200]$, and region 2, $I \in [1201, 2000]$. Black dots represent male agents, grey dots represent female agents. Figure 9b shows the simulated income distribution in period $t = 200$. Figure 9c shows the simulated per capita income trajectories of a particular rich agent in each of both regions. Figure 9d plots the implied evolution in gross regional product (GRP), Y_r, for regions 1 and 2. The distribution of policy preferences over the optimal location of the public goods bundle, g^*, in periods $t = 1$ and $t = 200$ are respectively shown in figures 9e and 9f.

4.3 Results

I first present the results for the benchmark scenario detailed in the previous subsection, with special attention to the fiscal channels in the model. After establishing how well the model predicts the Belgian past, I also consider two extensions. First, I contrast the benchmark results with alternative predictions when setting political heterogeneity to zero in each period, to evaluate whether the model on balance detects centripetal or centrifugal effects of political heterogeneity in the Belgian case. In a second extension, I numerically resolve the model allowing for fiscal decentralization from period 170 onwards, to verify to what extent Belgian federalism is predicted to stabilize existing country borders.

4.3.1 Benchmark analysis

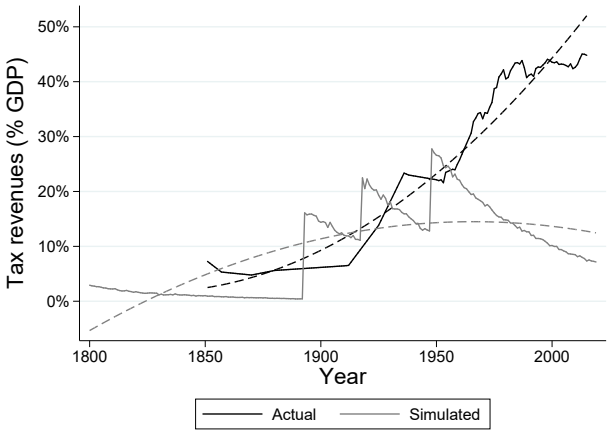
Figure 10 shows the first result of the benchmark analysis, demonstrating that the model is well able to explain the rise in government size over the previous double-century. More specifically, as predicted by Acemoglu and Robinson (2000), Belgium's drastic tax revenue increase between 1830 and 1945 primarily emerged as a byproduct of the Belgian democratization process, which increasingly gave the poor the opportunity to weigh in on fiscal policy choices, resulting in a gradual expansion of the welfare state. Once the expansionary effect of democracy peters out, however, income growth reduces tax preferences in our model and observed and simulated tax rates start to diverge.[23] Nevertheless, I later show how the continual expansion of government size post-1945 can be rationalized when the model is extended model to reflect the Belgian decentralization process, which further increased the willingness to pay taxes by bringing government closer to the people.

Subsequently, figure 11 shows that the model is also well able to match the historical evolution of political heterogeneity in Belgium.[24] According to the model, preference heterogeneity was initially rather limited and restricted to regional disagreements on public policy, as only a small part of the population was eligible to vote and economic precarity made them much more likely to primarily favor public policy narrowly aimed at increasing economic and physical security. Subsequently, income growth and democratization gradually extended the number of salient policy issues. Interestingly, conditioning the support

[23] It may be useful to point out that the mismatch between simulated and observed tax revenue shares largely stem from the simplifying assumption that the marginal utility from public consumption, θ, is time-constant. The existing literature nevertheless suggests that the taste for government likely went up over time, resulting from changing views on the fiscal role of government towards Keynesianism and mounting political pressures for increased social programs summarized in Wagner's law (Tanzi & Schuknecht, 2000). In this light, the discrepancy between both could be eliminated by endogenously determining the value of θ in each period, t, such that it rationalizes the observed data, $(d_{m,r}, y_{m,r}, Y_r, t_c^*)$, in accordance with equation (3). In terms of figure 10, this would amount to gradually increasing the 'taste for government' in the postwar period to rationalize the increased willingness to pay taxes. As the main purpose of the calibration exercise is to determine the extent to which two exogenous laws of motion can explain observed fiscal and political developments, I refrain from doing so in this exercise.

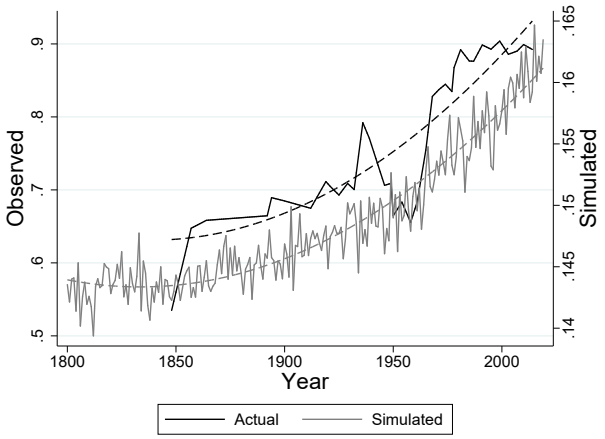
[24] Note that the observed indicator of electoral heterogeneity and the simulated distance between preferred and provided public goods bundles have no one-to-one mapping, as the former are indirectly derived from aggregate voting patterns, which is why both variables are plotted on different axes.

Figure 10: **Observed and predicted tax revenue shares**



of the policy preference distribution on income enables the model to also explain why new political parties continued to emerge and electoral heterogeneity continued to increase after the introduction of universal suffrage, because income growth brought about a cultural shift to postmaterialist values, which extended the breadth of politically relevant issues.

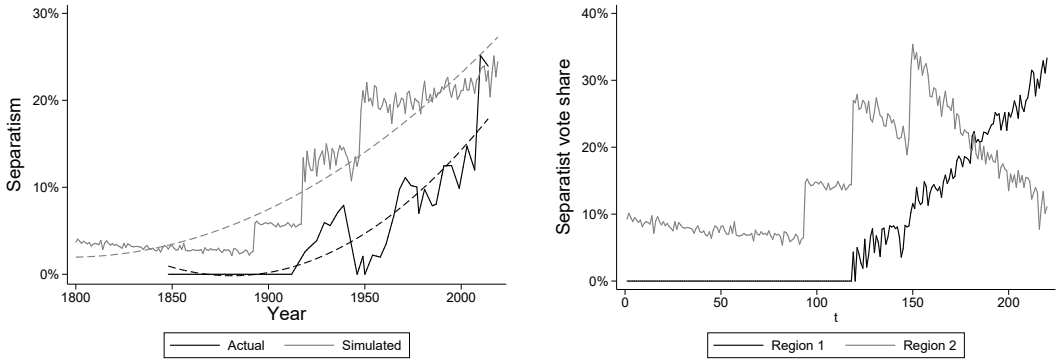
Figure 11: **Observed and predicted political distinctiveness**



Finally, figure 12a shows that the model also correctly anticipates a first peak of separatism around the interwar period followed by a reappearance and gradual increase in separatism throughout the postwar period. The primary reason is that, prior to universal suffrage, government size was comparatively small and the lack of voting power reduced the potential fiscal gains of secession for a majority of the inhabitants, while this situation reversed after the extension of suffrage from 1949 onwards. Interestingly, figure 12b also shows that the model mispredicts the separatist threat to first arise in Wallonia ('region 2'), *not* in Flanders. Nevertheless, the benchmark model does correctly anticipate the eventual drastic increase in separatism in Flanders up until its current peak today.

Figure 13 further investigates these issues by tracing them to the fiscal channels in the model. First of all, figure 13a gives a clear explanation for the continuous rise of

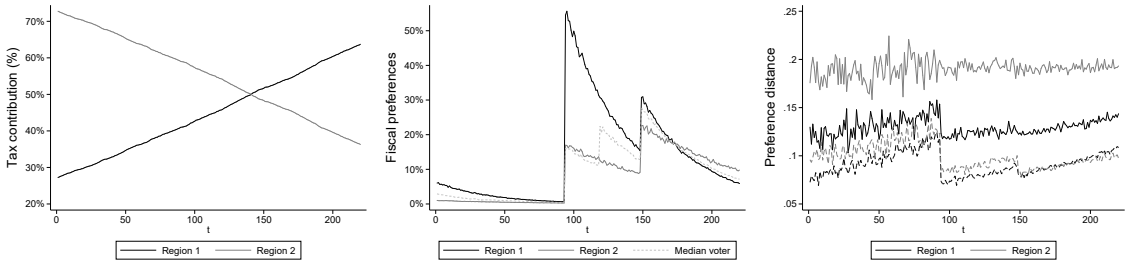
Figure 12: **Observed and predicted separatism**



(a) Belgium (b) Regional predictions

Flemish separatism ever since the early 20^{th} century, as its economic overperformance gradually increased its share in the central government budget, which went up from just short of 30% in period 1 to just over 60% in period 220. Going from a net receiver to a net contributor to the central government budget naturally increased the attractiveness of the outside option, especially in a context of increasing preference distances to public policy. Wallonia experienced the opposite evolution and saw the relative fiscal costs of maintaining the union gradually decrease, which in turn gradually increased the welfare gains it derives from the union in the form of the higher tax base it can provide.

Figure 13: **Channels**



(a) Tax contribution (%) (b) Fiscal preferences (c) Preference distance

Figure 13b complements this with a closer look at the intensity of redistributive conflict between both regions, by plotting the tax rates preferred by the national and regional median eligible voters. Interestingly, the model predicts fiscal preferences to almost coincide throughout the 19^{th} century, when census voting ensured that eligible voters shared a preference for low taxes. This implies that the potential fiscal gains of secession were close to zero for both regions. Between 1893 and 1918, the introduction of plural male voting for the first time magnified the discrepancy between majority preferred and actually implemented income taxes, by extending the suffrage yet giving the wealthy disproportional electoral weight. During this period, Flemish voters were especially ill-served under the existing fiscal policy, as the poor Flemish voters were the staunchest supporters of extend-

ing the welfare state, but were overruled by the combined electoral weight of rich agents and the more wealthy Walloon electorate. Nevertheless, its small tax base prevented the simultaneous emergence of a Flemish separatist threat, as testified in figure 12b.

The brief surge in separatism in the interwar years is then traced to the introduction of equal male voting in 1918, which led to the most pronounced discrepancy between regionally and nationally preferred tax rates. As the voting system gradually shifted power to the more numerous poor Flemish voters, the extension of the suffrage primarily disadvantaged then-richer Wallonia, as it allowed the Flemish median voter to increase equilibrium tax rates against its will. As Wallonia still was the country's richest region, its loss of fiscal control coincides with a simulated interwar rise in separatism, partially due to the tax base advantage the region could still leverage when declaring independence. Finally, the introduction of universal suffrage pacified this interregional redistributive conflict by giving more electoral weight to agents with similar income levels and hence tax preferences, namely women. Nevertheless, the growing income differential between the north and the south led to a situation where Wallonia became a net demander of extending the welfare state while Flanders increasingly wanted the exact opposite. In terms of the model, the relative erosion of the Walloon tax base thus simultaneously leads to increasing interregional disagreements over optimal government size and a growing separatist demand in the region with a persistently growing tax base advantage, namely Flanders.

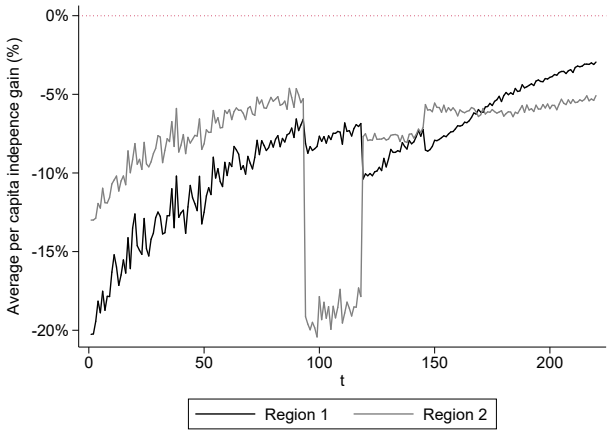
The reason for this latter result is also related to figure 13c, which shows the average preference distance to public policy for both regions under the union (full lines) and under secession (dashed lines). The figure first offers a partial explanation for the initial rise in Walloon separatism in figure 12b, which occurred after the introduction of census voting in 1893. Although figure 13b shows that income taxes are very similar both under the union and under secession, figure 13c nevertheless shows that, as the extension of the suffrage considerably increased popular control over the location of the public goods bundle, it also substantially reduced the average preference distance to public policy under secession. Combined with the general increase in government size induced by the extension of the suffrage, which increased the relative importance of maintaining control over the location of public policy, this fueled separatism in Belgium's richest Walloon region. Interestingly, the extension of suffrage mainly increased the demand for independence in Wallonia's richest strata, which not only saw Walloon secession as an opportunity to reduce the gap between goverment and the people, but also as a way to reduce income taxes by simultaneously reducing economies of scale (and thus preferred taxes) as well as the number of poor voters.

The figure also clearly shows that Flanders experienced the smallest preference distance to public policy throughout the period under consideration, largely thanks to its population advantage. Also apparent is that an independent Flanders would be better capable to tailor public policy to Flemish preferences, especially after the introduction of universal suffrage extended policy influence to the entire population. The postwar rise in Flemish separatism in the model thus not only has to do with the increasing fiscal disagreements with Wallonia

shown in figure 13b, but also with the welfare gain of a government closer to the people under secession shown in figure 13c. As this latter benefit becomes especially attractive when the fiscal costs of secession decline, as has been the case in the Flemish region due to the persistent rise of its relative economic size, it is easy to see why separatism primarily grew in postwar Flanders at a time of growing economic dominance and full democracy.

To conclude, figure 14 sheds some light on the evolving intensity of the separatist threat in Belgium, by plotting the average per capita utility gain of independence in both regions. Interestingly, the payoff of separatism is lowest in both regions throughout the 19^{th} century, when census voting ensured that eligible voters shared a preference for low taxes and small governments both under union and secession and scale economies stacked the economic cards in favor of existing country borders. With very similar tax rates under union and secession, the population advantage of the union became comparatively more attractive, as it allowed the central government to offer a larger public goods bundle for a given preferred income tax. Thus, as long as public consumption remained trivial, the cost-efficiency of the public goods bundle mattered more than control over its location.[25]

Figure 14: **The net gain of independence in Belgium**



The figure also clearly shows how the Belgian democratization process affected the independence gain in both regions. Initially, the extension of the suffrage decreased independence gains by giving more inhabitants influence over public and fiscal policy. Thus, as the Walloon population benefited most from the introduction of plural male voting, their net independence gain fell substantially in the run-up to World War I. Nevertheless, the long run effect of democracy has clearly been to increase independence gains, and especially in Flanders, which saw its net independence gain rise consistently throughout the entire postwar period. The primary reason is that democracy first increased government size, by increasing the electoral weight of the poor, and for the same reason elevated the importance of maintaining control of public policy, which has obvious centrifugal effects in

[25]Note that the assumption of preferred tax rates decreasing in income and preference heterogeneity, see equation (3), mechanically reduces independence costs over time, by reducing the size of government and, hence, the importance of maintaining control over its location.

a polarized country. This channel is thus capable of explaining the timing of separatism in Belgium and, especially, why it was largely absent in the 19^{th} century in spite of economic and political polarization: government was so small that control over public policy was largely irrelevant. When the democratization process sufficiently increased the relative importance of public over private consumption, the model moreover predicts the location of separatism to occur within the largest economic region, which only has to give up a comparatively small share of its internal market and tax base to declare independence. This channel thus explains why separatism, when it occurred, primarily attracted support in Belgium's richest region of Flanders. Finally, consistent with the electoral results of figure 8e, the model also suggests Belgium to never have been as secession prone as today.

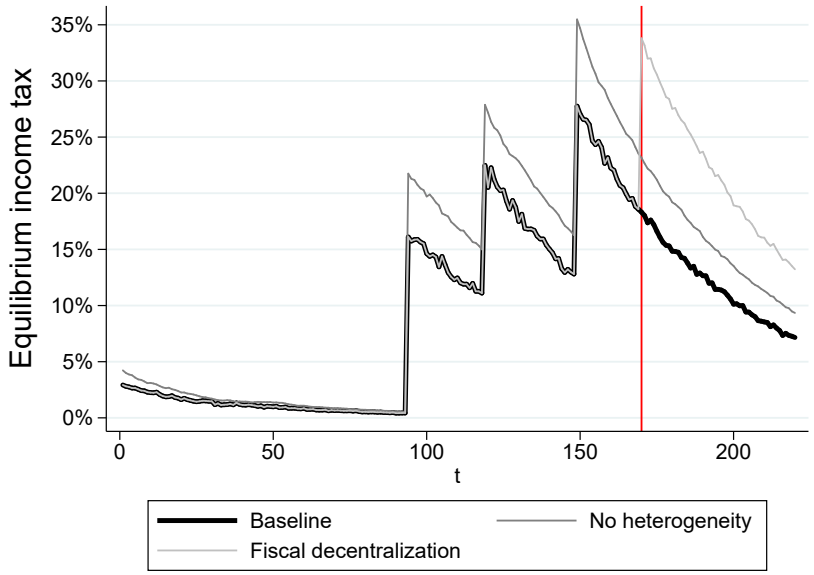
4.3.2 Extensions

I contrast these benchmark results with two sets of alternative predictions, corresponding to a first scenario where political heterogeneity is absent and policy preferences are perfectly congruent for all agents, and a second scenario where fiscal decentralization becomes possible from period 170 onwards and is modeled as in subsection 2.4. While a graphical summary of the primary results is relegated to figure A4, I briefly discuss the most relevant results here. Figures containing regional estimates plot the results for Flanders ('region 1') in black, while results for Wallonia ('region 2') are then plotted in grey.

First of all, figure 15 confirms that the willingness to pay income taxes effectively increases in both alternative scenarios. More specifically, the results imply that equilibrium income tax rates on average would lie 30% higher if policy preferences were fully congruent in Belgium, while they would lie 80% higher when fiscal decentralization bridged part of the gap between government and the people. Moreover, figure A4 shows that under decentralization, roughly 30% of the tax revenues would be decentralized, which is quite close but not equal to its observed level of around 20% as reported by the OECD (2019).

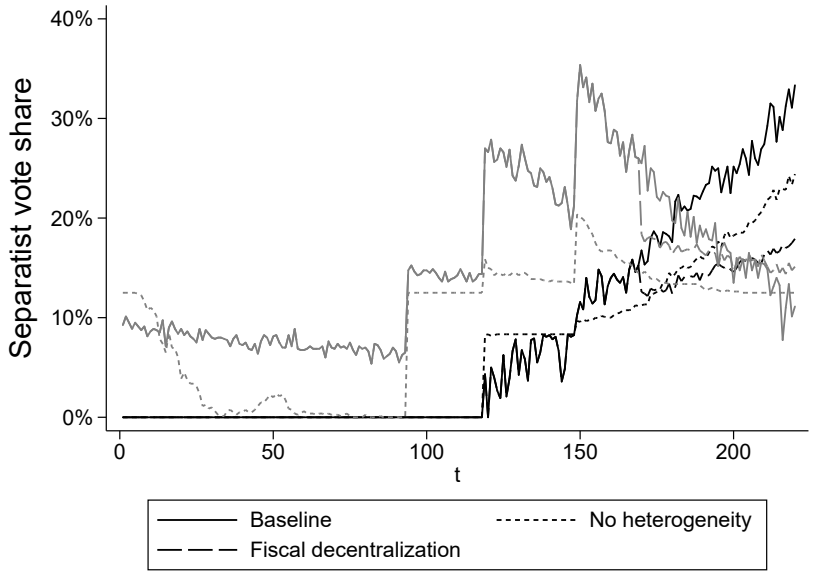
Figure 16 subsequently shows that separatism would generally become *less* prevalent in a scenario absent political heterogeneity. Apart from the brief interbellum period, when the existence of the gap between preferred and provided public policy actually reduces Flemish support for the welfare state and hence the incentives to secede to be able to implement it, secession is always less frequent in the scenario without political heterogeneity. According to this metric, the model thus suggests that political heterogeneity on average has had centrifugal effects in Belgium. Also visible on the figure is that fiscal decentralization has a strong centripetal effect, slashing popular support for independence roughly by a third the moment that it is introduced in period 170. The centripetal effect of fiscal federalism is only durable in Flanders, however, as its disproportional tax contribution ensures an ever larger amount of regional expenditures for any given degree of fiscal decentralization. For similar reasons, fiscal federalism even becomes disadvantageous for slower-growing Wallonia, which would prefer to preserve access to the common tax pool for government

Figure 15: **Tax revenue shares under different scenarios**



expenditures, even if it implies accepting a larger preference distance to public policy.

Figure 16: **Separatism under different scenarios**

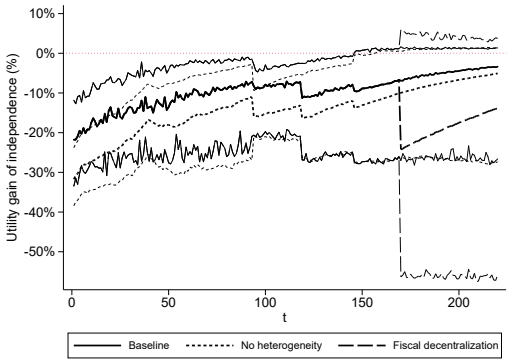


Finally, figure 17 sheds more light on the heterogeneity in separatist tendencies across income groups, by analyzing the evolution in individual secession payoffs for the bottom-90% and top-10% earners under different scenarios. Figures 17a and 17b confirm that the intensity of the separatist threat unambiguously decreases in a scenario without political heterogeneity, underscoring the centrifugal effect of political heterogeneity for the bottom-90% earners in both regions. Moreover, fiscal decentralization also unambiguously lowers independence gains for the majority of the population in both Flanders and Wallonia,

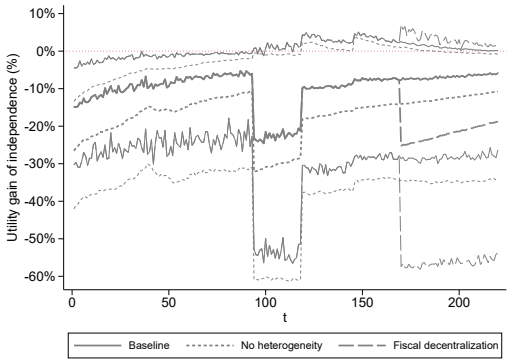
by allowing them to benefit some of the advantages of a government closer to the people even in the federal union. Nevertheless, the tax-increasing effect of fiscal decentralization does overcompensate the welfare gain of a government closer to the people for the highest earners even in the bottom-90%, as testified by the maximum independence gain growing more positive after the introduction of decentralization in $t = 170$.

Figure 17: **Distribution of independence gains under different scenarios**

Bottom-90% earners

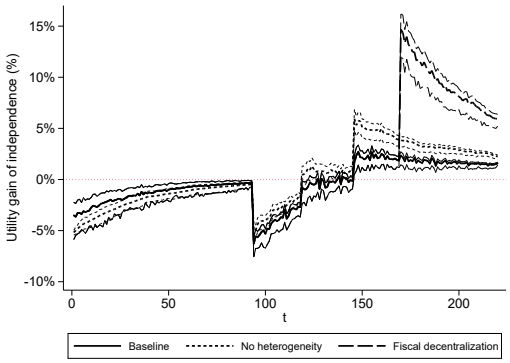


(a) Flanders ('region 1')

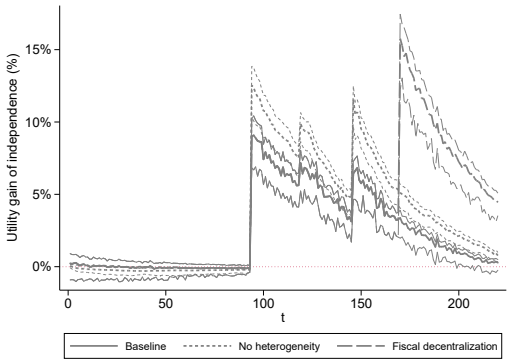


(b) Wallonia ('region 2')

Top-10% earners



(c) Flanders ('region 1')



(d) Wallonia ('region 2')

Note: This figure plots the simulated evolution in the utility gain of secession for the bottom-90% and top-10% earners in Flanders ('region 1') and Wallonia ('region 2') between $t = 1$ and $t = 220$, in three scenarios: the benchmark scenario (full lines), the 'no heterogeneity' scenario (short dashed lines) and the 'fiscal decentralization' scenario (long dashed lines). The thin lines track maximum and minimum independence gains within each group.

In sharp contrast, the top-10% of earners turn out to experience more positive independence gains in a scenario without political heterogeneity, as their lack of electoral power would then force them to accept higher equilibrium income tax rates without the tempering effect of preference heterogeneity, see figures 17c and 17d. According to this metric, political heterogeneity *did* have centripetal effects at the top of the income distribution, by lowering the independence gains of the wealthy, and *did* reduce separatist tendencies

in the most secession-prone strata of the Belgian population. By the same token, fiscal decentralization increased the willingness to pay taxes and drastically increased independence gains for the wealthy, which are disproportionally located in Flanders, contributing to country border *in*stability on the intensive margin for this particular group.

4.3.3 Summary

Although the model inevitably abstracts from a host of factors that affected Belgium's secession proneness, including Flemish disillusionment that their loyalty during World War I was not rewarded by meeting their linguistic demands or the introduction of constitutional minority protections, it nevertheless turns out to be well capable of matching the timing, if not the location, of all the main fiscal and political developments in Belgium since its inception.[26] It explains the absence of 19^{th} century separatism through the system of census voting, which restricted the electorate to a small number of wealthy voters favoring small government in both regions. Limited government presence in everyday life, both under separation and union, restricted the potential welfare gains of secession and stabilized country borders by elevating the concern of cost efficiency to the most pressing one, cementing the union through the superior economies of scale that it could offer.

Separatism could only emerge after the extension of the suffrage in 1893 and 1918 led to an increase in government size. The shifting electoral power from rich to poor voters resulted in an extension of the welfare state and an increase in government tasks, raising the relative importance of maintaining control over the location of public policy. Thus, even if the preference distance *between* both regions remained stable over time, support for separatism surged for a first time in the interwar period, when the redistributive preferences of the poor first became politically salient and regionally differing pressures to extend the welfare state began to fuel redistributive conflicts between the poorer and richer regions. After the introduction of universal suffrage, separatism grew in tandem with government size and the Flemish tax base, as increasing government size increased the welfare gain of closing the gap between government and the Flemish people by declaring independence while the growing Flemish tax base reduced the fiscal costs of doing so. As the discontinuous jumps in secession proneness correspond more closely to the discontinuous fiscal pressures induced by the democratization process, rather than the steady increase in preference heterogeneity, according to the model, separatism in Belgium thus mainly seems to be driven by fiscal pressures and less so by preference heterogeneity in itself.

The findings also imply that while political heterogeneity slightly increased independence gains for the majority of the population, by increasing the gap between government and the people, it also lowered the independence gain for the wealthy, by lowering equilib-

[26] According to De Wever, Verdoodt, and Vrints (2011), flamingant disillusionment with having contributed 70% of the first World War casualties yet being denied more linguistic autonomy after the war, epitomized in the slogan *"hier ons bloed, wanneer ons recht"* (*"here our blood, when our right"*), explains the brief flirt between Flemish nationalism and anti-democratic fascism during the interwar years.

rium tax rates, reducing the secession payoff for the most secession-prone Flemish part of the Belgian population. To the extent that the wealthy have a disproportionate influence on the political process, political heterogeneity may have actually increased country border stability in Belgium by bringing government size more in line with the fiscal preferences of its most secession-prone inhabitants. Finally, I also find that the Belgian fiscal decentralization process can plausibly explain why the observed willingness to pay taxes continued to increase post-1945, as it reduced the welfare costs of public policy under the union. I find suggestive evidence that federalism had strong centripetal effects, and mainly so for the most secession-prone region, as the ability to autonomously spend a fixed share of the tax contribution reduced the common pool problem by enabling the rich region to at least partially enjoy the benefits of its higher tax base within the union. Nevertheless, as fiscal decentralization also increased the willingness to pay taxes in the union, it also increased the independence payoff for the wealthy, who are disproportionally located in Flanders.

5 Conclusion

This papers develops a fiscal bargaining model for divided countries to analyze how fiscal policy mediates the relation between preference heterogeneity and country border stability. Its central finding is that preference heterogeneity has a tax-reducing effect and may therefore act as a centrifugal or centripetal force, depending on whether the beneficial effects of reducing incongruent government policies overcompensates for the cost of having a government further from the people in the most discontented regions. It also shows that stable and efficient country borders may require a minimal degree of preference heterogeneity to voluntarily keep discontented regions in or, lacking that, credible commitments to a tax ceiling, compensating transfers or, in some cases, fiscal decentralization. I provide empirical evidence for the model's predictions in a large panel of countries over the past half century. I also show its capacity to match some of the main fiscal and political evolutions in Belgium, namely the emergence of the welfare state, the gradual increase in political heterogeneity, the timing of separatism and its persistent postwar rise in Flanders.

One intriguing implication of these findings is that the current episode of political polarization in the West may well be instrumental in preserving efficient country borders after decades of divergent regional development and democratic extensions of the welfare state. That is, some discontented regions may efficiently utilize distinctive voting patterns to complicate policy formulation in the union, with the aim of keeping taxes low enough for them to voluntarily remain in it.[27] One plausible alternative explanation, however, is that discontented regions oversupport regionalist movements to improve their bargaining position and extract rents from the central government, for instance in the form of compensating transfers or favorable tax treatments. An extension of the model that re-

[27] Note that discontented regions by a similar chain of reasoning also benefit from fueling heterogeneity in the rest of the country, for instance by trying to pit other regions against each other.

laxes the implicit assumption of full information á la Fearon and Van Houten (2002), to account for the potential of regional strategic electoral behavior, could therefore paint a more complete picture of the welfare effects of fiscal bargaining under separatist threats.

In addition, the framework as it stands could also be further employed to assess its capacity to explain global patterns of state formation, through the two exogenous forces in our numerical application, namely the unprecedented growth of democracy and increasing economic security. According to the model, governments first had to grow in size before becoming salient enough in everyday life to incite separatist tendencies. The global democratization process may therefore have been an essential precondition for the centrifugal effects of the 20^{th} century increase in preference heterogeneity, as the urge to maintain influence over public policy should be proportional to its presence in everyday life. It would therefore be interesting to relate these findings to those of Gancia et al. (2017), where state formation is exclusively driven by the exogenous increase in globalization and trade. More specifically, incorporating the effects of globalization in the model discussed here could shed some light on the relative importance of domestic and international channels as well as their interaction in shaping existing country borders over the past two centuries.

References

Acemoglu, D., & Robinson, J. A. (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *The Quarterly Journal of Economics*, *115*(4), 1167–1199.

Adserà, A., & Boix, C. (2002). Trade, Democracy, and the Size of the Public Sector: The Political Underpinnings of Openness. *International Organization*, *56*(2), 229–262.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics*, *114*(4), 1243–1284.

Alesina, A., Baqir, R., & Hoxby, C. (2004). Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy*, *112*(2), 348–396.

Alesina, A., & Spolaore, E. (1997). On the Number and Size of Nations. *The Quarterly Journal of Economics*, *112*(4), 1027–1056. doi: https://doi.org/10.1162/003355300555411

Alesina, A., & Spolaore, E. (2003). *The Size of Nations*. Cambridge: MIT Press.

Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, Social Divisions, and Public Goods in Rural India. *Journal of the European Economic Association*, *3*(2-3), 639–647.

Banerjee, A., Iyer, L., & Somanathan, R. (2007). Chapter 49 Public Action for Public Goods. In P. T. Schultz & J. Strauss (Eds.), *Handbook of development economics* (Vol. 4, pp. 3117–3154). Oxford.

Bolton, P., & Roland, G. (1997). The Breakup of Nations: a Political Economy Analysis. *The Quarterly Journal of Economics*, *112*(4), 1057–1090. doi: https://doi.org/10.1162/003355300555420

Buchanan, J., & Faith, R. (1987). Secession and the Limits of Taxation: Toward a Theory of Internal Exit. *American Economic Review*, *77*(5), 1023–1031.

Buyst, E. (2010). Reversal of Fortune in a Small, Open Economy: Regional GDP in Belgium, 1896–2000. *Rivista di Storia Economica*, *26*, 75–92.

Caramani, D. (2004). *The Societies of Europe*. New York: Palgrave Macmillan.

Claeys, P., & Martire, F. (2015). 'If You Want Me To Stay, Pay': a Model of Asymmetric Federalism in Centralised Countries. *Environment and Planning C: Government and Policy*, *33*(2), 305–320.

Dalton, R. J. (1996). Political Cleavages, Issues, and Electoral Change. In L. LeDuc, R. G. Niemi, & P. Norris (Eds.), *Comparing democracies: Elections and voting in global perspective* (pp. 319–342). Thousand Oaks, CA: Sage Publications.

De Wever, B., Verdoodt, F.-J., & Vrints, A. (2011). De Vlaamse Patriotten en de Natievorming. Hoe de Vlaamse Natie Ophield 'Klein' te Zijn. *Wetenschappelijke Tijdingen*, *4*, 217–248.

de Smaele, H. (2011). How 'real' is Right-Wing Flanders. In B. De Wever (Ed.), *Re-bel e-book 12: Right-wing flanders, left-wing wallonia? is this so? if so, why? and is it*

a problem? (pp. 6–10). Brussels: Re-Bel Initiative.

Desmet, K., Le Breton, M., Ortuño-Ortín, I., & Weber, S. (2011). The Stability and Breakup of Nations: A Quantitative Analysis. *Journal of Economic Growth*, *16*(3), 183–213.

Desmet, K., Ortuño-ortín, I., & Wacziarg, R. (2017). Culture, Ethnicity, and Diversity. *American Economic Review*, *107*(9), 2479–2513.

Desmet, K., Ortuño-ortín, I., & Weber, S. (2017). Peripheral Diversity: Transfers Versus Public Goods. *Social Choice and Welfare*, 1–37.

Etro, F. (2006). Political Geography. *Public Choice*, *127*(3-4), 321–343.

Fearon, J., & Van Houten, P. (2002). The Politicization of Cultural and Economic Difference: a Return to the Theory of Regional Autonomy Movements. Paper Presented at the Fifth Meeting of the Laboratory in Comparative Ethnic Processes.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, *105*(10), 3150–3182. doi: https://doi.org/10.1257/aer.20130954

Flora, P., Kraus, F., & Pfenning, W. (1983). *State, Economy, and Society in Western Europe 1815-1975*. Frankfurt: Campus Verlag.

Gancia, G. A., Ponzetto, G. A. M., & Ventura, J. (2017). Globalization and Political Structure. *CEPR Discussion Paper No. 111592*, 1–51.

Gehring, K., & Schneider, S. A. (2020). Regional Resources and Democratic Secessionism. *Journal of Public Economics*, *181*, 1–18.

Gennaioli, N., & Tabellini, G. (2018). Identity , Beliefs, and Political Conflict. *CEPR Discussion Paper 13390*, 1–64.

Gerard, M. (2014). Economic Aspects of Constitutional Change: the Case of Belgium. *Oxford Review of Economic Policy*, *30*(2), 257–276.

Griffiths, R. D. (2014). Secession and the Invisible Hand of the International System. *Review of International Studies*, *40*(3), 559–581.

Griffiths, R. D. (2015). Between Dissolution and Blood: How Administrative Lines and Categories Shape Secessionist Outcomes. *International Organization*, *69*(3), 731–751. doi: https://doi.org/10.1017/s0020818315000077

Griffiths, R. D. (2016). *Age of Secession: the International and Domestic Determinants of State Birth*. Cambridge: Cambridge University Press.

Haimanko, O., Le Breton, M., & Weber, S. (2005). Transfers in a Polarized Country: Bridging the Gap Between Efficiency and Stability. *Journal of Public Economics*, *89*(7), 1277–1303.

IMF. (2019). *The IMF Fiscal Decentralization Dataset.* Retrieved 2019-07-24, from https://data.imf.org/?sk=1C28EBFB-62B3-4B0C-AED3-048EEEBB684F

Inglehart, R. (1971). The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies. *The American Political Science Review*, *65*(4), 991–1017.

Inglehart, R. (1981). Post-Materialism in an Environment of Insecurity. *The American Political Science Review*, *75*(4), 880–900.

Inglehart, R., & Norris, P. (2016). Trump, Brexit and the Rise of Populism: Economic Have-Nots and Cultural Backlash. In *American political science association annual meeting* (pp. 1–52). Philadelphia.

Kollman, K., Caramani, D., Backer, D., & Lublin, D. (2019). *Constituency-level Elections Archive [data file and codebook]*. Retrieved from http://www.electiondataarchive.org

La Porta, R., Lopez-de Silanes, F., Shleifer, A., & Vishny, R. (1999). Quality of Government. *The Journal of Law, Economics and Organization*, *15*(1), 222–279.

Lake, D. A., & O'Mahony, A. (2004). The Incredible Shrinking State: Explaining Change in the Territorial Size of Countries. *Journal of Conflict Resolution*, *48*(5), 699–722.

Le Breton, M., & Weber, S. (2003). The Art of Making Everybody Happy : How to Prevent a Secession. *IMF Staff Papers*, *50*(3), 403–435.

Maslow, A. H. (1954). *Motivation and Personality*. New York: Harper & Brothers.

Miguel, E., & Gugerty, M. K. (2005). Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics*, *89*(11-12), 2325–2368.

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic Polarization, Potential Conflict, and Civil Wars. *The American Economic Review*, *95*(3), 796–816.

Nieuw-Vlaamse Alliantie. (2015). *Nieuw-Vlaamse Alliantie - Statuten.* Retrieved 2016-06-01, from http://www.n-va.be/sites/default/files/documents/PDF/statuten20151012.pdf

OECD. (2019). *Revenue Statistics.* Retrieved 2019-07-26, from https://stats.oecd.org/Index.aspx?DataSetCode=REV

Rodríguez-pose, A. (2017). *When Rich Places Want to Secede.* Retrieved 2019-07-29, from https://www.theatlantic.com/business/archive/2017/10/catalonia-secede-rich-region/544244/

Rodríguez-pose, A. (2018). The Revenge of the Places That Don't Matter (and What To Do About It). *Cambridge Journal of Regions, Economy and Society*, *11*, 189–209.

Rodrik, D. (2017). Populism and the Economics of Globalization. *NBER Working Paper 23559*, 1–43.

Rosés, J. R., & Wolf, N. (2019). *The Economic Development of Europe's Regions: A Quantitative History since 1900* (J. R. Rosés & N. Wolf, Eds.). Oxon: Routledge.

Spolaore, E. (2009). Efficiency and Stability of National Borders. *Review of Business and Economic Literature*, *54*(1), 4–22.

Spolaore, E. (2010). Federalism, Regional Redistribution, and Country Stability. In N. Bosch, M. Espasa, & A. Solé Ollé (Eds.), *The political economy of inter-regional fiscal flows: Measurement, determinants and effects on country stability.* (pp. 329–356). Massachusetts: Edward Elgar Publishing.

Spolaore, E. (2016). The Economics of Political Borders. In E. Kontorovich & F. Parisi

(Eds.), *Economic analysis of international law* (pp. 1–304). Cheltenham: Edward Elgar Publishing.

Staal, K. (2010). Incentives for Separation and Incentives for Public Good. *Public Choice*, *145*, 531–546.

Suesse, M. (2018). Adjusting the Size of Nations: Empirical Determinants of Separatism and the Soviet Breakup. *Journal of Comparative Economics*, *47*(1), 1–15.

Tanzi, V., & Schuknecht, L. (2000). *Public Spending in the 20th Century A Global Perspective*. Cambridge: Cambridge University Press.

The Economist. (2016). *Rage Against the Dying of the Light.* Retrieved 2019-07-29, from https://www.economist.com/finance-and-economics/2016/12/15/place-based-economic-policies-as-a-response-to-populism?fsrc=scn/tw/te/bl/ed/

The Maddison Project. (2017). *The Maddison Project Database 2013.* Retrieved from http://www.ggdc.net/maddison/maddison-project/data.htm.

Van Zanden, J. L., Baten, J., Foldvari, P., & Van Leeuwen, B. (2014). The Changing Shape of Global Inequality 1820-2000; Exploring a New Dataset. *Review of Income and Wealth*, *60*(2), 279–297. Retrieved from https://clio-infra.eu/data/IncomeInequality{_}Broad.xlsx

Vanhanen, T. (2000). *Polyarchy Dataset.* Retrieved 2019-03-01, from https://www.prio.org/Global/upload/CSCW/Data/Governance/file42535{_}polyarchy{_}v2.dta

Vanschoonbeek, J. (2020). Regional (In)Stability in Europe: a Quantitative Model of State Fragmentation. *Journal of Comparative Economics*, 1–53.

World Bank. (2019). *World Development Indicators.* Retrieved 2017-01-10, from http://data.worldbank.org/data-catalog/world-development-indicators

A Proofs

A.1 Section 2.1

Substituting for t_c^* in equation (6), I get that

$$\begin{aligned}\Delta_{i,R} &= U_{i,r}^{R\notin c} + \left(\frac{1-d_P}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R - \frac{(1-d_R)\left(\left(\frac{1-d_P}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} Y_c\right)^{\theta}}{\theta} \\ &= U_{i,r}^{R\notin c} + \left(\frac{1-d_P}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R - \frac{(1-d_R)\left(\frac{1-d_P}{y_P}\right)^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}}{\theta}\end{aligned} \tag{1A}$$

Partially deriving this expression with respect to d_P yields

$$\begin{aligned}\frac{\partial \Delta_{i,R}}{\partial d_P} &= \frac{1}{1-\theta}\left(\frac{1-d_P}{y_P}\right)^{\frac{\theta}{1-\theta}}\left(-\frac{1}{y_P}\right) Y_c^{\frac{\theta}{1-\theta}} y_R - \frac{\frac{\theta}{1-\theta}(1-d_R)\left(\frac{1-d_P}{y_P}\right)^{\frac{2\theta-1}{1-\theta}}\left(-\frac{1}{y_P}\right) Y_c^{\frac{\theta}{1-\theta}}}{\theta} \\ &= -\left(\frac{1-d_P}{y_P}\right)^{\frac{\theta}{1-\theta}} \frac{Y_c^{\frac{\theta}{1-\theta}}}{(1-\theta)}\frac{y_R}{y_P} + \frac{(1-d_R)(1-d_P)^{\frac{2\theta-1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}}{(1-\theta)\, y_P^{\frac{\theta}{1-\theta}}} \\ &= \frac{(1-d_R)(1-d_P)^{\frac{2\theta-1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_P - (1-d_P)^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R}{(1-\theta)\, y_P^{\frac{1}{1-\theta}}}\end{aligned} \tag{2A}$$

which reaches a critical point at

$$\begin{aligned}(1-d_R)(1-d_P)^{\frac{2\theta-1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_P &= (1-d_P)^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R \\ \frac{(1-d_R)}{(1-d_P)} y_P &= y_R \\ (1-d_P) &= (1-d_R)\frac{y_P}{y_R} \\ d_P^* &= 1-(1-d_R)\frac{y_P}{y_R}\end{aligned} \tag{3A}$$

Subsequently, the second order derivative can be computed as

$$\begin{aligned}\frac{\partial \Delta_{i,R}^2}{\partial^2 d_P} &= \frac{\frac{2\theta-1}{1-\theta}(1-d_R)(1-d_P)^{\frac{3\theta-2}{1-\theta}}(-1) Y_c^{\frac{\theta}{1-\theta}} y_P - \frac{\theta}{1-\theta}(1-d_P)^{\frac{2\theta-1}{1-\theta}}(-1) Y_c^{\frac{\theta}{1-\theta}} y_R}{(1-\theta)\, y_P^{\frac{1}{1-\theta}}} \\ &= \frac{\theta(1-d_P)^{\frac{2\theta-1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R - (2\theta-1)(1-d_R)(1-d_P)^{\frac{3\theta-2}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_P}{(1-\theta)^2\, y_P^{\frac{1}{1-\theta}}}\end{aligned} \tag{4A}$$

which is positive as long as

$$\begin{aligned}
\theta\left(1-d_P\right)^{\frac{2\theta-1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R &> \left(2\theta-1\right)\left(1-d_R\right)\left(1-d_P\right)^{\frac{3\theta-2}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_P \\
\theta\left(1-d_P\right) y_R &> \left(2\theta-1\right)\left(1-d_R\right) y_P \\
\frac{y_R}{y_P}\frac{\left(1-d_P\right)}{\left(1-d_R\right)} &> 2-\frac{1}{\theta}
\end{aligned} \tag{5A}$$

As the left hand side of the last inequality is strictly positive by assumption, it is trivial to see that the second order derivative is always positive for $\theta < \frac{1}{2}$. In this case, $\Delta_{i,R}$ is a convex function of d_P such that increasing political heterogeneity initially decreases the welfare gain of secession until d_P reaches d_P^*, after which further increases in political heterogeneity increase the welfare gain of secession for the rich region.

More formally, when $\theta < \frac{y_P(1-d_R)}{2y_P(1-d_P)-y_R(1-d_P)}$, there exists a non-empty set of political heterogeneity levels $d_P \in [0, d_P^*]$ where increases in political heterogeneity lower the secession pay-off for the rich region as described in equation (6). Whether intensifying heterogeneity in the rest of the country can actually avoid separatism in the rich region crucially depends on the utility value of the outside option. That is, intensifying political heterogeneity in the rest of the country can only stabilize country borders if it manages to turn a positive welfare gain of secession in the rich region negative: in such a scenario, the rich region would want to secede if political heterogeneity in the rest of the country is too low and causes too large a discrepancy between regionally preferred and nationally levied income taxes, but would refrain from doing so once the rest of the country becomes sufficiently heterogeneous to bring equilibrium national tax rates down to agreeably low levels. Formally, political heterogeneity can stabilize country borders only if there exist a set of parameter values for which $\Delta_R > 0$ for $d_P = 0$ and $\Delta_R \leq 0$ for $d_P = d_P^*$.

To prove the existence of such scenario's, note that the former condition hold whenever

$$\begin{aligned}
U_{i,r}^{R\notin c} + \left(\frac{1}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R - \frac{\left(1-d_R\right)\left(\frac{1}{y_P}\right)^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}}{\theta} &> 0 \\
U_{i,r}^{R\notin c} &> \frac{\left(1-d_R\right) Y_c^{\frac{\theta}{1-\theta}} - Y_c^{\frac{\theta}{1-\theta}} \frac{y_R}{y_P}}{\theta y_P^{\frac{\theta}{1-\theta}}} \\
U_{i,r}^{R\notin c} &> \frac{\left(1-d_R\right) Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(1-\frac{\theta y_R}{\left(1-d_R\right) y_P}\right)
\end{aligned} \tag{6A}$$

while the latter condition is satisfied whenever

$$U_{i,r}^{R\notin c} + \left(\frac{1-d_P^*}{y_P}\right)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} y_R - \frac{\left(1-d_R\right)\left(\frac{1-d_P^*}{y_P}\right)^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}}{\theta} < 0$$

$$U_{i,r}^{R\notin c}+\left(\frac{1-\left[1-(1-d_R)\,\frac{y_P}{y_R}\right]}{y_P}\right)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}y_R-\frac{(1-d_R)\left(\frac{1-\left[1-(1-d_R)\frac{y_P}{y_R}\right]}{y_P}\right)^{\frac{\theta}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{\theta}<0$$

$$U_{i,r}^{R\notin c}+\left(\frac{1-d_R}{y_R}\right)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}y_R-\frac{(1-d_R)\left(\frac{1-d_R}{y_R}\right)^{\frac{\theta}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{\theta}<0 \tag{7A}$$

$$U_{i,r}^{R\notin c}<\frac{(1-d_R)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}-\theta\,(1-d_R)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{\theta y_R^{\frac{\theta}{1-\theta}}}$$

$$U_{i,r}^{R\notin c}<\frac{(1-d_R)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{\theta y_R^{\frac{\theta}{1-\theta}}}(1-\theta)$$

implying both conditions are simultaneously met for the set of parameter values satisfying

$$\frac{(1-d_R)\,Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(1-\frac{\theta y_R}{(1-d_R)\,y_P}\right)<\frac{(1-d_R)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{\theta y_R^{\frac{\theta}{1-\theta}}}(1-\theta)$$
$$\left(\frac{y_R}{(1-d_R)\,y_P}\right)^{\frac{\theta}{1-\theta}}\left(1-\frac{\theta y_R}{(1-d_R)\,y_P}\right)+\theta<1 \tag{8A}$$

Noting that $y_R > y_P$ and $\theta \leq \frac{1}{2}$ by assumption, it is trivial to see that the set of parameter values satisfying this latter inequality is effectively non-empty.

Finally, I can also specify the range of values for d_P, $[d^*, d^{**}]$, for which the existing country borders are stable. To do so, first note that the lower threshold is reached when the representative agent of the rich region becomes indifferent between preserving the union and the secession of its region, as their control over equilibrium national tax rates ensures that poor regions always benefit from the union for low values of d_P (eg. see figure 1). Consequently, d^* can be computed as the lowest value of d_P for which $\Delta_R = 0$, or

$$\Delta_R=(t_c^*-t_R^*)\,y_R+\frac{(t_R^*Y_R)^\theta-(1-d_R)\,(t_c^*Y_c)^\theta}{\theta}=0$$

$$\left(\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}-\frac{Y_R^{\frac{\theta}{1-\theta}}}{y_R^{\frac{1}{1-\theta}}}\right)y_R+\frac{\left(\frac{Y_R}{y_R}\right)^{\frac{\theta}{1-\theta}}-(1-d_R)\left(\frac{[1-d_P]Y_c}{y_P}\right)^{\frac{\theta}{1-\theta}}}{\theta}=0$$

$$\frac{(1-d_P)^{\frac{1}{1-\theta}}\,y_R}{y_P^{\frac{1}{1-\theta}}}-\frac{(1-d_R)\,(1-d_P)^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}+\frac{\alpha_R^{\frac{\theta}{1-\theta}}}{\theta y_R^{\frac{\theta}{1-\theta}}}(1-\theta)=0$$

$$\frac{(1-d_R)\,(1-d_P)^{\frac{\theta}{1-\theta}}-\theta\,(1-d_P)^{\frac{\theta}{1-\theta}}\,(1-d_P)\,\frac{y_R}{y_P}}{\theta y_P^{\frac{\theta}{1-\theta}}}=\frac{1}{\theta}\left(\frac{\alpha_R}{y_R}\right)^{\frac{\theta}{1-\theta}}(1-\theta)$$

$$(1-d_P)^{\frac{\theta}{1-\theta}}\left(\frac{[1-d_R]\,y_P-\theta\,(1-d_P)\,y_R}{y_P}\right)=\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}}(1-\theta) \tag{9A}$$

$$(1-d_P)^{\frac{\theta}{1-\theta}}=\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}}\frac{(1-\theta)\,y_P}{(1-d_R)\,y_P-\theta\,(1-d_P)\,y_R}$$

$$1-d_P=\frac{\alpha_R y_P}{y_R}\left(\frac{[1-\theta]\,y_P}{[1-d_R]\,y_P-\theta\,(1-d_P)\,y_R}\right)^{\frac{1-\theta}{\theta}}$$

$$d_P=1-\alpha_R\frac{y_P}{y_R}\left(\frac{[1-\theta]\,y_P}{[1-d_R]\,y_P-\theta\,(1-d_P)\,y_R}\right)^{\frac{1-\theta}{\theta}}$$

$$d^*\approx 1-\left([1-d_R]-\left[\alpha_R\frac{y_P}{y_R}\right]^{\frac{\theta}{1-\theta}}[1-\theta]\right)\frac{y_P}{\theta y_R}$$

where $\alpha_R=\frac{\sum_{i=1}^{I} y_R}{Y_c}<1$ captures the relative economic size of the rich region in the union and, hence, the economies of scale it can retain after secession. The last approximation is obtained by putting the left hand side to zero and is reasonably accurate for low threshold values of $d^*\approx 0$. One takeaway from equation (9A) is that the minimal level of political heterogeneity required to keep the rich region in unambiguously increases in the relative economic size of the rich region, α_R. In other words, if the economy of scale cost of secession for the rich region declines, its outside option becomes more attractive, lowering the maximal tax burden it is willing to accept within the existing union.

Following a similar chain of logic, I can also compute the maximal amount of political heterogeneity in the poor regions that satisfies a no-secession constraint by computing the critical value of d_P for which even poor regions would be indifferent between secession and union, see figure 1. Formally, this would be the case whenever $\Delta_P=0$, or

$$\begin{aligned}
(t_c^*-t_P^*)\,y_P+\frac{(t_P^*Y_P)^\theta-(1-d_P)\,(t_c^*Y_c)^\theta}{\theta}&=0\\
\left(\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}-\left[\frac{1}{y_P}\right]^{\frac{1}{1-\theta}}Y_P^{\frac{\theta}{1-\theta}}\right)y_P+&\\
\frac{\left(\left[\frac{1}{y_P}\right]^{\frac{1}{1-\theta}}Y_P^{\frac{\theta}{1-\theta}}Y_P\right)^\theta-(1-d_P)\left(\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}Y_c\right)^\theta}{\theta}&=0\\
\left([1-d_P]^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}-[\alpha_P Y_c]^{\frac{\theta}{1-\theta}}\right)\frac{1}{y_P^{\frac{\theta}{1-\theta}}}+\frac{(\alpha_P Y_c)^{\frac{\theta}{1-\theta}}-(1-d_P)^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}\theta}&=0\\
(1-d_P)^{\frac{1}{1-\theta}}-\alpha_P^{\frac{\theta}{1-\theta}}+\frac{\alpha_P^{\frac{\theta}{1-\theta}}-(1-d_P)^{\frac{1}{1-\theta}}}{\theta}&=0\\
(1-d_P)^{\frac{1}{1-\theta}}\left(1-\frac{1}{\theta}\right)=\alpha_P^{\frac{\theta}{1-\theta}}\left(1-\frac{1}{\theta}\right)&\\
1-d_P=\alpha_P^\theta&\\
d^{**}=1-\alpha_P^\theta&
\end{aligned}\tag{10A}$$

where $\alpha_P=\frac{\sum_{i=1}^{I} y_P}{Y_c}<1$ captures the relative economic size of a poor region in the union and, hence, the economies of scale it can retain after the rich region' secession. The upper threshold of political heterogeneity above which state breakup becomes unavoidable, d^{**}, thus also decreases if poor regions are better able to retain more economies of scale, α_P, after breakup and derive less utility from the consumption of public goods, θ.

Finally, using these heuristics to compute the threshold values for d_P for the numerical example in section 2.1 yields that $\hat{d}^*\approx 1-\left([1-.25]-\left[.182\frac{4000}{8000}\right]^{\frac{.45}{1-.45}}[1-.45]\right)\frac{4000}{.45\times 8000}=$

.253 and $\hat{y}^{**} = 1 - .091^{.45} = .66$, which are indeed arbitrarily close to their observed counterparts, which are equal to [.27, .67] and also plotted in figure 1.

A.2 Section 2.2

Denote the tax ceiling implicitly defined by $\{t_c \in [0,1] : \Delta_R(t_c) = 0\}$ by $\bar{t_c^*}$. Note that agents in poor regions will only respect this ceiling if doing so yields higher utility than ignoring the no-secession constraint to independently set national income tax rates. This implies that country borders are stable for low values of $d_P \leq d^*$ only when

$$\Delta_P^{\bar{t_c^*}} = \underbrace{(1-\bar{t_c^*})\, y_P + \frac{(1-d_P)\,(\bar{t_c^*} Y_c)^\theta}{\theta}}_{\text{Utility under tax ceiling, } U^{\bar{t_c^*}}} - \underbrace{(1-t_c^*)\, y_P + \frac{(1-d_P)\,(t_c^* \bar{\alpha}_P Y_c)^\theta}{\theta}}_{\text{Utility ignoring no-secession constraint, } U^{R \notin C}} \geq 0 \quad \text{(11A)}$$

where $\bar{\alpha}_P = \frac{\sum_{r \neq R}^{R} \sum_{i=1}^{I} y_P}{Y_c}$ tracks GDP in the mother country after the rich region secedes.

The net gain of accepting this tax ceiling remains positive as long as $\Delta_P^{\bar{t_c^*}} \geq 0$, or

$$\begin{aligned}
&\left(\left[\frac{1-d_P}{y_P} \right]^{\frac{1}{1-\theta}} [\bar{\alpha}_P Y_c]^{\frac{\theta}{1-\theta}} - \left[\frac{1-d^*}{y_P} \right]^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} \right) y_P \\
&+ \frac{(1-d_P)\left(\left[\frac{1-d^*}{y_P} \right]^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} - \left[\frac{1-d_P}{y_P} \right]^{\frac{\theta}{1-\theta}} [\bar{\alpha}_P Y_c]^{\frac{\theta}{1-\theta}} \right)}{\theta} \geq 0 \\
&\frac{(1-d_P)^{\frac{1}{1-\theta}} (\bar{\alpha}_P Y_c)^{\frac{\theta}{1-\theta}} - (1-d^*)^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}} \\
&+ \frac{(1-d_P)\left([1-d^*]^{\frac{\theta}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}} - [1-d_P]^{\frac{\theta}{1-\theta}} [\bar{\alpha}_P Y_c]^{\frac{\theta}{1-\theta}} \right)}{\theta y_P^{\frac{\theta}{1-\theta}}} \geq 0
\end{aligned} \quad \text{(12A)}$$

$$\theta (1-d_P)^{\frac{1}{1-\theta}} \bar{\alpha}_P^{\frac{\theta}{1-\theta}} - \theta (1-d^*)^{\frac{1}{1-\theta}} + (1-d_P)(1-d^*)^{\frac{\theta}{1-\theta}} - (1-d_P)^{\frac{1}{1-\theta}} \bar{\alpha}_P^{\frac{\theta}{1-\theta}} \geq 0$$

$$(1-d_P)^{\frac{1}{1-\theta}} \bar{\alpha}_P^{\frac{\theta}{1-\theta}} (\theta - 1) + (1-d^*)^{\frac{1}{1-\theta}} \left(\frac{1-d_P}{1-d^*} - \theta \right) \geq 0$$

This expression is clearly positive at $d_P = d^*$, as $(1-\theta)(1-d^*)^{\frac{1}{1-\theta}} \left(1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}} \right) > 0$ by assumption, confirming that poor agents will always prefer to keep the rich region in if they can freely determine income tax rates. To verify whether the tax ceiling remains welfare-optimal if it becomes binding, note that we can approximate the effect of a small

decrease in d_P by partially deriving this expression with respect to $(1-d_P)$, yielding

$$\begin{aligned}\frac{\partial \Delta_P^{\bar{t}_c^*}}{\partial (1-d_P)} &= \frac{1}{1-\theta}(1-d_P)^{\frac{\theta}{1-\theta}}\,\bar{\alpha}_P^{\frac{\theta}{1-\theta}}(\theta-1)+(1-d^*)^{\frac{\theta}{1-\theta}} \\ &= (1-d^*)^{\frac{\theta}{1-\theta}}-(1-d_P)^{\frac{\theta}{1-\theta}}\,\bar{\alpha}_P^{\frac{\theta}{1-\theta}}\end{aligned} \tag{13A}$$

Note that $(1-d^*)^{\frac{\theta}{1-\theta}}\left(1-\bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right)>0$ by assumption, such that this first order derivative is unambiguously positive at $d_P=d^*$. The net welfare gain of holding on to the tax ceiling thus grows larger when political heterogeneity, d_P, falls below the lower threshold value, d^* and the tax ceiling becomes binding. This ensures that poor agents can credibly commit to the tax ceiling whenever this is necessary to keep the country together.

A.3 Section 2.3

Equation (9A) of appendix A.1 computes the rich region's net gain of declaring independence when $d_P \leq d^*$ and, consequently, the magnitude of the fiscal transfer it would require from the center to accept the existing country borders. Based this equation, the required fiscal transfer, T_R, for a heterogeneity level of $d_P \leq d^*$ can be slightly rewritten as

$$T_R \geq \Delta_R = \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(\left[\alpha_R\frac{y_P}{y_R}\right]^{\frac{\theta}{1-\theta}}[1-\theta]+[1-d_P]^{\frac{1}{1-\theta}}\left[\theta\frac{y_R}{y_P}-\frac{1-d_R}{1-d_P}\right]\right) \tag{14A}$$

The willingness to pay for transfers in the poor regions equals the amount that renders them indifferent between the rich region's secession and the preservation of the union. Assuming they agree to each pay proportional shares, their willingness to pay, T_P, equals

$$\begin{aligned}T_P \leq \underbrace{\frac{T_R}{R-1}}_{\text{fiscal transfer}} &= \underbrace{(1-t_c^*)\,y_P+\frac{(1-d_P)\,(t_c^*Y_c)^\theta}{\theta}}_{\text{Utility under transfer scheme, } U^T}-\underbrace{(1-t_{R\notin c}^*)\,y_P-\frac{(1-d_P)\,(t_{R\notin c}^*\bar{\alpha}_P Y_c)^\theta}{\theta}}_{\text{Utility under secession, } U^{R\notin c}} \\ &= \left(\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}}[\bar{\alpha}_P Y_c]^{\frac{\theta}{1-\theta}}-\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}}Y_c^{\frac{\theta}{1-\theta}}\right)y_P \\ &+\frac{(1-d_P)}{\theta}\left(\frac{[1-d_P]^{\frac{\theta}{1-\theta}}\,Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}-\frac{[1-d_P]^{\frac{\theta}{1-\theta}}\,[\bar{\alpha}_P Y_c]^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}\right) \\ &= \frac{(1-d_P)^{\frac{1}{1-\theta}}\,Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}(\bar{\alpha}_P-1)+\frac{(1-d_P)^{\frac{1}{1-\theta}}\,Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(1-\bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right) \\ &= \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(\theta\,[1-d_P]^{\frac{1}{1-\theta}}\,[\bar{\alpha}_P-1]+[1-d_P]^{\frac{1}{1-\theta}}\left[1-\bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right]\right) \\ &= \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}[1-d_P]^{\frac{1}{1-\theta}}\left[1-\bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right][1-\theta]\end{aligned} \tag{15A}$$

Equation (9A) of appendix A.1 also reveals that the required fiscal transfer to keep the rich region in, T_R, will be exactly equal to zero at $d_P = d^*$. Nevertheless, at $d_P = d^*$, there is a positive willingness to pay fiscal transfers in the poor regions, as by assumption $T_P = \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}} [1-d^*]^{\frac{1}{1-\theta}} \left[1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right] [1-\theta] \geq 0$. This is another confirmation that if the union is welfare-optimal at low levels of political heterogeneity, $d_P < d^{**}$, secession would induce an aggregate welfare loss.

To verify whether there is sufficient willingness to pay in poor regions to fully stabilize efficient country borders, let $\frac{y_R}{y_P} = \beta \geq 1$ denote the degree of inter-regional income inequality and note that there is positive willingness to pay as long as $T_P \geq T_R$, or

$$\frac{Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}} \Bigg(\underbrace{[R-1]\,[1-d_P]^{\frac{1}{1-\theta}} \left[1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right] [1-\theta]}_{\text{Poor region's WTP, } T_P} - \underbrace{\left[\frac{\alpha_R}{\beta}\right]^{\frac{\theta}{1-\theta}} [1-\theta] - [1-d_P]^{\frac{1}{1-\theta}} \theta\beta + [1-d_P]^{\frac{\theta}{1-\theta}} [1-d_R]}_{\text{Rich region's compensation, } T_R} \Bigg) \geq 0 \tag{16A}$$

In order to get a tractable solution for d_P, I approximate $(1-d_P)^{\frac{\theta}{1-\theta}}$ by 1, which is reasonably accurate for low values of $d_P \approx 0$ and $\theta \leq \frac{1}{2}$. Solving for the threshold value of d_P satisfying the inequality in equation (16A), I find that

$$\begin{aligned}
(R-1)\,(1-d_P)^{\frac{1}{1-\theta}} \left(1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right)(1-\theta) - (1-\theta)\left(\frac{\alpha_R}{\beta}\right)^{\frac{\theta}{1-\theta}} - (1-d_P)^{\frac{1}{1-\theta}}\theta\beta + 1 - d_R &\geq 0 \\
(1-d_P)^{\frac{1}{1-\theta}} \left[(1-\theta)\left(1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right)(R-1) - \theta\beta\right] &\geq \frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\,(1-d_R)}{\beta^{\frac{\theta}{1-\theta}}} \\
(1-d_P)^{\frac{1}{1-\theta}} &\geq \frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\,(1-d_R)}{\beta^{\frac{\theta}{1-\theta}}\left[(1-\theta)\left(1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right)(R-1) - \theta\beta\right]} \\
d_P &\leq 1 - \left(\frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\,(1-d_R)}{\beta^{\frac{\theta}{1-\theta}}\left[(1-\theta)\left(1 - \bar{\alpha}_P^{\frac{\theta}{1-\theta}}\right)(R-1) - \theta\beta\right]}\right)^{1-\theta}
\end{aligned} \tag{17A}$$

So as long as d_P lies below the threshold value identified in the last inequality, poor region's willingness to pay exceeds the required transfer to keep the rich region in. I already mentioned this to be the case for $d_P = d^*$; this last inequality moreover guarantees that this will also hold for all possible values of $d_P < d^*$.

To verify when poor region's prefer a fiscal transfer scheme over a tax ceiling, note that individual utility under both scenario's is respectively given by

$$U_{i,P} = \begin{cases} U_{i,P}^{T} = (1-t_c^*)\,y_P + \dfrac{(1-d_P)\,(t_c^* Y_c)^{\theta}}{\theta} - \left(\dfrac{T_R}{R-1}\right) & \text{if } i \notin R \;\&\; T_R \geq 0 \\ U_{i,P}^{\bar{t_c^*}} = (1-\bar{t_c^*})\,y_P + \dfrac{(1-d_P)\,(\bar{t_c^*} Y_c)^{\theta}}{\theta} & \text{if } i \notin R \;\&\; t_c^* = \bar{t_c^*} \end{cases} \tag{18A}$$

It is straightforward to see that poor regions will prefer a fiscal transfer scheme over a tax ceiling whenever $U_{i,P}^{T} \geq U_{i,P}^{t_c^*}$, or

$$\begin{aligned}
&\left(1-\left[\frac{1-d_P}{y_P}\right]^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}\right) y_P + \frac{(1-d_P)^{\frac{1}{1-\theta}}\, Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}} \\
&- \frac{Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}\left(\left[\frac{\alpha_R}{\beta}\right]^{\frac{\theta}{1-\theta}}[1-\theta] + [1-d_P]^{\frac{1}{1-\theta}}\left[\theta\beta - \frac{1-d_R}{1-d_P}\right]\right) \\
&- \left(1-\left[\frac{1-d^*}{y_P}\right]^{\frac{1}{1-\theta}} Y_c^{\frac{\theta}{1-\theta}}\right) y_P - \frac{(1-d_P)\,(1-d^*)^{\frac{\theta}{1-\theta}}\, Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}} \geq 0 \\
\Leftrightarrow & y_P + \frac{Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}\left(-[1-d_P]^{\frac{1}{1-\theta}}\,\theta + [1-d_P]^{\frac{1}{1-\theta}}\right) \\
&- \frac{Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}\left(\left[\frac{\alpha_R}{\beta}\right]^{\frac{\theta}{1-\theta}}[1-\theta] + [1-d_P]^{\frac{1}{1-\theta}}\left[\theta\beta - \frac{1-d_R}{1-d_P}\right]\right) \\
&- y_P + \frac{Y_c^{\frac{\theta}{1-\theta}}}{y_P^{\frac{\theta}{1-\theta}}}\left(-[1-d^*]^{\frac{1}{1-\theta}}\,\theta + [1-d_P]\,[1-d^*]^{\frac{\theta}{1-\theta}}\right) \geq 0 \\
\Leftrightarrow & (1-d_P)^{\frac{1}{1-\theta}}\left(-\theta + 1 - \theta\beta - \left[\frac{1-d^*}{1-d_P}\right]^{\frac{\theta}{1-\theta}}\right) \geq \\
&\left(\frac{\alpha_R}{\beta}\right)^{\frac{\theta}{1-\theta}}(1-\theta) - (1-d_P)^{\frac{\theta}{1-\theta}}\,(1-d_R) - \theta\,(1-d^*)^{\frac{1}{1-\theta}}
\end{aligned} \tag{19A}$$

Once again letting $(1-d_P)^{\frac{\theta}{1-\theta}} \approx 1$ allows me to rewrite this as

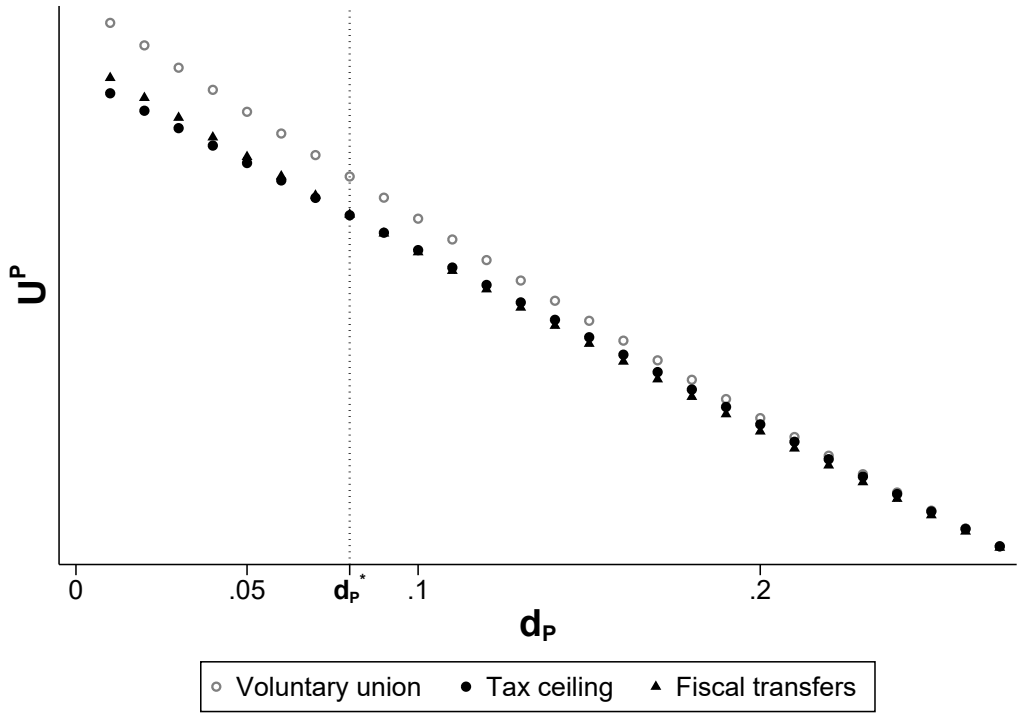
$$\begin{aligned}
(1-d_P)^{\frac{1}{1-\theta}}\left(1-[\beta+1]\,\theta - [1-d^*]^{\frac{\theta}{1-\theta}}\right) &\geq \frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\left[1-d_R+\theta\,(1-d^*)^{\frac{1}{1-\theta}}\right]}{\beta^{\frac{\theta}{1-\theta}}} \\
(1-d_P)^{\frac{1}{1-\theta}} &\geq \frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\left[1-d_R+\theta\,(1-d^*)^{\frac{1}{1-\theta}}\right]}{\beta^{\frac{\theta}{1-\theta}}\left(1-[\beta+1]\,\theta - [1-d^*]^{\frac{\theta}{1-\theta}}\right)} \\
d_P &\leq 1 - \left[\frac{(1-\theta)\,\alpha_R^{\frac{\theta}{1-\theta}} - \beta^{\frac{\theta}{1-\theta}}\left[1-d_R+\theta\,(1-d^*)^{\frac{1}{1-\theta}}\right]}{\beta^{\frac{\theta}{1-\theta}}\left(1-[\beta+1]\,\theta - [1-d^*]^{\frac{\theta}{1-\theta}}\right)}\right]^{1-\theta}
\end{aligned} \tag{20A}$$

As can be seen, poor agents will prefer a fiscal transfer scheme over a tax ceiling whenever political heterogeneity is below this threshold value. One interpretation of this result is that more homogeneous regions also have a stronger preference to remain in control of the size of the government, even if this requires considerable transfers to keep discontented regions in, because political consensus more easily allows the central government to locate the public goods bundle such that all poor agents derive maximal utility from public consumption. If political differences also divide the poor regions, on the other hand, they are more likely to prefer the hard constraint on fiscal policy to prevent fiscal spending that comes at an extra costs of increasing the necessary stabilizing transfers.

For completeness sake, note that this threshold amounts to 0.08 in the numerical

example, close to its approximated value of $1 - \left[\frac{(1-.45).182^{\frac{.45}{1-.45}} - 2^{\frac{.45}{1-.45}}\left[1-.25+.45(1-.27)^{\frac{1}{1-.45}}\right]}{2^{\frac{.45}{1-.45}}\left(1-[2+1].45-[1-.27]^{\frac{.45}{1-.45}}\right)}\right]^{1-.45} = .1$. This implies that poor regions will prefer to commit to a tax ceiling to keep the rich region in if $d_P \in [.08, .27]$, but will rather set up a fiscal transfer scheme whenever political heterogeneity reduces to $d_P \in [0, .08]$. The figure below provides a graphical summary of this finding by plotting how utility in the poor regions evolves when their political heterogeneity drops to sufficiently low values to require either a tax ceiling or fiscal transfers to stabilize country borders. It is immediately obvious that poor agents initially derive higher utility under a tax ceiling, but that they reach superior utility with fiscal transfers whenever political heterogeneity approaches zero. The difference between both scenario's is, however, comparatively small such that both scenario's guarantee a similar level of individual utility for poor as well as rich agents.

Figure A1: **Poor region's utility under a tax ceiling and a fiscal transfer scheme**



Note: This figure plots how individual utility in the poor regions evolves as poor regions become increasingly politically homogeneous, d_P. The hollow dots, solid dots and triangles show individual utility under the voluntary preservation of the union (U^c), a tax ceiling ($U^{\bar{t}^*_c}$) and a fiscal transfer scheme (U^T) as defined in equations (1), (11A) and (18A). d_P^* captures the critical level of political heterogeneity in the poor regions to make them indifferent between a tax ceiling and a fiscal transfer scheme.

A.4 Section 2.4

Assume that fiscal decentralization allows each region to spend a share D, $0 \leq D \leq 1$, of its tax contribution on the provision of a regional public goods bundle and that there are no losses in economies of scope or externalities. Moreover, assume that utility derived

from regional and national public consumption is additively separable. This implies that individual utility under fiscal decentralization of inhabitant i in region r will amount to

$$\begin{aligned} U_{i,r}^{D} &= \left(1-t_c^D\right) y_{i,r} + \frac{(1-d_r)\left(\left[1-D\right] t_c^D Y_c\right)^{\theta} + \left(D t_c^D \alpha_r Y_c\right)^{\theta}}{\theta} \\ &= \left(1-t_c^D\right) y_{i,r} + \frac{\left(t_c^D Y_c\right)^{\theta}\left(\left[1-d_r\right]\left[1-D\right]^{\theta} + \alpha_r^{\theta} D^{\theta}\right)}{\theta} \end{aligned} \tag{21A}$$

with α_r denoting the regional share in country GDP as before.

If taxes are decided by majority voting, conditional on the degree of fiscal decentralization D, the poor region's median tax voter prefers the national income tax equal to

$$t_c^D = \frac{Y_c^{\frac{\theta}{1-\theta}}\left(\left[1-d_P\right]\left[1-D\right]^{\theta} + \alpha_P^{\theta} D^{\theta}\right)^{\frac{1}{1-\theta}}}{y_P^{\frac{1}{1-\theta}}} \tag{22A}$$

Assuming that the degree of decentralization is also chosen by majority voting, note that substituting for equilibrium taxes in equation (21A) allows me to rewrite individual utility of the poor region's median voter conditional on optimal taxation as

$$U_{i,r}^{D} = y_P + \frac{Y_c^{\frac{\theta}{1-\theta}}\left(\left[1-d_P\right]\left[1-D\right]^{\theta} + \alpha_P^{\theta} D^{\theta}\right)^{\frac{1}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(1-\theta\right) \tag{23A}$$

such that the median voter will decide to set the degree of decentralization, D^*, at

$$\begin{aligned} \frac{\partial U_P^D}{\partial D} =& \frac{Y_c^{\frac{\theta}{1-\theta}}\left(\left[1-d_P\right]\left[1-D\right]^{\theta} + \alpha_P^{\theta} D^{\theta}\right)^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left(1-\theta\right)\left(\theta\left(1-d_P\right)\left(1-D\right)^{\theta-1}\left(-1\right) + \theta\alpha_P^{\theta} D^{\theta-1}\right) = 0 \\ & \frac{\left(1-d_P\right)}{\left(1-D\right)^{1-\theta}} = \frac{\alpha_P^{\theta}}{D^{1-\theta}} \\ & \left(\left[1-d_P\right]^{\frac{1}{1-\theta}} + \alpha_P^{\frac{\theta}{1-\theta}}\right) D = \alpha_P^{\frac{\theta}{1-\theta}} \\ & D_P^* = \frac{\alpha_P^{\frac{\theta}{1-\theta}}}{\left(1-d_P\right)^{\frac{1}{1-\theta}} + \alpha_P^{\frac{\theta}{1-\theta}}} \end{aligned} \tag{24A}$$

such that equilibrium decentralization increases in the amount of scale economies that are present in the poor regions, α_P, and their degree of political heterogeneity, d_P.

To verify the extent to which fiscal decentralization can effectively deter inefficient separatism, note that, conditional on the equilibrium income tax of equation (22A) and the equilibrium amount of decentralization of equation (24A), individual utility of the

inhabitants of the rich region can be written as

$$U_{i,R}^{D} = y_R + \frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_P^{\frac{\theta}{1-\theta}}}\left((1-d_R)(1-d_P)^{\frac{\theta}{1-\theta}} + \alpha_R^{\theta}\alpha_P^{\frac{\theta^2}{1-\theta}} - (1-d_P)^{\frac{1}{1-\theta}}\theta\frac{y_R}{y_P} - \alpha_P^{\frac{\theta}{1-\theta}}\theta\frac{y_R}{y_P}\right) \tag{25A}$$

On the other hand, recall that equation (5) stipulates that rich agents could obtain the following utility under secession:

$$U_{i,R}^{R\notin c} = y_R + \frac{1}{\theta}\left(\frac{\alpha_R Y_c}{y_R}\right)^{\frac{\theta}{1-\theta}}(1-\theta) \tag{26A}$$

This implies that the rich region's median voter will not favor seceding from the federal union whenever $\Delta_R^D = U_{i,R}^{R\notin c} - U_{i,R}^{D} \leq 0$, or

$$\begin{aligned}
&\frac{Y_c^{\frac{\theta}{1-\theta}}}{\theta y_R^{\frac{\theta}{1-\theta}}}\left((1-\theta)\alpha_R^{\frac{\theta}{1-\theta}} - \left(\frac{y_R}{y_P}\right)^{\frac{\theta}{1-\theta}}\left[(1-d_R)(1-d_P)^{\frac{\theta}{1-\theta}} + \alpha_R^{\theta}\alpha_P^{\frac{\theta^2}{1-\theta}} - (1-d_P)^{\frac{1}{1-\theta}}\theta\frac{y_R}{y_P} - \alpha_P^{\frac{\theta}{1-\theta}}\theta\frac{y_R}{y_P}\right]\right) \leq 0 \\
&(1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}} \leq (1-d_P)^{\frac{1}{1-\theta}}\left(\frac{1-d_R}{1-d_P} - \theta\frac{y_R}{y_P}\right) - \alpha_P^{\frac{\theta}{1-\theta}}\left(\theta\frac{y_R}{y_P} - \left(\frac{\alpha_R}{\alpha_P}\right)^{\theta}\right) \\
&(1-d_P)^{\frac{1}{1-\theta}}\left(\theta\frac{y_R}{y_P} - \frac{1-d_R}{1-d_P}\right) \leq \alpha_P^{\frac{\theta}{1-\theta}}\left(\left(\frac{\alpha_R}{\alpha_P}\right)^{\theta} - \theta\frac{y_R}{y_P}\right) - (1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}}
\end{aligned} \tag{27A}$$

Note that fiscal decentralization thus fully stabilizes efficient country borders as long as $\left\{\forall d_P \in [0, d^*] : (1-d_P)^{\frac{1}{1-\theta}}\left(\theta\frac{y_R}{y_P} - \frac{1-d_R}{1-d_P}\right) \leq \alpha_P^{\frac{\theta}{1-\theta}}\left(\left(\frac{\alpha_R}{\alpha_P}\right)^{\theta} - \theta\frac{y_R}{y_P}\right) - (1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}}\right\}$. To derive an explicit solution for the critical value of d_P for which this inequality is satisfied, I approximate $(1-d_P)^{\frac{1}{1-\theta}}$ by 1, which should produce reasonably accurate results for low critical values of $d_P \approx 0$. Solving for the threshold value of d_P, I find that

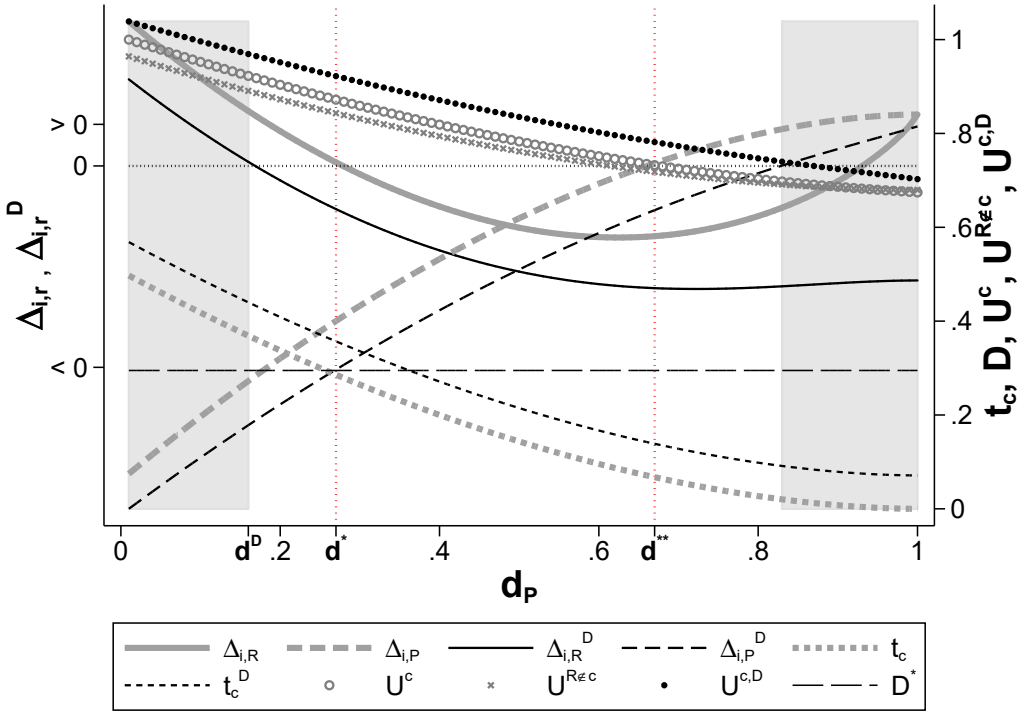
$$\begin{aligned}
\frac{1-d_R}{1-d_P} &\geq \theta\frac{y_R}{y_P} - \alpha_P^{\frac{\theta}{1-\theta}}\left(\left(\frac{\alpha_R}{\alpha_P}\right)^{\theta} - \theta\frac{y_R}{y_P}\right) + (1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}} \\
\underline{d^D} &\geq 1 - \frac{1-d_R}{\theta\frac{y_R}{y_P} + (1-\theta)\left(\frac{\alpha_R y_P}{y_R}\right)^{\frac{\theta}{1-\theta}} + \alpha_P^{\frac{\theta}{1-\theta}}\left(\theta\frac{y_R}{y_P} - \frac{\alpha_R^{\theta}}{\alpha_P^{\theta}}\right)}
\end{aligned} \tag{28A}$$

Thus, if the right side of this inequality exceeds zero, it is impossible to fully stabilize efficient country borders using fiscal decentralization alone. Note that the numerical application in section 2.4 offers an example where fiscal decentralization can not stabilize efficient country borders if political heterogeneity in the poor regions becomes *too low*, as $\underline{d^D} \approx 1 - \frac{1-.25}{.45\frac{8000}{4000} + (1-.45)\left(\frac{.18\times 4000}{8000}\right)^{\frac{.45}{1-.45}} + .09^{\frac{.45}{1-.45}}\left(.45\frac{8000}{4000} - \left(\frac{.18}{.09}\right)^{.45}\right)} = .177$, which is arbitrarily close to the observed threshold value of .182 plotted in figure 4.

Furthermore, if political heterogeneity in the poor regions falls below 0.18, it may even become impossible to determine a level of fiscal decentralization to convince the rich region to voluntarily remain in the union. To demonstrate this, figure A2 plots how the

net gain of independence for the rich region, Δ_R^D, evolves if income taxes are determined by majority voting *but the rich region can subsequently unilaterally set the degree of fiscal decentralization*. As the figure shows, even if the rich region receives full control over the degree of decentralization, inefficient break-up *still* occurs whenever $d_P < d^D$.[28]

Figure A2: **Stability when the rich region determines fiscal decentralization,** D^*



Note: This figure plots how the net gain of independence for the rich and poor regions, $\Delta_{i,R}^D$ and $\Delta_{i,P}^D$, and the equilibrium national tax rate, t_c^D, evolve under fiscal decentralization, D^*, when it is not determined by majority voting but unilaterally determined by the rich region's median voter. The small black dots show the evolution of aggregate welfare in the federal union, $U^{c,D}$, as defined in equation (8), normalized to reach a maximal value of 1 on the right axis. The grey areas depict levels of preference heterogeneity driving at least one region to declare independence, with d^D and d^{**} the corresponding critical values of d_P.

[28] Note that, as the numerical exercise keeps the rich regions preference distance to public policy, d_R, fixed, its preferred degree of fiscal decentralization, D^*, will be kept constant, see equation (10).

B The fiscal effects of electoral polarization

Section 2 demonstrates how the combination of two canonical models of country border formation of Alesina and Spolaore (1997) and Bolton and Roland (1997) yields the comparative static result that preference heterogeneity has a tax-reducing effect. To empirically verify this result, the crucial question is how to best measure preference heterogeneity using available data. Section 3 develops a simple electoral model of collective action to answer this question, suggesting an electoral fractionalization index to be the most appropriate quantitative measure to capture the intensity of real-world preference heterogeneity based on electoral data. Montalvo and Reynal-Querol (2005) nevertheless argue that political diversity could also be measured by an alternative polarization index. This appendix sheds further light on why the choice between a fractionalization and a polarization index matters and why a polarization index would be inappropriate in the current context.

Figure A1 immediately illustrates why the choice between a fractionalization and a polarization index is not only theoretically relevant, but also practically important. It more specifically shows the empirical relation between electoral fractionalization and electoral polarization in our data, when electoral polarization, EP, in the notation of section 3.1, is computed á la Montalvo and Reynal-Querol (2005, p. 798) as[29]

$$PP_{c,t} = 4 \sum_{p=1}^{P} v_{c,t,p}^2 \left(1 - v_{c,t,p}\right) \tag{29A}$$

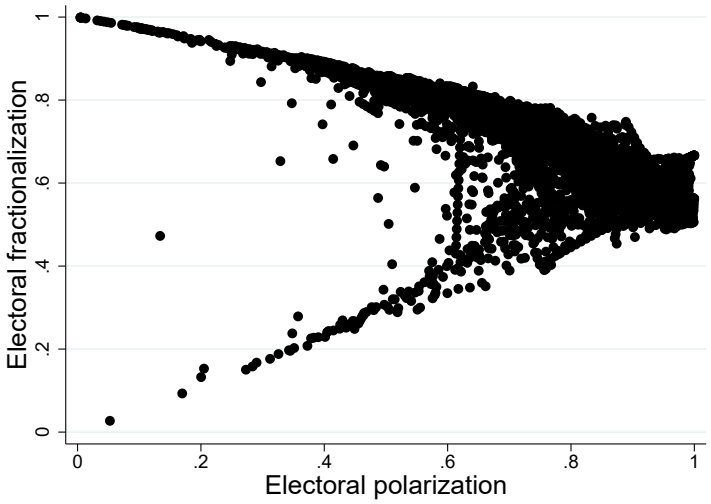
The figure clearly demonstrates that electoral fractionalization and electoral polarization represent different concepts, as the correlation between both indexes turns from positive for low levels of fractionalization to negative for high levels of fractionalization.

To see why a polarization index would be an inappropriate measure for preference heterogeneity in the context of the voting model of section 3.1, it is useful to point out that the vast majority of over 99% of the party vote shares in the electoral data are below 50%. Consequently, an electoral polarization index will predominantly have low values when voting patterns are fragmented over many small parties while reaching its maximal value when party votes are concentrated in a small number of disproportionally large parties. For this reason, the polarization index is strongly *negatively* correlated with the fractionalization index for our dataset, with a correlation coefficient of -.66 for the 6015 available country-year observations. In this sense, the electoral polarization index could perhaps be better described as a preference *homogeneity* index, as high values are associated with a lack of small parties and a strong concentration of votes in a few large parties, which will therefore have an easier time to satisfy a large fraction of the electorate with their choice of the location of the public goods bundle when becoming pivotal.[30]

[29] As discussed in section 3, implicitly assuming that preference heterogeneity evolves slowly between election years, non-election years are linearly interpolated whenever possible.

[30] Moreover, note that the lack of politically relevant small parties implies the lack of sharply dissenting political views in the electorate in terms of the voting model outlined in section 3.1.

Figure A3: **Electoral fractionalization versus electoral polarization**



Note: This figure plots the empirical relation between electoral fractionalization, as defined in equation (13), and electoral polarization, as defined in by equation (29A).

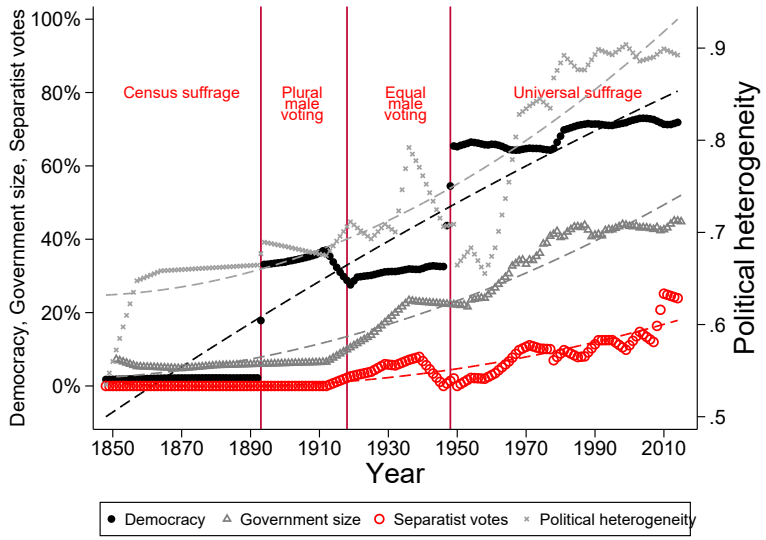
For this reason, it is also easy to see that the electoral model of section 3.1 would in fact predict the willingness to pay taxes to *increase* - rather than decrease - in measured electoral polarization. Given the strong negative correlation between electoral fractionalization and polarization, it comes as no surprise that this prediction is also borne out in the data, see table A1. More specifically, this table reports how the coefficients of interest in table 1 are reversed when substituting the polarization for the fractionalization index. In other words, the results now indicate that electoral polarization *increases* the willingness to pay taxes and especially so in centralized countries. This puzzling finding can be rationalized by the fact that the polarization index in this particular context measures preference homogeneity rather than heterogeneity, for the reasons discussed above.

Table A1: **The fiscal effects of electoral polarization**

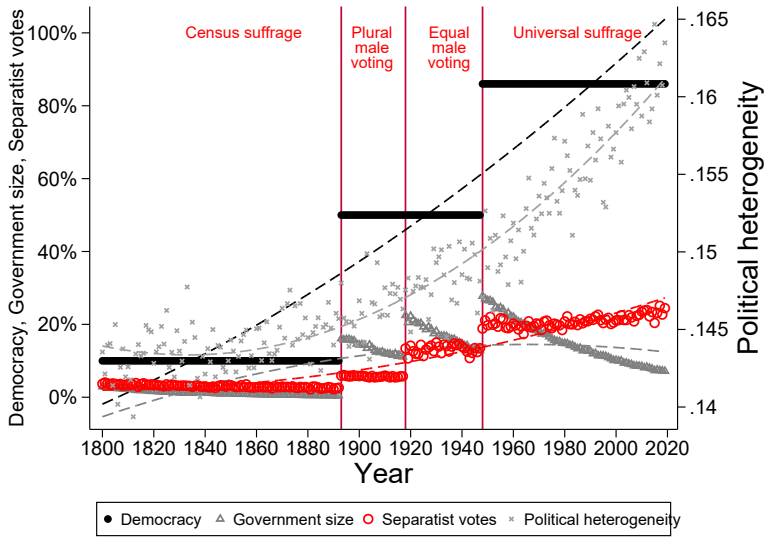
Channel	*(1)*	*(2)*	*(3)*	*(4)*
Electoral polarization	.043***	.133***	.062***	.164***
	(.008)	*(.02)*	*(.021)*	*(.031)*
Intermediate decentralized		.076***	.049***	.074***
		(.013)	*(.013)*	*(.018)*
Most decentralized		.072***	.029*	.108***
		(.016)	*(.017)*	*(.023)*
Electoral polarization × Intermediate decentralized		-.127***	-.072***	-.125***
		(.021)	*(.021)*	*(.033)*
Electoral polarization × Most decentralized		-.177***	-.101***	-.204***
		(.024)	*(.026)*	*(.04)*
Population size (log)	-.026***	-.03***	-.034***	-.039***
	(.002)	*(.004)*	*(.004)*	*(.004)*
GDP (log)	.01***	.016***	.019***	.031***
	(.002)	*(.003)*	*(.004)*	*(.004)*
Trade openness	.008**	.003	-.008**	.005
	(.004)	*(.005)*	*(.004)*	*(.007)*
Democracy	.001	.001	.001	.002
	(0)	*(0)*	*(0)*	*(0)*
Income inequality			0	.002
			(0)	*(0)*
Subnational payroll share				-.024***
				(.006)
Observations [# countries]	2165 [94]	1243 [55]	1132 [53]	461 [30]
Adjusted R-squared	.329	.356	.275	.469
Year fixed effects	Yes	Yes	Yes	Yes

Note: Bootstrapped standard errors, based on 500 iterations, between brackets.

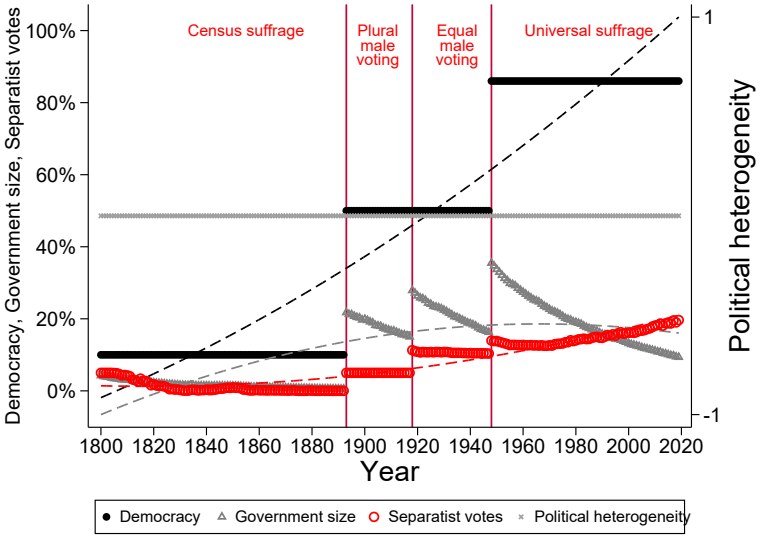
Figure A4: **Observed and simulated results**



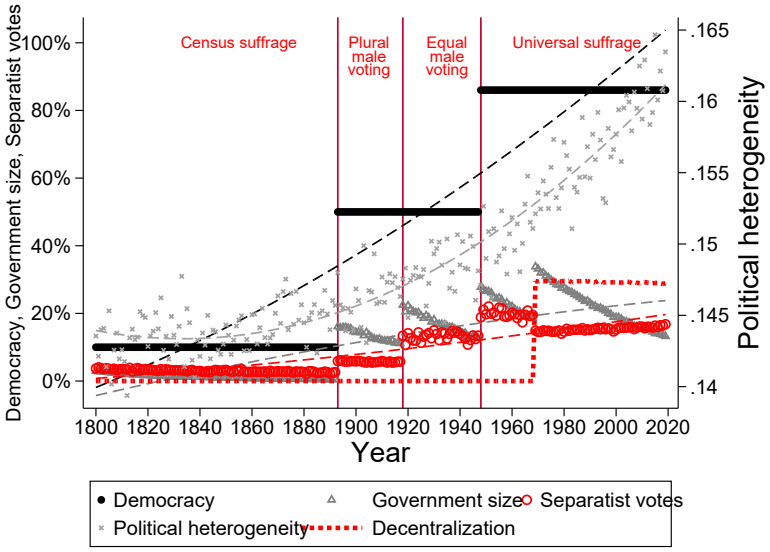
(a) Observed data



(b) Benchmark simulation



(c) Extension without heterogeneity



(d) Extension with fiscal federalism

Note: This figure plots the observed and simulated trajectories of democratization, government size, separatist vote shares, political heterogeneity and decentralization in Belgium between 1800 ($t = 1$) and 2020 ($t = 220$), as explained in section 3. The black dots plot the population share eligible to vote, the triangles capture the tax revenue share, the X-es show preference heterogeneity and the hollow dots track (predicted) separatist vote shares. In figure A4d, the dashed thick red line shows the optimal degree of decentralization. The red vertical lines track the democratization waves in Belgium.

Table A2: **The fiscal effects of electoral fractionalization**

Channel	*(1)*	*(2)*	*(3)*	*(4)*
Electoral fractionalization	-.045***	.441***	.281***	.402*
	(.012)	*(.101)*	*(.104)*	*(.212)*
Tax revenue centralization		.664***	.471***	.509***
		(.095)	*(.098)*	*(.184)*
Electoral fractionalization × Tax revenue centralization		-.589***	-.342***	-.505**
		(.13)	*(.131)*	*(.238)*
Population size (log)	-.026***	-.037***	-.042***	-.04***
	(.002)	*(.003)*	*(.003)*	*(.004)*
GDP (log)	.01***	.027***	.031***	.034***
	(.002)	*(.003)*	*(.003)*	*(.004)*
Trade openness	.005*	.002	-.012***	-.014**
	(.003)	*(.005)*	*(.004)*	*(.006)*
Democracy	.001	.002	.001	.002
	(0)	*(0)*	*(0)*	*(0)*
Income inequality			0	.001
			(0)	*(0)*
Subnational payroll share				-.014**
				(.006)
	2165 [94]	1243 [55]	1132 [53]	461 [30]
	.324	.457	.404	.477
	Yes	Yes	Yes	Yes

Note: This table reports how the results reported in table 1 change when replacing the categorical version of the tax revenue centralization index with a continuous version. Bootstrapped standard errors, based on 500 iterations, between brackets.