Who Sent You?
Extreme Voting, Transfers and Bailouts in a Federation

Gianmarco Daniele[a]
Bocconi University & IEB

Amedeo Piolatto[b]
UAB & BGSE & Move & IEB

Willem Sas[c]
University of Stirling & KU Leuven

Abstract

Lower-level governments often receive federal support through transfers or bailouts. We study how the regional ties of federal politicians can steer this process. We build a two-tier model of government, where regionally-elected federal legislators bargain over federal support aimed at their own constituency. This leads to strategic voting on the regional level. Federal legislators are strategically elected to watch over the interests of their own region, cushioning shocks to local consumption and driving down borrowing costs. Lower-level legislators anticipate this, which sets the stage for regional over-borrowing. Voters strategically select federal representatives with more extreme positions than the median voter, as long as federal co-funding schemes imply some degree of interregional redistribution. These theoretical predictions are confirmed by our empirical analysis, where we compare the political extremism of representatives elected to the EU Parliament with that of representatives elected to national Parliaments.

[a]Bocconi University. E-mail: gianmarco.daniele@unibocconi.it

[b]Autonomous U. of Barcelona (UAB), Barcelona Graduate School of Economics (BGSE), MOVE and Barcelona Economics Institute (IEB). E-mail: amedeo.piolatto@uab.es

[c]Economics Division, University of Stirling & Department of Economics, KU Leuven. E-mail: willem.sas@stir.ac.uk

1 Introduction

When political authority is shared across multiple levels of government – for example in federal countries or political unions such as the EU – lower-level jurisdictions (henceforth, regions) usually finance only a part of their expenditures with own taxes or user fees. The remainder is paid by a higher-level jurisdiction (henceforth, the federation), or borrowed from financial markets. Across OECD countries, for example, regional revenue-raising only covers 65% of lower-level spending, a number which is even lower in the developing world (World Bank, 2000).[1] For such a system to function, federal governments allocate grants to their regions, which usually implies a certain degree of interregional redistribution. They may even introduce bailout schemes if regional spending is such that regions are close to bankruptcy. Now, while convincing arguments in favour of this constellation exist, there are also significant drawbacks. In this paper we focus on 'soft budget constraints', i.e. when regions have a tendency to over-spend and over-borrow, because they expect to be financially supported by the federal government.[2]

The literature on soft budget constraints has so far understated the relevance of political economy elements that, we argue, are key to understand the problem. If politicians respond to incentives, what drives a central government to support local governments? The usual argument is that federal politicians simply have no choice, and have *given* motives for intervention. This may be intuitive in the presence of sudden shocks and spillovers – when for example police departments, schools, or hospitals start closing – or when the success of local governments is seen as a shared federal responsibility. In the more frequent cases where federal governments gradually scale up grant finance for no immediate reason, however, this point is less convincing. It furthermore only partially explains why regional politicians can expect to be supported.

Our aim is to study how specific political and voting processes – unique to multi-layered democracies – can affect public spending and bailout expectations of regional governments. We argue that both the bargaining strategies of federal politicians and their regional ties have a vital role to play in obtaining federal support, which is the essential link in the chain. Indeed, if redistribution mechanisms on the federal level could not be manipulated, forward-looking regional politicians would have no direct incentive to overspend in the first place. To fully understand these dynamics, our theoretical model opens up the black box of political and voter incentives at the federal and sub-national level. Our argument will be

---

[1] Own calculations for 2012: http://www.oecd.org/tax/federalism/oecdfiscaldecentralisationdatabase.htm.

[2] Budgets are called 'soft' when organisations *expect*, in case of financial trouble, to be bailed out by a supporting organisation. Kornai et al. (2003) or Qian and Roland (1998) were first to cast these dynamics as a moral hazard problem where an agent ex-ante anticipates the motive of the principal ex-post.

that voters strategically try to stack interregional redistribution in their favour, by voting for more extreme representatives in federal elections.

We model a setting where the federal government is composed of a coalition of regionally elected politicians, whose electoral survival depends on voter welfare in their respective regional constituency. An extreme example of this constellation would be the European Council within the current EU framework. Policy decisions at the highest level are made in a more or less cooperative way, so that the elected representatives must reach a mutually advantageous agreement. Our inter-temporal setting allows regions to tax and borrow, and the federal government to co-finance regional spending in a redistributive way. Then, anticipating how the grant or bailout scheme is negotiated, regional politicians and voters alike are in a position to take advantage of the redistributive system. If so, it can also be in their strategic interest to make (socially) ill-considered decisions such as over-borrowing.

Our model shows that voters strategically elect federal candidates that are more extreme in safeguarding regional welfare than they themselves would be. This makes federal representatives more likely to compensate for regional spending and borrowing. Indeed, whenever the subsequent debt repayments start weighing on general welfare in the region they represent, they will readily bargain for increased federal support. For net receiving regions this support comes in the form of higher federal grants and bailout funds, and vice-versa for net contributing regions. Anticipating the bargaining process at the heart of federal politics, voters thus turn this behaviour to their advantage by voting for precisely these more supportive and extreme types.[3] Interestingly, voters' strategic behaviour is not limited to the choice of the federal representative. When it comes to local policies, voters push for the excessive provision of local public goods which is financed by above-optimal borrowing. This way, voters strategically ensure the hands of their federal representatives are tied once negotiations at the federal level start, which pushes them to be tough negotiators.

As a result, the theoretical model predicts excessive political extremism and polarisation at the federal level, as well as inefficiently high spending and borrowing at the regional level. Our model also shows that the strategic distortion of voter preferences, and the resulting regional overspending, is present both in net receiving and net contributing regions, but is more pronounced in the poorer ones. When federal decisions are taken by a minimum winning coalition instead of the grand coalition lastly, our results are quantitatively affected, but qualitative patterns are not.

We model the electoral process as follows. Citizens in each region cast their vote both in regional as well as federal elections, in a "citizen-candidate" type of setting. Federally

---

[3]An alternative interpretation here would be that federal candidates market themselves as advocates of regional interests, to attract more votes.

elected representatives then go on to defend the interests of their respective regions in the federal government. To varying degrees, the bargaining process in the federal legislation and government thus reflects regional considerations, a phenomenon that is often observed in real-world federations. Moreover, and also argued by Rodden (2002) and Rodden et al. (2003), the fact that the legislative process in federations is characterised by log-rolling and complex regional bargaining can lead to vote-trading on interregional redistribution as well. Indeed, whenever decisions taken on the federal level follow more from consensus than a simple majority voting rule,[4] and when locally elected federal politicians start behaving as regional representatives for electoral reasons, we can expect region-oriented bargaining to direct most of distributive federal policy-making. We formalise this process in our model, and thus analyse the unique territorial representation of federalism in the context of interregional redistribution.

The federal coalition in our model can be thought of as the result of a specified institutional arrangement, as e.g. is the case for the German 'Bundesrat' where every member state has a guaranteed seat at the table.[5] More broadly, however, the direct representation of regions in the central government can also be seen as the product of the political constellation at hand. Nation-wide political parties will always consist of candidates elected in all regions of the federation, hence any coalition formed on the federal level to a certain extent represents each region when it comes to making decisions on federal taxation and transfers.[6] Furthermore, our qualitative results go through when the federal bargaining process takes place in a smaller coalition representing only a subset of all regions.

Finally, we provide an empirical validation of the first main prediction of the model regarding political extremism, focusing on the European Union (EU). The EU is the world biggest supra-national federation, where representatives of member states continuously bargain over a common budget – which in 2018 amounted to 160 billion Euros. Moreover, some countries are net contributors to the common budget (e.g. the Netherlands or Germany), others are net receiving ones (e.g. Poland or Bulgaria), while still others contribute as much as they receive. This constellation shows many similarities to our model, once we conceive of the European elections as applying to the federal level, and elections in each EU member state to our regional level. Following our theoretical predictions therefore, we should expect strategic

---

[4]Duchacek (1985) describes the variation in decision-making at the highest level of government in a multi-tiered state, ranging from purely consensus-driven 'consociationalism', to 'compound' or outright majoritarian formulas. See also Lijphart (1985) or Andeweg (2000) on the various gradations in 'consociational' democracy.

[5]A drawback of this interpretation is the fact that in these councils there is no direct election of delegates, rather they are assigned by lower-level governments (Bundesrat) or are composed of the heads of government themselves (EU). Hence, the same politicians can be thought of as having a regional and a federal agenda, between which voters can distinguish.

[6]This also goes for countries with region-specific parties, such as Belgium, Spain or Canada. If such parties are part of the federal government their region is automatically represented as well.

voting behaviour to mark European elections but leave national elections unaffected. Indeed, and as compared to national election outcomes, our empirical analysis shows that voters elect relatively more extreme and Eurosceptic delegates to the European parliament. These results are robust to different specifications, definitions of the key variables and a varied set of controls. Crucially, and fully in line with our predictions, this strategic voting only takes place in net receiving and net contributing member states, but not in countries where EU contributions and expenditures are more or less balanced.

In what follows, we present a general setting in Section 2 which is applied to multiple forms of interregional redistribution. In Section 3 the federal government simply puts a grant system in place which co-finances regional spending. Starting from a simplified setting where the federation counts only two regions in Section 3.1, we derive the core results described above. In Section A.1 of the appendix we prove that these results remain valid – but are attenuated – when voters expect their representative to have a potentially limited influence on federal policies. This could occur, for example, when the process of forming a coalition at the federal level implies some randomness, which may cause regions to be excluded from the coalition with some probability. Such a specific legislative setting, which requires at least three regions, is formalised in sections Section A.2 and Section A.3 of the appendix. Next, Section A.4 shows that our results emerge independently of whether interregional redistribution occurs through a continuous grant-based transfer or one-off rescues using a bailout fund. We modify some core elements of the baseline model and introduce some uncertainty about the future, subject to income shocks, and are able to show that the two models converge to the same results. Lastly, the empirical findings based on the European Union elections are reported in Section 4.

Related Literature

Our paper first of all contributes to the ongoing research on soft budget problems in multi-layered systems of government. Empirical work in this area has mostly investigated whether a central government does indeed cover regional deficits. Case studies focusing on specific countries abound,[7] but broader cross-country studies by e.g. Rodden (2002) or Eyraud and Lusinyan (2013) have been undertaken as well. Most of these arrive at more or less the same conclusion that there is a relationship between what is labelled as bailouts and regional deficits or debt – where bailouts can vary from explicit rescue attempt in the face of imminent defaults, to a more gradual build-up of existing grant programs. Since the relation is conditional on there being a significant degree of grant-finance in place, the results are often interpreted as

---

[7]See e.g. Jones et al. (1999) for Argentina; Buettner and Wildasin (2006), Poterba (1995) and Von Hagen (1991) for the US; Borge and Rattsø (2003) for Norway; Faguet (2004) for Columbia and Venezuela; Von Hagen and Dahlberg (2004) for Sweden; and Sorribas-Navarro (2011) for Spain.

proof of bailout expectations as well. However, most of these studies do not identify bailout expectations explicitly, lacking clear and testable theoretical predictions in this respect. In that light, our theoretical results also add to a more recent strand of the empirical literature, which tries to achieve exactly this more explicit identification: e.g. Baskaran (2012) for Germany, Pettersson-Lidbom (2010) for Sweden and Bordignon and Turati (2009) for Italy. The downside of these studies is that they are context-dependent, however, so that results have limited external validity. Further theoretical and empirical insights are therefore needed to gain a broader understanding of the problem.

The theoretical work on soft budgets cited earlier relies on the decentralised leadership literature pioneered by Caplan et al. (2000), and later on extended by Köthenbürger (2004) and others, which is founded on two assumptions. First, lower-level governments are first movers in a sequential game, with the central government responding to policy choices made by lower levels. Second, the central government acts as a unified and single body, usually as a benevolent planner maximising the sum total of regional utilities.[8] Bailout expectations and soft budgets automatically materialise in such a setting, since regional politicians are assumed to move first and can anticipate a given degree of federal redistribution ex-ante. Such a framework thus abstracts from the underlying reasons why federal governments would in fact intervene – and why the latter can anticipate this support – as it was mainly used to explore other, more specific questions.[9] Instead, by opening up the black box of benevolent governments and modelling the relevant political channels explicitly, our model endogenises interregional redistribution and bailout expectations.

Second, our results regarding strategic voting lie at the intersection of the broader subfields of fiscal federalism and legislative bargaining, as well as political economy in general. Knight (2008) recognises that grant receipts are the outcome of a bargaining game at the federal level, and may reflect constituent preferences through their elected representatives. His model shows that, when forming majority coalitions, the committee chair (formateur) prefers to include those delegates with relatively strong preferences for public goods since their vote is cheaper to secure. The latter are not elected strategically, however. In Besley and Coate (2003) and Harstad (2010) strategic delegation does enter the electoral decision-making of

---

[8]In (Caplan et al., 2000; Sanguinetti and Tommasi, 2004; Köthenbürger, 2004; Besfamille and Lockwood, 2008; Akai and Sato, 2008; Breuillé and Vigneault, 2010; Akai and Sato, 2011) for example, a benevolent federal government equalises marginal utilities across regions. A smaller literature adds a channel of interregional spillovers where regions become 'too big' (Wildasin, 1999) or 'too small' (Crivelli and Staal, 2013) to fail, depending on the way the horizontal interactions are modelled.

[9]In Goodspeed (2002) regional utility is multiplied by a probability of re-election which is exogenously assumed different across states, and which diverts federal bailouts to the states where the most votes are to be found. Robinson and Torvik (2009) also conceptualise bailouts as a means to transfer income to potential supporters, namely entrepreneurs. Our point in this paper however, is that political decision making is rarely this top-down oriented in federations, and federal politicians will rarely have the same objectives.

voters, in a similar setting of higher-level co-financing. In the former, strategic incentives only materialise when the grand coalition is assumed to form. In the latter, the strategic motive to vote progressively – which comes closest to our notion of extreme voting here – arises to enhance the chances of a given delegate to be included in the coalition, of which all members share the same goal: the expropriation of non-coalition members. Similarly, in Krasa and Polborn (2015) the end-goal of voters, if they value the national party position, is dominating the national election at the expense of the opposition, and at the possibly very high cost of electing an extreme party candidate in each district. Inversely, in our model the strategic reflex of voters derives from the intention to mitigate the objectives of coalition partners because – being on either side of the redistributive mechanism – these will have *opposite* goals.[10] In Mattozzi and Snowberg (2018) decisions on federal taxation and regionally targeted spending are made in a two stage process, where a majority of legislators of the same type set their preferred tax rate in the first stage, and strategically elected legislators can attract more pork to their constituency in the second stage. Our model can be thought of as simultaneously combining both stages, since the proposing power of each representative is equally weighted and redistributive shares are fixed, so that strategic delegation mostly centres on what happens once a coalition of opposing types is formed.[11]

Third, the tendency of voters in our model to condone over-spending relates to Gavazza and Lizzeri (2009), where this kind of "fiscal churning" is caused by the imperfect observability of electoral promises. This generates an incentive for candidates to offer excessive transfers even if voters are homogeneous. Our result that over-spending also has a strategic element, to constrain the options of federal negotiators in the future, is similar to the findings of Hatfield and Padró i Miquel (2012). In their model, however, voters adjust the federal structure of the constitution to partially tie their own hands ex-ante, to rein in their ex-post desire for public spending. Lastly, and to the extent that stronger national parties undercut the regional ties of federal representatives, our predictions also coincide with those of Enikolopov and Zhuravskaya (2007). This kind of stronger political centralisation would indeed better align local political incentives with national interests, as federal legislators in our model become less malleable to regional electoral pressures.

---

[10]In Eerola et al. (2004), voters elect "stubborn" representatives in the context of the European Union framework. Here, threats of secession posed in an ultimatum game played at the European level are the main drivers of strategic voting. There is no legislative coalition formation process underlying these dynamics, however. Back in the context of legislative bargaining, and adding a delegation stage to the model of Volden and Wiseman (2007), Christiansen (2013) studies strategic voting when legislators face a trade-off between public goods provision and targeted spending on pork projects. In our model there is no such trade-off, as redistributive shares are fixed whilst the budget is endogenous, so that side-payments in the form of particularistic spending are ruled out.

[11]In line with political practice in many federations, our model does rely on a two-stage process regarding the coalition formation process itself. In a first stage the formateur chooses his coalition partners, which in a second stage – and among others – bargain on the federal tax rate and implicitly, interregional redistribution.

2 The Model

We consider a federation where each region $r \in R$ is inhabited by a unit mass of citizens. We focus on a two-period time span $s = 1, 2$. Then, Y_r^s is total income in period s in region r and it is normalised in period $s = 1$ so that $\sum_r Y_r^1 = 1$, whereas income in period $s = 2$ is $Y_r^2 = Y_r^1 + \epsilon_r$. Income can be used to finance either a local public good G_r or private consumption C_r^s.

At the beginning of period 1, local elections are held: voters in each region independently select their preferred level of local public good G_r, as well as the tax t_r^1 that is levied in period 1 to finance it. The difference between what is raised through the local tax and what is needed to finance the public good is obtained through borrowing B_r, which must be paid back at the end of period 2. Regions are residual claimants and therefore in period 2 they must raise a local tax t_r^2 such that they can repay the debt.

In period 1, voters also select a federal representative that joins the federal government. Each region thus marks a constituency where every single voter is a potential candidate, from which one candidate is chosen to represent their region at the federal level. In period 2, a coalition of federal representatives bargains on interregional transfers T_r, which are modulated either according to a grant-based mechanism or a bailout-based mechanism as further explained below. We denote by $k \in K \subseteq R$ the regions/representatives that belong to the coalition. We assume that, within the federal coalition, decisions are taken cooperatively so that federal transfers maximise the joint utility of the coalition members. When $K = R$, a grand coalition forms that includes all regions. As mentioned earlier, such a coalition can be thought of as the result of a specified institutional arrangement or as the product of the political constellation at hand. Either way, cooperative bargaining is driven by a threat point which we conceive as the risk of re-election after a continued disagreement, with the eventual loss of power as a result. This translates into the disutility of losing political benefits, wages, or more generally, all possible kinds of (ego)rents enjoyed simply by staying in office.

Voters differ in the weight that they put on future consumption (θ) and in their aversion (η) to be represented by politicians with preferences different from theirs. The latter thus expresses the disutility of being represented by someone in a different ideological camp and can be interpreted as a measure of strategic tolerance. Both θ_r and η_r are exogenous and weakly positive. Parameter η_r is region-dependent. We do not impose further restrictions on θ_r, that can be heterogeneous within and across regions and follow any distribution with weakly positive support. We denote by $\hat{\theta}_r$ the preferences (hence the type) of the regional representative within the federal government.

The utility of voter j living in region r is then defined as

$$U_r^j = g(G_r) + c(C_r^1) + \theta_r^j c(C_r^2) + M_r, \tag{1}$$

where functions g and c are increasing and strictly concave and $M_r = -\eta_r \left(\hat{\theta}_r - \theta_r^j \right)^2$ is the function that measures the utility of the "political match", that is, the distance for a voter between their type θ and the one of the elected federal representative of their region.

The variables in the model are then

$$G_r = B_r + t_r^1 Y_r^1, \tag{2}$$

$$C_r^1 = \left(1 - t_r^1\right) Y_r^1, \tag{3}$$

$$C_r^2 = \left(1 - t_r^2\right) \left(Y_r^2 + \Delta_r^f\right), \tag{4}$$

$$B_r = t_r^2 \left(Y_r^2 + \Delta_r^f\right). \tag{5}$$

Eq. (2) defines the amount of public good G_r, financed either through debt B_r or through the local tax raised in period 1. Eq. (3) and Eq. (4) define consumption in each period as the net, after-tax disposable income. Eq. (4) in particular states that in period 2 the local tax is levied on the total gross income computed as the sum of regional income Y_r^2 and Δ_r^f, which is the *net* value of federal transfers. The latter can take either positive or negative values, with region k being a net recipient whenever $\Delta_r^f > 0$ and a net contributor otherwise. Eq. (5) lastly, ensures that the local tax t_r^2 is such that the debt B_r is repaid in the second period.

Combining Eqs. (2) and (5) with Eq. (4), it is possible to rewrite Eqs. (2) to (5) as

$$B_r = G_r - t_r^1 Y_r^1, \tag{6}$$

$$t_r^2 = \frac{B_r}{\left(Y_r^2 + \Delta_r^f\right)} = \frac{G_r - t_r^1 Y_r^1}{\left(Y_r^2 + \Delta_r^f\right)}, \tag{7}$$

$$C_r^1 = \left(1 - t_r^1\right) Y_r^1, \tag{8}$$

$$C_r^2 = Y_r^2 + \Delta_r^f - B_r = Y_r^2 + \Delta_r^f + t_r^1 Y_r^1 - G_r. \tag{9}$$

Clearly, Eq. (1) meets all the requirements in Gans and Smart (1996) and therefore the median voter theorem applies. We solve the model backward, first computing the federal transfers selected by the federal government in period 2. Subsequently, we analyse the decisions made by the median voter of each region r in period 1.

In Section 3 – the core of the paper – we introduce some simplifying assumptions to the setting that we just illustrated. We interpret the resulting model as a case in which the federal government allows redistribution across regions in the form of grants. As previously discussed, Sections A.1 to A.4 in the online appendix extend the model further.

3 The grants system

In the grant setting there is no need to assume income shocks. For the sake of parsimony, we assume that $\epsilon_r = 0$ and hence that $Y_r^1 = Y_r^2 \equiv Y_r$. Under the grants system, the federal government raises a federation-wide tax t_f, which is used to finance public spending on interregional redistribution in its most general sense.[12] Federal tax proceeds $\sum_r (t_f Y_r) = t_f$ are thus fully shared across regions, so that region r receives a proportion $\lambda_r \in [0, 1]$. We assume λ_r to be exogenous, which reflects the fact that finding federation-wide consensus on how to split national proceeds is a difficult and long-winded process. Indeed, either because modifying the status quo requires a qualified majority in the legislature, because negotiations have to be approved by another board or council, or because of other political constraints, the allocation rule of grant schemes will in the short run often be stickier than absolute budgetary volumes. This directly applies to any country where the grant system is at least partially set up as a revenue-sharing scheme using an explicit or implicit sharing rule, such as in Spain, Germany, or Belgium. Yet it would also indirectly apply to any decision on the level of existing grant systems, where the allocation will always – and for various reasons – involve a kind of regional redistribution. In both cases, changes in grant design tend to be more politically feasible when they concern adjustments on the overall grant volumes, rather than the allocation rule itself.

Under the assumption that in this setting $\epsilon_r = 0$, the net value of the federal transfer Δ_r^f in Eq. (5) becomes $\Delta_r^f = (\lambda_r - Y_r)t_f$, where $\lambda_r t_f$ expresses the federal grant that region r receives while $Y_r t_f$ is its own contribution. Introducing the notation $\Delta_r = (\lambda_r - Y_r)$, federal budget balance implies that $\sum_r \Delta_r = 0$. Region r is a net-recipient if $\Delta_r > 0$ and a net-contributor if $\Delta_r < 0$. The federal transfer becomes

$$\Delta_r^f = \Delta_r t_f, \tag{10}$$

and Eqs. (6) to (9) can be written as

$$B_r = G_r - t_r^1 Y_r, \tag{11}$$

$$t_r^2 = \frac{G_r - t_r^1 Y_r}{(Y_r + \Delta_r t_f)}, \tag{12}$$

$$C_r^1 = \left(1 - t_r^1\right) Y_r, \tag{13}$$

$$C_r^2 = (1 + t_r^1)Y_r + \Delta_r t_f - G_r. \tag{14}$$

---

[12]In many countries, this kind of redistribution implies federal grants which are – in one way or another – used to finance public provision, in our case paying back part of the debt incurred in period 1. Introducing restrictions on how federal grants can be used would have no effect at any interior solution.

3.1 Two-region setting

With two regions, $|R| = 2$. In this scenario, we focus on the grand coalition case, that is $K = R$ and both regions are represented in the federal coalition. We denote regions in the coalition as α and β, so that $K = \{\alpha, \beta\}$. With only two regions, if one is a net-recipient the other must be a net contributor, with $\Delta_\alpha = -\Delta_\beta$.

Solving the model backward, Section 3.1.1 analyses the second (last) period decisions, while Section 3.1.2 solves the equilibrium in the first period and discusses the findings.

3.1.1 The Federal Government

We assume this federal government will seek the (weighted) utilitarian bargaining solution to decide on public spending – which under the grants system boils down to a decision on the federal tax t_f to co-finance public or private consumption – and therefore maximises

$$\max_{t_f} \sum_{k \in K} \omega_k \left(\hat{U}_k - F_k \right) = \sum_{k \in K} \omega_k \left(g\left(G_k\right) + c\left(C_k^1\right) + \hat{\theta}_k c\left(C_k^2\right) - F_k \right), \tag{15}$$

where $\hat{U}_k$ is the utility of the representative elected in region k, rewriting Eq. (1) for $\theta_k = \hat{\theta}_k$, ω_k is the weight of region k in the bargaining process, and $F_k = \begin{cases} 0, & \text{if a coalition is formed} \\ \bar{F}_k, & \text{otherwise} \end{cases}$ represents the cost, for an elected politician, of not being able to form a coalition and hence the threat of losing power and its associated benefits. We assume $\bar{F}_k$ to be sufficiently large for coalitions to form. When this is not the case, the solution of the model is isomorphic to the case of no federal government that is analysed at the end of Section 3.1.2.

How a specific candidate rose to power in state k in period 1, in other words why their preferences were chosen above the preferences of the citizens living in a specific region k, will be dealt with in Section 3.1.2 where we discuss the local voting mechanism. For now, it suffices to see that the utility of an elected legislator $\hat{U}_k$ is defined not just by their appreciation of public and private consumption in the region they represent, but also by the weight $\hat{\theta}_k$ in Eq. (15) given to future consumption.

Now, deriving and rewriting the first order conditions of the optimisation problem denoted by Eq. (15) and subject to Eqs. (11) to (14), the optimal federal tax rate t_f is implicitly expressed by

$$\omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c'\left(C_\alpha^2\right) + \omega_\beta \hat{\theta}_\beta \Delta_\beta c'\left(C_\beta^2\right) = 0. \tag{16}$$

Since $\Delta_\beta = -\Delta_\alpha$, we can rewrite Eq. (16) as $\Delta_\alpha \left(\omega_\alpha \hat{\theta}_\alpha c'\left(C_\alpha^2\right) - \omega_\beta \hat{\theta}_\beta c'\left(C_\beta^2\right) \right) = 0$, which leads to

$$\begin{aligned} \phi_f &= 0 \\ \text{with } \phi_f &\equiv \omega_\alpha \hat{\theta}_\alpha c'\left(C_\alpha^2\right) - \omega_\beta \hat{\theta}_\beta c'\left(C_\beta^2\right). \end{aligned} \tag{17}$$

Eq. (17) implicitly captures the way in which period 2 federal tax decisions respond to period 1 regional decisions. In our soft budget constraint setting, this is an important relation. Indeed, whether governments are tempted to over-borrow in period 1 ultimately depends on federal behaviour in period 2. Analysing this reaction with respect to regional public provision G_α, we obtain Lemma 1.

Lemma 1. *Given the optimal federal tax rate t_f, defined by Eq. (17), we find that*

$$\frac{\partial t_f}{\partial G_\alpha} = \frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right) \Delta_\alpha} \tag{18}$$

and

$$\frac{\partial t_f}{\partial t_\alpha^1} = \frac{-\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) Y_\alpha}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right) \Delta_\alpha}. \tag{19}$$

Notice that $\mathrm{sgn}\left(\frac{\partial t_f}{\partial G_\alpha}\right) = \mathrm{sgn}\left(\Delta_\alpha\right)$. *That means that the federal government raises (reduces) its tax rate in response to an increase in local public provision in period 1 in the net recipient (net contributing) region, that is when $\Delta_\alpha > 0$ ($\Delta_\alpha < 0$). The opposite occurs with a change in local taxation t_α^1, given that* $\mathrm{sgn}\left(\frac{\partial t_f}{\partial t_\alpha^1}\right) = -\mathrm{sgn}\left(\Delta_\alpha\right)$. *Also notice that the magnitude of the effect (in absolute terms) is smaller for t_α^1 than for G_α.*

Proof. See appendix B. □

We have thus uncovered a situation where the federal government finds it optimal to increase federal support to the net receiving region, whenever the latter borrows additional funds to finance higher levels of the public good. Since the debt needs to be paid back in period 2, regional borrowing will erode consumption and welfare in that same period, as shown by Eq. (9). The federal government compensates for this drop in consumption with a federal tax hike, financing a larger share of public provision or private consumption via its grant system. The opposite is true for the net contributing region: lowering the federal tax results in higher consumption in period 2 as this will pull federal co-financing down, which serves as a similar compensating measure as this region now has to contribute less. We refer to both of these federal responses as "soft budget constraint" policies, since these will in effect lower the opportunity cost of borrowing for the regions, as we will show in section Section 3.1.2.

What we also learn from Eq. (17) and Lemma 1, is that the federal response to regional borrowing depends on the preference types $\hat{\theta}_\alpha$ and $\hat{\theta}_\beta$ of the federal legislators elected respectively in regions α and β. This is where our model clearly diverges from previous approaches. If for example a region is represented by a legislator who for some reason puts more weight

11

on consumption in period 2, this can be expected to have an effect on debt relief, and consequently, on federal tax responses.[13] We elaborate in Lemma 2.

Lemma 2. *Given the optimal federal tax rate t_f, defined by Eq. (17), we find that*

$$\frac{\partial t_f}{\partial \hat{\theta}_\alpha} = -\frac{\omega_\alpha c'(C_\alpha^2)}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_\alpha}, \tag{20}$$

with $\text{sgn}\left(\frac{\partial t_f}{\partial \hat{\theta}_\alpha}\right) = \text{sgn}(\Delta_\alpha)$. *Hence, for representatives elected in the net recipient (contributing) regions, the more they are in favour of future consumption, the more the federal government raises (reduces) its tax rate in response.*

Proof. See appendix B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The intuition here is that representatives that find consumption in period 2 more important, will generally want to compensate for regional spending on public provision as well, since the latter has to be partially repaid out of period 2 incomes. As a result, and ceteris paribus, they are more generous when compensating for regional borrowing – which is used to finance the additional regional spending. The representative of the net recipient region gladly taps the national common pool a little more and bargains for higher federal tax levels to support period 2 consumption of their region. Inversely, yet similarly in a conceptual sense, the representative of the net contributor will bargain for even lower federal taxes to uphold period 2 consumption in their region as well. Clearly, the question at this point becomes which preference type will emerge from the local election process. Section Section 3.1.2 provides us with an answer.

3.1.2 Local Elections and Government

Local elections define the quantity of public good G_k to be produced in each region, as well as the local tax t_k^1 that partially finances the public good. Voters also select the representative $\hat{\theta}_k$ who joins the federal government and negotiates t_f in the name of region k, staying true to their preference type.

The voters' objective function, defined by Eq. (1), is optimised under constraints imposed by Eqs. (11) to (14). Lemma 3 presents the results of the maximisation.[14]

---

[13]By a less strict interpretation of the model, a legislator with a preference for future consumption could be also replaced by a legislator with a taste for redistribution.

[14]Notice that in Eqs. (21) to (23) both $\hat{\theta}_\alpha$ and θ_α^m appear. This comes from the fact that the median voter is decisive in the election, hence the presence of θ_α^m but they anticipate the strategic role of the federal politician at the federal level. Therefore, $\hat{\theta}_\alpha$ follows from the strategic and forward looking behaviour of the local median voter.

Lemma 3. *Eqs. (21) to (23) implicitly define the preferences of the median voter in region α, in terms of the representative selected to join the federal legislature ($\hat{\theta}_\alpha$), the preferred quantity of public good (G_α) and the local tax (t^1_α) .*

$$\frac{-\omega_\beta \hat{\theta}_\beta c''(C^2_\beta)}{\omega_\alpha \hat{\theta}_\alpha c''(C^2_\alpha) + \omega_\beta \hat{\theta}_\beta c''(C^2_\beta)} = -\frac{c'\left(C^1_\alpha\right)}{\theta^m_\alpha c'\left(C^2_\alpha\right)} \tag{21}$$

$$\frac{-\omega_\beta \hat{\theta}_\beta c''(C^2_\beta)}{\omega_\alpha \hat{\theta}_\alpha c''(C^2_\alpha) + \omega_\beta \hat{\theta}_\beta c''(C^2_\beta)} = -\frac{g'\left(G_\alpha\right)}{\theta^m_\alpha c'\left(C^2_\alpha\right)} \tag{22}$$

$$\frac{-\omega_\alpha c'\left(C^2_\alpha\right)}{\omega_\alpha \hat{\theta}_\alpha c''(C^2_\alpha) + \omega_\beta \hat{\theta}_\beta c''(C^2_\beta)} = \frac{2\eta^m_\alpha \left(\hat{\theta}_\alpha - \theta^m_\alpha\right)}{\theta^m_\alpha c'\left(C^2_\alpha\right)}. \tag{23}$$

Proof. See appendix B. $\qquad\square$

Combining Eq. (21) with Eq. (22), we obtain that for any k it must be that $c'\left(C^1_k\right) = g'\left(G^1_k\right)$. Furthermore, from Eq. (23) it is immediate to observe that a solution can only exist if $\hat{\theta}_k \geq \theta^m_k$. Proposition 1 and its corollaries develop around that result.

Proposition 1 (Strategic Voting)**.** *In a federation, voters strategically elect candidates with a higher preference for future consumption than their own preference when a cooperative legislature of (locally elected) representatives bargains over federal grants that co-finance regional public provision.*

Corollary 1. *The strategic voting of the median voter is characterised by*

$$\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)} = \frac{\omega_\alpha}{\omega_\beta} \frac{\left(c'(C^2_\alpha)\right)^2 / \eta^m_\alpha}{\left(c'(C^2_\beta)\right)^2 / \eta^m_\beta} \tag{24}$$

where $\mu(\theta_k) = \left(\hat{\theta}_k - \theta^m_k\right)/\theta^m_k$ is the mark-up in terms of θ_k, that is, the percentage increase in θ_k from the median voter to the federal representative in region k.

Corollary 2. *If there is no redistribution at the federal level – so that each region contributes to federal tax proceeds the exact same amount that it receives via federal co-financing of public provision – then local elections will not be distorted by strategic voting.*

Proof. See appendix B. $\qquad\square$

According to Proposition 1, hence, the elected federal representative always has a weakly-stronger preference for consumption in period 2 than the median voter. This may be surprising at first, for it says that the median voter is bound to select a representative with preferences different from theirs. And this occurs despite the fact that the median voter's utility depends negatively on the distance between their type and the elected representative type. This is

explained by the fact that the median voter anticipates the federal bargaining process, and hence pre-commits to a more extreme stand to manoeuvre their representative in a – at least for them – more favourable bargaining position at the outset of negotiations.

By symmetry, median voters in both regions vote for candidates that are inclined to safeguard consumption in their own region in period 2, more than they would do themselves $(\hat{\theta}_k \geq \theta_k^m)$. This evidences the *strategic* behaviour of both median voters, who make full use of the knowledge that if they vote for a candidate with preferences of this kind, their welfare will go up. The reason for the latter is given by Lemma 2, which states that the federal tax – and thus the federal contribution to financing the debt – is increasing in $\hat{\theta}_k$ for the net-receiving region, and vice versa for the net-contributing region. From the median voter's perspective therefore, a more extreme representative will always bargain for more consumption in period 2 as shown by Eq. (14), resulting in higher utility following Eq. (1). As emphasised earlier, higher $\hat{\theta}_k$ types mind less when the region they represent accumulates debt, are more generous when compensating for state borrowing, and gladly negotiate for higher, or lower federal tax rates to bring this about. Knowing this in advance, voters turn this behaviour to their advantage by voting in precisely such generous types. Importantly, this kind of strategic voting occurs both in net contributing and net receiving states, but for different reasons. In the net receiving region voters will distort local elections to guarantee more federal co-financing through higher taxes, whilst the net contributing region strategically sends a representative keen on lowering federal taxes. The key channel here is the fact that there is some redistribution at the federal level, which is confirmed by Corollary 2. Indeed, when federal co-financing is conceived as strictly non-redistributive, the level of federal taxation loses its strategic appeal to voters who will then vote sincerely. When there is some redistribution organised at the federal level, however, the question becomes which region will distort the elections the most, and why.

Eq. (24) helps us to analyse the determining factors driving the strategic voting defined in Proposition 1 and to be able to compare these factors between regions. Results imply that the strategic pre-commitment of the median voter (measured by the left-hand side of the equation) depends on both the relative size of the disutility η_k from being 'misrepresented' as well as the relative marginal utility of consumption in period 2. Clearly, for the symmetric case in which $\eta_\alpha^m = \eta_\beta^m$ and $\omega_\alpha = \omega_\beta$, Eq. (24) simplifies to $\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)} = (c'(C_\alpha^2))^2/(c'(C_\beta^2))^2$, which means that the distortion is relatively larger in the region in which period 2 consumption is smaller. Because C_k^2 is concave in regional income, we conclude that, ceteris paribus, the poorer a region the more extreme will be the federal negotiator which is selected.

It is relevant to compare the equilibrium characterised by Lemma 3 to i) a scenario without federal intervention, as well as ii) the outcome when the median voter is naive (as opposed

to strategic). For that, we rewrite Eq. (21) and Eq. (22) for $k = \alpha$ to arrive at

$$\left| MRS_\alpha^{1,2}(\theta_\alpha^m) \right| = \left| MRS_\alpha^{G,2}(\theta_\alpha^m) \right| = \left(\frac{\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} \right) \tag{25}$$

with $\left| MRS_k^{1,2}(\theta_k^m) \right| \equiv {c'(C_k^1)}/{c'(C_k^2)\theta_k^m}$ being the absolute value of the marginal rate of sub-stitution for the median voter in region k between current and future consumption, and $\left| MRS_k^{G,2}(\theta_k^m) \right| \equiv {g'(G_k)}/{c'(C_k^2)\theta_k^m}$ the marginal rate of substitution for the median voter in re-gion k between the local public good and future consumption. Both expressions combined implicitly capture the opportunity cost of borrowing, which will be a valuable measure for fin-ancial federal support in what follows. Now, in order to compare the equilibrium expressed by Eq. (25) to other scenarios, we need a similar characterisation for the benchmark case without federal co-financing. We therefore first derive the equilibrium marginal rates of substitution under such a scenario in Lemma 4.

Lemma 4. *In a setting where the federal government does not co-finance regional public goods, we have that for any k*

$$\left| MRS_k^{1,2}(\theta_k^m) \right| = \left| MRS_k^{G,2}(\theta_k^m) \right| = 1. \tag{26}$$

Or, when the federal government does not intervene in regional finance, on the margin, the local public good is valued to the same degree as private consumption in both periods.

Proof. See appendix B. □

Lemma 4 will be crucial to our reasoning below, where we compare our earlier results to two scenarios of interest.

Comparison to the scenario without federal intervention.   If we define over-borrowing as a situation where regions borrow more funds from financial markets than they otherwise would have in the absence of federal co-funding, we can characterise the equilibrium defined by Eq. (25) as follows in Proposition 2.

Proposition 2 (Overborrowing). *In a federation where a cooperative legislature of locally elected representatives bargains over federal redistribution, regional governments face a soft budget constraint and over-borrow.*

Proof. See appendix B. □

Comparing when the federal government does, or does not, co-finance regional public goods by use of grants – respectively equations Eq. (25) and Eq. (26) – it is immediate to

notice that $\left|MRS_k^{1,2}(\theta_k^m)\right|$ and $\left|MRS_k^{G,2}(\theta_k^m)\right|$ are smaller in the co-financing equilibrium. This, implicitly, expresses the incentive to over-borrow. To understand this, first notice that $\left|MRS_k^{1,2}(\theta_k^m)\right|$ and $\left|MRS_k^{G,2}(\theta_k^m)\right|$ decrease when consumption respectively of C_k^1 and of G_k increases (at least relative to period 2 consumption C_k^2). Then, the way to obtain a lower equilibrium value is to increase consumption in the first period and consumption of the public good. However, from Eq. (3), the only way to increase consumption in the first period is by reducing taxation while, by Eq. (2), the only way to increase public good consumption is through borrowing or taxing in the first period. Then, the only way to have a decrease in both marginal rates of substitution is to simultaneously decrease taxation (t_k^1) and increase borrowing (B_k). Since the conditions in Eq. (25) and Eq. (26) imply that the marginal rate of substitution must be lower under co-financing, it follows that borrowing must increase.

In other words, current consumption and overall public provision are valued more than future consumption when the federal government intervenes in regional finance. This is because the median voter can anticipate the compensating effect of the simultaneous decision to strategically elect a federal representative who will safeguard future consumption. As shown by Proposition 1, the median voter indeed sends a representative to the federal level with an extreme preference for future consumption, thus reducing the opportunity cost of borrowing as this representative will bargain for federal policies cushioning any drop in future consumption. The logical result will be higher borrowing levels compared to the setting where the federal government would not intervene.

Consequently, and as proven in Proposition 2, the full extent of the soft budget constraint problem is maintained in our setting. Regions not only face a soft budget constraint but also over-borrow because of it. The expectations to receive a grant materialise, and lead to more current consumption and public provision of the local public good. Because of this, consumption in period 2 goes down in both regions, since the borrowing is financed by second-period tax hikes. However, and depending on the relative proportions of the strategic mark-up as characterised by Corollary 1, one region will enjoy a net gain from its strategic vote and receive federal aid – mitigating the drop in future consumption – and vice versa for the other region.

Comparison to the scenario with naive (non-strategic) median voters. We can study the equilibrium under Eq. (25) further by comparing it to the case in which there is federal redistribution but median voters are naive and hence do not strategically distort their choice when voting for the federal representative. Also here, the interesting question remains whether borrowing costs, and thus borrowing decisions, will be different. To size up a situation where the federal government consists of the median voters – sincerely elected in

each region – we simply need to replace $\hat{\theta}_k$ with θ_k^m in Eqs. (21) to (25).

The main difference between the scenario here, in which the median voter is naive, and the previous one in which there was no federal government intervention is obvious, yet relevant. In the previous setting, without federal co-financing, there was no redistributive grant mechanism, which entirely undermined the incentive of median voters to strategically distort their type of representative in order to take full advantage of – or to protect themselves from – federal redistribution. Not just because in this case there was no redistribution to begin with. Also, and more generally, because there was no federal response to decisions made in period 1 – with Lemma 1 no longer holding as a result – it never paid off to strategically influence the type of elected federal politicians. This then had regions operate under a 'hard' budget constraint. Inversely, in this scenario the median voter is naive but still has the federal government responding to local decisions, even though voting is assumed non-strategic. In other words, it is not because voters do not vote strategically that the soft budget constraint ceases to exist. The federal government remains responsive to regional borrowing decisions, as laid out by Lemma 1, so that local governments will still be tempted to over-borrow. Compared to the equilibrium characterised by Eq. (25) however, over-borrowing will be less pronounced as federal politicians now have less extreme – median – preferences θ_k^m over period 2 consumption. We summarise in Proposition 3.

Proposition 3. *When the median voter acts in a naive manner, regional governments over-borrow as compared to the case with no federal intervention, but less so than when the median voter acts strategically.*

Proof. See appendix B.   □

Interestingly, what Proposition 3 suggests is that when regional affiliation – leading to strategic voting – occurs, the soft budget constraint could pose an even more pressing problem than is commonly understood. Compared to a federal government which behaves in the usual Downsian fashion and follows the median voter, a cooperative legislature leads to more over-borrowing.

At first glance, such a result may readily be harnessed as a further critique on vertical fiscal imbalances. On the other hand, and in light of the arguments in favour of imbalances as referred to in our introduction, a more constructive stance is also at hand. From this perspective, Proposition 3 can be seen as adding more weight to recent calls for installing *federation-wide* constituencies. Indeed, when federal politicians are held accountable by all voters of a federation, and not just by a favoured regional fraction, we return to the outcome obtained in a median voter model, or a benevolent government model with representative

consumers. The excess in inefficiency would thus be undone without altering grant policies, which can be welfare enhancing for a variety of other reasons.[15]

Lastly, and for completeness, we investigate the role of risk aversion – i.e. the concavity of utility – with respect to our findings sketched out above.

Risk aversion. The analysis of the result in Eq. (25) can be pushed even further by re-writing it as

$$\left|MRS_\alpha^{1,2}(\theta_\alpha^m)\right| = \left|MRS_\alpha^{G,2}(\theta_\alpha^m)\right| = \frac{A_\beta^2}{A_\alpha^2 + A_\beta^2} \tag{27}$$

where $A_k^2 = -c''(C_k^2)/c'(C_k^2)$ is the Arrow-Pratt measure of concavity in region k for consumption in period 2. Eq. (27) depicts the median voter's decision as a function of the risk aversion in one region relative to the other, or more generally they relate the median voter decision to the concavity of function c in each region, measured at the equilibrium.

Combining Eq. (27) for regions α and β, we then obtain

$$\frac{\left|MRS_\alpha^{1,2}(\theta_\alpha^m)\right|}{\left|MRS_\beta^{1,2}(\theta_\beta^m)\right|} = \frac{\left|MRS_\alpha^{G,2}(\theta_\alpha^m)\right|}{\left|MRS_\beta^{G,2}(\theta_\beta^m)\right|} = \frac{A_\beta^2}{A_\alpha^2}. \tag{28}$$

Eq. (28) suggests that the more concave the $c(C_k^2)$ function in equilibrium, the more agents in region k want to replace future consumption with both current consumption and public spending. Hence, our results regarding regional over-borrowing and strategic voting – in particular Proposition 2, Proposition 3 and Corollary 1 – will be more pronounced as a result.

Soft budgets and gerrymandering Our results show that a multi-layered structure is likely to create the conditions for soft budget issues to take shape, with all the undesired consequences of over-borrowing and over-provision of public goods at lower levels of jurisdictions and – something that was previously unknown – a rise of political extremism at higher levels of government. Several components determine the proportion of this phenomenon: 1) the relative bargaining power of regions within the federal government (ω) is a primitive of the model that can hardly be modified by the legislator; 2) the level of interregional redistribution (γ) depends on the institutional setting and can be modified by the legislator but often it is a politically very costly task; 3) several attributes of the median voter in each jurisdiction: their preferences θ over future consumption or redistribution, the concavity $c(\cdot)$ of their consumption preferences, their income Y and their disutility η from political diversity.

---

[15]See Boadway and Shah (2009) for an elaborate account.

Now, politicians in power often have the authority to, quite arbitrarily, modify the geographical boundaries of jurisdictions. This is done in order to modify the composition of the electorate within a jurisdiction and, hence, enhance the chances of being re-elected. Such practice is known as gerrymandering. By realigning the electoral base, gerrymandering is likely to affect the identity of the median voter who, in turn, is the ultimate cause of the undesired distortions. This leads us to conclude that, on the one hand, gerrymandering could be used strategically to manipulate the extent to which soft budgets affect total welfare. Meanwhile, any gerrymandering decision will, as a side-effect, influence the magnitude of the soft budget distortions. This way, an increase (decrease) of the median income in a net-contributing (-receiving) jurisdiction would increase distortions as well as political extremism.

4   Empirical Analysis: European vs National Elections

4.1   Hypothesis

In this section, we provide evidence on one of our key predictions. In the model, we show how citizens, when casting their vote for federal elections, prefer candidates that have more extreme preferences than their own. Conversely, we do not expect the same strategic behaviour to characterise regional elections, in which case citizens will vote sincerely. An important corollary is that the strategic effect occurs only in net receiving and net contributing regions, but not in regions where contributions are netted out by federal expenditures.

We assess these predictions focusing on elections in the European Union (EU). This is an ideal setting for our model, as i) the EU is the world largest supra-national federation, enveloping the fiscal policy of 28 different countries with respect to justice, home affairs, trade, agriculture and regional development; ii) and EU citizens are called to elect national as well as European delegates, where the former might be interpreted as the regional politicians on our model, and the latter as the federal politicians. While in the model we consider voters' support for parties holding more extreme views, here we consider the aggregate performances of moderate and extreme parties. Specifically, following Proposition 1, our main hypothesis is that voters will anticipate the bargaining over the EU policymaking taking place in Brussels, and consequently favour extreme parties relatively more at the European (federal) than at the national (regional) level.

The European setting thus forms an ideal testing ground for our predictions. From a different perspective, previous studies have also focused on differences in parties performance between national and European elections. The "Second-Order" theory is the most established approach, studying empirical regularities first observed by Reif and Schmitt (1980). This theory has been empirically validated by more recent studies, e.g. Clark and Rohrschneider

19

(2009); Hobolt and Wittrock (2011); Hix and Marsh (2011), and has shown three main broad patterns: i) turnout is lower in European than in national elections; ii) citizens prefer smaller parties at the European level, and iii) they tend to penalise parties leading their respective national government. The theory is based on the idea that European elections represent a second-order election, in which citizens cast their vote based on domestic preferences. This leads to a general punishment of the leading parties, especially when the European elections take place during the mid-term of the national election cycle. In the empirical analysis, we include some covariates to exclude the possibility that our findings might be driven by this Second-Order theory.

A peculiarity of the EU setting is the presence of extreme parties, which in some cases are openly advocating the dissolution of the EU. Therefore, voting for extreme and Eurosceptic parties might be considered as casting a preference for politicians who aim to demolish the European integration project as a whole, rather than bringing about more favourable conditions for their constituencies – which is our interpretation here. This would not be in contrast with our interpretation, if we consider that the most extreme parties are pursuing such a hard (and potentially fatal for the federation) bargaining strategy just to obtain more favourable conditions for their countries. Indeed, many analysts would argue that most Eurosceptics are not out to destroy EU institutions and funding mechanisms, but rather want to turn these to their advantage (Henceroth, 2017; Vasilopoulou, 2011, 2013).

4.2 Data

First, we create a dataset including the vote share of all parties at national and European elections, for parties running in at least one national and one European election.[16] The data are based on the Election and Referendum Database, which provides elections results for European countries, starting from 1990. Second, based on Algan et al. (2017) and the Chapel Hill Expert Survey, we classify parties in moderate or extreme ones. Extreme parties are defined as radical left or radical right ones. While the Chapel Hill Expert Survey is an established source, which estimates party positioning on European integration and ideology[17], Algan et al. (2017) extends it including some brand-new and small parties. As European and national elections very often do not take place in the same year, we consider a time window of 5 years for each political party. Then, we calculate the difference between a party votes share at the European and at the national elections within this period. Therefore, positive (negative) values imply that a party gets higher votes share at the European (national) elections. While European elections take place every 5 years, national elections do not have

---

[16]A clear limitation of this approach is that we cannot detect parties' performances when they enter a coalition, as we only observe the overall coalition votes share. This leads to an overall smaller sample size.

[17]https://www.chesdata.eu/

the same schedule across EU countries. If more than one national election took place in a specific time window, we consider a party average vote share. If a party runs only for one national or one European election in a specific 5-years time period, we would have a missing observation. We consider the first 5-years period starting from the first European election in our dataset (in 1994). Therefore, the periods are 1990-1994; 1995-1999; 2000-2004; 2005-2009; 2010-2013. Different definitions of these 5-years time windows do not affect our findings. In the Appendix, we consider 3-years and 10-years time windows. Our results are very similar to the ones based on a 5-years time window. We prefer the 5-years one as this maximises the number of observations in our sample.

Fig. 1 reports the distribution of this variable (*DiffEU-Nat*). The figure highlights a different distribution for parties considered as extreme, which generally receive a higher vote share in European than in national elections. Figs. 6 and 7 highlight a very similar pattern when considering 3-years and 10-years time windows. To validate this descriptive evidence, we estimate a simple OLS model, in which the dependent variable is the above-mentioned variable (*DiffEU-Nat*) and the main explanatory variable is a dummy set equal to 1 for extreme parties, based on Algan et al. (2017) or the Chapel Hill Expert Survey. To control for differences across countries and common time trends, we include (5-years) time-windows and country fixed effects. In a more demanding specification, we include time-country fixed effects. The Second Order theory might affect these estimates in three ways: i) extreme parties are often small ones, and if small parties are more likely to be voted for at the European level, we might just be capturing this effect. Therefore, we include a variable measuring the average vote share of a party in a specific time window. Note that this might be an endogenous control variable since, if our theory is correct, our model might explain why small parties – if they are also extreme – are rewarded in European elections in the first place; ii) extreme parties might be systematically rewarded or punished if they are always part of the incumbent or in the challenger coalition. To allay this concern, we include a dummy set equal to 1 for the most voted party in each country-time window period; iii) lower turnout at the European elections might differently affect moderate and extreme parties. The inclusion of time-country fixed effects controls for this, as well as for all other country level time-varying changes. These include, for instance, different electoral laws across and within countries, which might differently affect extreme parties.[18]

Figure 1: Differences in Voting between EU and National Elections across Moderate and Extreme Parties
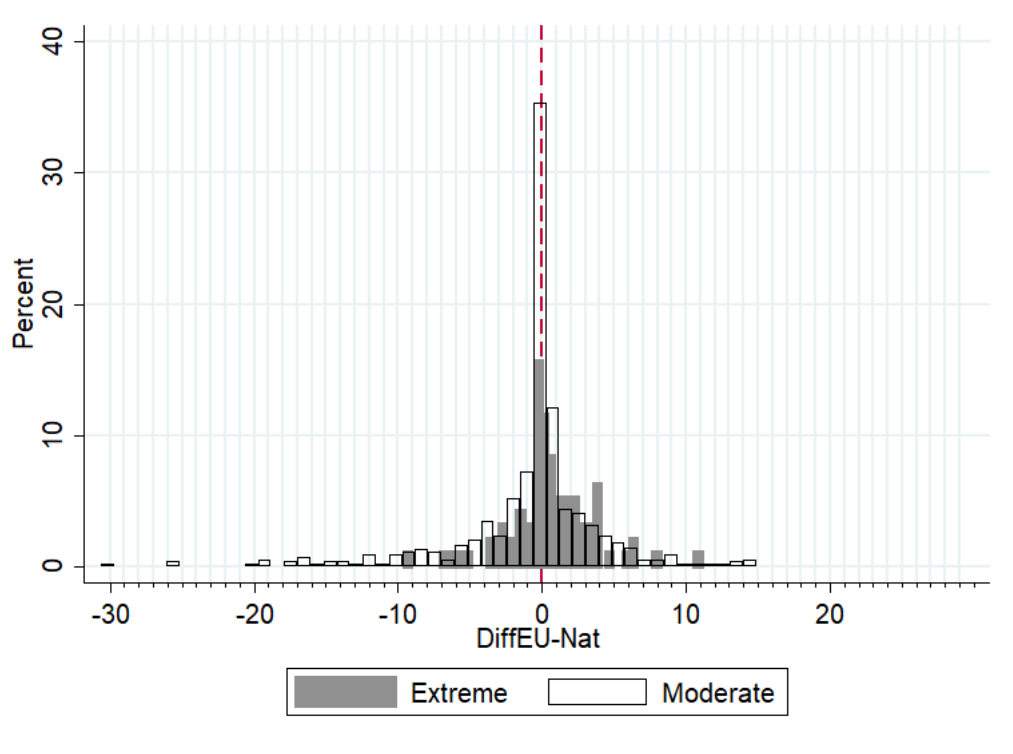
The figure shows the distributions of votes for parties participating in European and National elections in the period 1990-2013, differentiating between moderate and extreme parties.

4.3 Results

We report our findings in Table 1. From Column 1 to 4, we classify extreme parties based on Algan et al. (2017), from Columns 5 to 8 based on the Chapel Hill Expert Survey. In both cases, the first column does not include control variables, which are gradually included over the subsequent columns. The results strongly suggest that extreme parties are more likely to be elected to European parliament than to its national counterparts. This effect is sizable (about 30% of the standard deviation) and statistically significant across all models. The coefficient is reduced only when including the full set of controls, which represents a very demanding specification considering the limited sample size. Specifically, the estimated effect is reduced when controlling for party size. However, as explained above, this could be an endogenous control. Table 5 in the Appendix considers a winsorised version at 99% of the dependent variable, while Tables 6 and 7 test the same specification on a 3-years and a 10-years time windows. In all cases, we find similar results. Still in the Appendix,

---

[18]Note also that i) national and European electoral laws can differ within a country and, ii) most EU countries, in the period of interest, held European elections under some form of proportional representation.

Table 1: Extreme Voting in EU vs National Elections

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme (Algan et. al)	1.277***	1.281***	1.330***	0.758*				
	(0.365)	(0.366)	(0.461)	(0.429)				
Party Size				-0.111***				-0.113***
				(0.0341)				(0.0341)
Leading Party				-0.144				-0.113
				(1.310)				(1.310)
Extreme (Chapel Hill)					1.302***	1.309***	1.269***	0.949**
					(0.406)	(0.407)	(0.487)	(0.467)
Observations	639	639	639	639	639	639	639	639
R-squared	0.009	0.009	0.070	0.157	0.008	0.008	0.068	0.158
Country FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in party vote shares between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy set equal to 1 for extreme parties; *Party Size* is a continuous variable measuring party vote share; *Leading Party* is a dummy set equal to one for the main incumbent party. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Eurosceptic Voting at EU vs National Elections

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Eurosceptic	1.722***	1.734***	1.891***	1.534***	1.849**	2.512***	2.098***
	(0.410)	(0.412)	(0.490)	(0.473)	(0.852)	(0.774)	(0.708)
Party Size				-0.109***			
				(0.0338)			
Leading Party				-0.129			
				(1.301)			
Eurosceptic*2^{nd}tertile					-1.920*	-2.578**	-2.725***
					(1.095)	(1.044)	(1.027)
Eurosceptic*3^{rd}tertile					-0.0712	-0.235	0.0777
					(1.076)	(1.180)	(1.175)
Observations	639	639	639	639	639	639	639
R-squared	0.017	0.018	0.081	0.166	0.031	0.087	0.173
Country FE	NO	NO	YES	YES	NO	YES	YES
Time Fe	NO	YES	YES	YES	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	YES
Maj*Eurosk	NO	NO	NO	NO	YES	NO	NO

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in party vote shares between European and National elections); *Eurosceptic* is a dummy set equal to 1 for Eurosceptic parties; *Party Size* is a continuous variable measuring party vote share; *Leading Party* is a dummy set equal to one for the main incumbent party. *Tertiles* split countries in three groups based on their level of net contributions to the EU budget. The omitted category is the first tertile (largest net contributors). Maj*Eurosk includes a set of interactions between *Tertiles* and countries holding a mixed or majoritarian electoral system. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8 highlights how the effect similarly arises for extreme left and right parties (based on the classification used by Algan et al. (2017)). Finally, Table 2 (Columns 1 to 4) focuses on a different definition of extreme parties, including only Eurosceptic parties. While not

all Eurosceptic parties are extreme ones (only 50%), most extreme parties are considered Eurosceptic ones (94%). Therefore, this variable includes a wider group of parties, which might still capture voters' preferences for parties aiming at challenging the status quo at the European level. This definition is included in Algan et al. (2017). Also in this case, our findings are confirmed.

The EU budget, which in 2015 was 145 billion Euros, represents a crucial source of financing for the poorest EU members, as well as for firms in several economic sectors (e.g. energy and agriculture) across all EU countries. Similar to our model specifications, some countries – such as Germany, the Netherlands and Sweden – are net contributors to the EU budget, others are net receiving members, and still others contribute as much as they receive. For instance, in the period 2000-2015, the net transfers received from the EU represented 3,53% of the GNI for Lithuania, 2,9% for Bulgaria and 2,11% for Poland.[19] Conversely, based on this measure, the Netherlands was the main EU net contributor (-0.41% of GNI). Following the model, we should expect a stronger bias toward extreme voting at the European elections, especially in net receiving but also net contributing member states, and less so in countries contributing about as much as they receive. To test this prediction, we consider the share of EU net contributions as a percentage of the gross national income (GNI) in the period 2000-2015. Fig. 2 shows these data. We then split countries in three groups, in which the top 33% includes the most generous net-contributors, as the Netherlands, Germany and Sweden and the bottom one the most net receiving countries. Interestingly, the split in tertiles allows for a U-shape test, as we expect a stronger effect in the top (net-contributors) and in the bottom tertile (net-receiving), while strategic extreme voting should matter less in countries in which contributions (to the EU) are, more or less, equal to the transfers (from the EU).

In Table 3, we interact the *Extreme* dummy with this categorical variable. The coefficient *Extreme* represents the effect of the omitted base category, i.e. the top 33% net contributors countries. In column 2 and 5, we also control for the interaction between *Extreme* and countries holding a mixed or majoritarian electoral system at the national level (instead of a proportional one). This classification is based on the International Institute for Democracy and Electoral Assistance (International IDEA). Conversely, at the European level, all countries hold a proportional system with national specificities.[20] Column 1, 2 and 3 in Table 3

---

[19]These data are available on the website of the European Commission: `http://ec.europa.eu/budget/financialreport/2015/revenue/index_en.html`

[20]In mixed or majoritarian systems, citizens might be more likely to strategically cast their vote at the national level, in turn, favouring centrist parties. Conversely, they might be more likely to express their ideological preferences at the European elections, based on a proportional electoral system. In this case, extreme parties might be favoured. Note that this specification is highly demanding as four out of six countries holding (at least for some years) a mixed or majoritarian system are in the first thirdile, i.e. France, Germany, Italy and United Kingdom. The other two are Hungary and Lithuania.

Figure 2: Net Transfers *from* the EU as % of GNI (2000-2015)
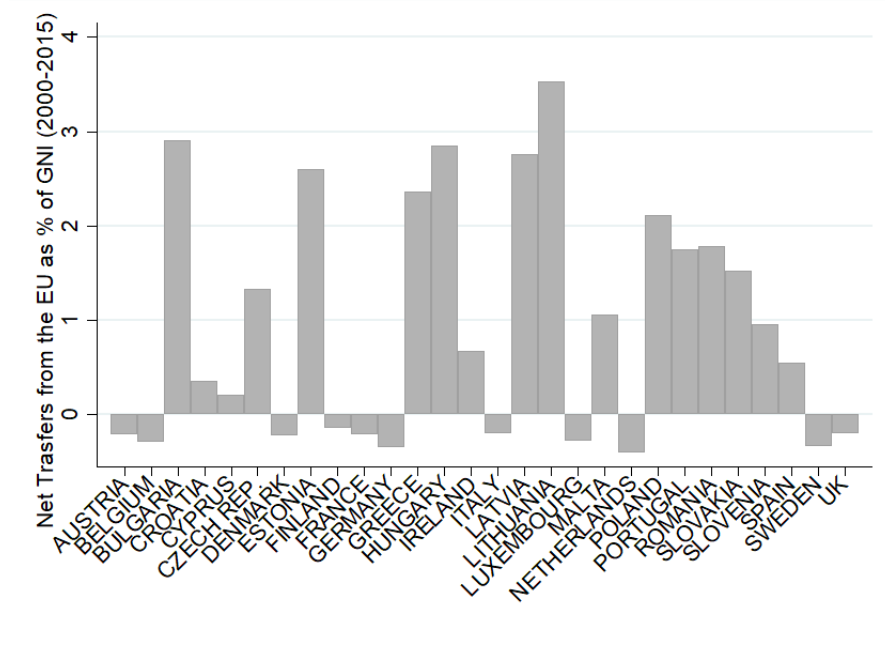
The figure shows the net transfers received from the EU budget as % of GNI in the period 2000-2015 for each EU country.

confirm our predictions. The effect materialises only in the top and the bottom tertile (line 1 and 3), while there is not such effect for extreme parties in the intermediate tertile (line 2): these countries represent an ideal control group, as they receive more or less as much from the EU budget as they contribute to it. Although we would expect a stronger effect in the third tertile (top net receiving countries), we observe a slightly higher effect among the top contributors (first tertile): this difference, however, is very far from being statistically significant. A similar pattern emerges from Columns 4, 5 and 6, which focuses on the Chapel Hill classification: in this case, however, the differences across tertiles weakly miss statistically significant levels. [21,22] These findings are also confirmed when looking at Eurosceptic parties, as shown in Columns 5, 6 and 7 of Table 2. Overall, extreme parties appear more successful on the European political arena, especially in countries in which the EU budget is more salient, such as in (highly) net receiving and in net contributing ones.[23]

---

[21]In the Appendix, Table 9 highlights similar findings when grouping countries based on their absolute average contribution in the period 2000-2015 (in this case, Germany is the top net-contributor with a yearly contribution of almost 9 billion Euros).

[22]This heterogeneous effect similarly arises for extreme left parties both in net receiving and contributing countries. Conversely, the effect is stronger for extreme right parties in net contributing countries. This could be explained by the fact that extreme right parties are in favour of downsizing the EU budget, a policy more likely to be supported by citizens in countries subsidising the EU project.

[23]In Appendix Section D, we investigate whether these effects depend on the size of the federation, focusing on the EU enlargement in 2004.

Table 3: Extreme Voting at EU vs National Elections: Net Contributors and Net Receivers

	(1)	(2)	(3)	(4)	(5)	(6)
Extreme (Algan et al.)	1.995***	1.229*	1.451**			
	(0.762)	(0.731)	(0.716)			
Extreme Alg*2.third_net	-2.214**	-1.511	-2.276**			
	(1.046)	(1.021)	(1.071)			
Extreme Alg*3.third_net	-0.641	0.0357	-0.680			
	(1.035)	(0.897)	(1.041)			
Extreme (Chapel Hill)				1.653**	0.572	1.260*
				(0.767)	(0.844)	(0.725)
Extreme CH*2^{nd}tertile				-1.884	-0.805	-1.685
				(1.284)	(1.300)	(1.331)
Extreme CH*3^{rd}tertile				-0.191	1.037	-0.0362
				(0.978)	(1.034)	(0.972)
Observations	639	639	639	639	639	639
R-squared	0.074	0.021	0.160	0.070	0.021	0.159
Country FE	YES	NO	YES	YES	NO	YES
Time Fe	YES	YES	YES	YES	YES	YES
Country-Time FE	NO	NO	YES	NO	NO	YES
Controls	NO	NO	YES	NO	NO	YES
Maj*Extr	NO	YES	NO	NO	YES	NO

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in party vote shares between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy set equal to 1 for extreme parties; *Tertiles* split countries in three groups based on their level of net EU contributions. The omitted category is the first tertile (largest net contributors). Maj*Eurosk includes a set of interactions between *Tertiles* and countries holding a mixed or majoritarian electoral system. *** p<0.01, ** p<0.05, * p<0.1.

5   Conclusion

Incentives play a major role when it comes to economic decision-making, as a badly designed incentive scheme may lead to undesired outcomes. This paper shows that in a multi-tiered country – as well as in a union of countries such as the EU – the presence of interregional redistribution or bailout schemes has firstly, and quite intuitively, the consequence that local jurisdictions adjust their behaviour. Levels of inter-temporal consumption, public good provision and borrowing are determined in such a way as to take advantage of the federal system. Whether they are net recipients or contributors, maximising or minimising the costs of interregional redistribution will drive local decisions. Either way, local governments will over-borrow when a transfer scheme is in place, as shown in Proposition 3.

Furthermore, and more interestingly, when voters are sophisticated, they anticipate the bargaining process that leads to the concession of grants or bailouts by the federal level. As a result, when they select their local representative within the federal government, Proposition 1 shows that they manipulate their choice in order to move the bargaining point in a more favourable direction. To do that, they vote for federal candidates that have more extreme preferences than their own. In other words, median voters in each region willingly elect delegates with preferences that are more extreme than their own. This is a strategic

choice, reflected by the preference type of the elected federal representatives. Net receiving regions will elect delegates which are more likely to condone profligate spending by lower-level governments, whilst net contributors are more likely to elect fiscal hawks. This kind of strategic voting is more pronounced in poorer regions. Such behaviour has direct and testable consequences on the level of polarisation and extremism. Our empirical analysis (section Section 4) – consistent with the model's predictions – shows that citizens in the European Union have voted for more extreme parties in EU elections than they did for their own national elections. This strategic effect arises only for net receiving and net contributing member states, but not in countries where EU contributions and expenditures are more or less balanced.

Over-borrowing by regions, as described in Proposition 2, is the result of two distortions: i) the redistribution scheme per-se, which induces over-borrowing, and ii) the strategic behaviour of regions, when they select their federal delegates, which induces the redistribution scheme to be overly generous, reinforcing the direct distortion.

We initially show these dynamics in a simplified setting with only two regions and a grant system in place, where we show that results are reinforced the more agents are risk-averse. In Sections A.1 and A.2 we extend our results in a few directions. Firstly, we show that results hold when we move to a 3-region setting in which one region is excluded from the ruling coalition at the federal level. Secondly, we prove that results would attenuate and yet remain valid when voters expect their representative to influence the federal policies only with some probability. This would occur, for example, when the process of forming a coalition at the federal level implies some randomness that may cause regions to be sometimes excluded from the coalition. Such a specific setting is formalised in Section A.3 of the appendix. Section A.4, instead, shows that the distortions and inefficiencies related to soft budgets follow from the redistribution schemes in place, regardless whether these emerge through a continuous grant-based transfer, or through one-off rescues defined by a bailout fund. Indeed, Section A.4 modifies some core elements of the baseline model, in order to allow it to be interpreted as a bailout scheme. We are able to show that the two models converge to the same results.

References

Akai, N., Sato, M., 2008. Too big or too small? a synthetic view of the commitment problem of interregional transfers. Journal of Urban Economics 64, 551 – 559.

Akai, N., Sato, M., 2011. A simple dynamic decentralized leadership model with private savings and local borrowing regulation. Journal of Urban Economics 70, 15 – 24.

Algan, Y., Guriev, S., Papaioannou, E., Passari, E., 2017. The european trust crisis and the rise of populism. Brookings Papers on Economic Activity 2017, 309–400.

Andeweg, R.B., 2000. Consociational democracy. Annual Review of Political Science 3, 509–536.

Baskaran, T., 2012. Soft budget constraints and strategic interactions in subnational borrowing: Evidence from the german states, 1975–2005. Journal of Urban Economics 71, 114–127.

Besfamille, M., Lockwood, B., 2008. Bailouts in federations: Is a hard budget constraint always best? International Economic Review 49, 577–593.

Besley, T., Coate, S., 2003. Centralized versus decentralized provision of local public goods: a political economy approach. Journal of Public Economics 87, 2611–2637.

Boadway, R., Shah, A., 2009. Fiscal federalism: principles and practice of multiorder governance. Cambridge University Press.

Bordignon, M., Turati, G., 2009. Bailing out expectations and public health expenditure. Journal of Health Economics 28, 305–321.

Borge, L.E., Rattsø, J., 2003. Spending growth with vertical fiscal imbalance: decentralized government spending in norway, 1880–1990. Economics & Politics 14, 351–373.

Breuillé, M.L., Vigneault, M., 2010. Overlapping soft budget constraints. Journal of Urban Economics 67, 259 – 269.

Buettner, T., Wildasin, D.E., 2006. The dynamics of municipal fiscal adjustment. Journal of Public Economics 90, 1115–1132.

Caplan, A.J., Cornes, R.C., Silva, E.C., 2000. Pure public goods and income redistribution in a federation with decentralized leadership and imperfect labor mobility. Journal of Public Economics 77, 265 – 284.

Christiansen, N., 2013. Strategic delegation in a legislative bargaining model with pork and public goods. Journal of Public Economics 97, 217–229.

Clark, N., Rohrschneider, R., 2009. Second-order elections versus first-order thinking: How voters perceive the representation process in a multi-layered system of governance. European Integration 31, 645–664.

Crivelli, E., Staal, K., 2013. Size, spillovers and soft budget constraints. International Tax and Public Finance 20, 338–356.

Duchacek, I.D., 1985. Consociational cradle of federalism. Publius: The Journal of Federalism 15, 35–48.

Eerola, E., Määttänen, N., Poutvaara, P., 2004. Citizens should vote on secession. Topics in Economic Analysis & Policy 4.

Enikolopov, R., Zhuravskaya, E., 2007. Decentralization and political institutions. Journal of public economics 91, 2261–2290.

Eyraud, L., Lusinyan, L., 2013. Vertical fiscal imbalances and fiscal performance in advanced economies. Journal of Monetary Economics 60, 571–587.

Faguet, J.P., 2004. Does decentralization increase government responsiveness to local needs?: Evidence from Bolivia. Journal of Public Economics 88, 867–893.

Gans, J.S., Smart, M., 1996. Majority voting with single-crossing preferences. Journal of Public Economics 59, 219–237.

Gavazza, A., Lizzeri, A., 2009. Transparency and economic policy. The Review of Economic Studies 76, 1023–1048.

Goodspeed, T.J., 2002. Bailouts in a federation. International Tax and Public Finance 9, 409–421.

Harstad, B., 2010. Strategic delegation and voting rules. Journal of Public Economics 94, 102–113.

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

Henceroth, N.W., 2017. The engaged eurosceptic: Explaining eurosceptic success and failures in european parliamentary elections .

Hix, S., Marsh, M., 2011. Second-order effects plus pan-european political swings: An analysis of european parliament elections across time. Electoral Studies 30, 4–15.

Hobolt, S.B., Wittrock, J., 2011. The second-order election model revisited: An experimental test of vote choices in european parliament elections. Electoral Studies 30, 29–40.

Jones, M.P., Sanguinetti, P., Tommasi, M., 1999. Politics, institutions, and public-sector spending in the argentine provinces, in: Fiscal Institutions and Fiscal Performance. University of Chicago Press, pp. 135–150.

Knight, B., 2008. Legislative representation, bargaining power and the distribution of federal funds: evidence from the us congress. Economic Journal , 1785–1803.

Kornai, J., Maskin, E., Roland, G., 2003. Understanding the soft budget constraint. Journal of economic literature 41, 1095–1136.

Köthenbürger, M., 2004. Tax competition in a fiscal union with decentralized leadership. Journal of Urban Economics 55, 498–513.

Krasa, S., Polborn, M., 2015. Political competition in legislative elections. University of Illinois .

Lijphart, A., 1985. Non-majoritarian democracy: A comparison of federal and consociational theories. Publius: The Journal of Federalism 15, 3–15.

Mattozzi, A., Snowberg, E., 2018. The right type of legislator: A theory of taxation and representation. Journal of Public Economics 159, 54–65.

Pettersson-Lidbom, P., 2010. Dynamic commitment and the soft budget constraint: an empirical test. American Economic Journal: Economic Policy 2, 154–79.

Poterba, J.M., 1995. Balanced budget rules and fiscal policy: Evidence from the states. National Tax Journal 48, 329–329.

Qian, Y., Roland, G., 1998. Federalism and the Soft Budget Constraint. American Economic Review 88, 1143–1162.

Reif, K., Schmitt, H., 1980. Nine second-order national elections–a conceptual framework for the analysis of european election results. European journal of political research 8, 3–44.

Robinson, J.A., Torvik, R., 2009. A political economy theory of the soft budget constraint. European Economic Review 53, 786–798.

Rodden, J., 2002. The dilemma of fiscal federalism: grants and fiscal performance around the world. American Journal of Political Science , 670–687.

Rodden, J., Eskeland, G.S., Litvack, J.I., 2003. Fiscal decentralization and the challenge of hard budget constraints. MIT press.

Sanguinetti, P., Tommasi, M., 2004. Intergovernmental transfers and fiscal behavior insurance versus aggregate discipline. Journal of International Economics 62, 149–170.

Sorribas-Navarro, P., 2011. Bailouts in a fiscal federal system: Evidence from spain. European Journal of Political Economy 27, 154–170.

Thorlakson, L., 2009. Patterns of party integration, influence and autonomy in seven federations. Party Politics 15, 157–177.

Vasilopoulou, S., 2011. European integration and the radical right: Three patterns of opposition. Government and Opposition 46, 223–244.

Vasilopoulou, S., 2013. Continuity and change in the study of euroscepticism: Plus ça change? JCMS: Journal of Common Market Studies 51, 153–168.

Volden, C., Wiseman, A.E., 2007. Bargaining in legislatures over particularistic and collective goods. American Political science review 101, 79–92.

Von Hagen, J., 1991. A note on the empirical effectiveness of formal fiscal restraints. Journal of Public Economics 44, 199–210.

Von Hagen, J., Dahlberg, M., 2004. Swedish local government: Is there a bailout problem?. Springer. pp. 47–76.

Wildasin, D.E., 1999. Externalities and Bailouts: Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations. The World Bank.

World Bank, 2000. Entering the 21st century: World development report 1999/2000.

Appendix A   Extensions

A.1   Strategic behaviour and beliefs

Until now we assumed that elected delegates have a certain say in federal negotiations, and that agents are fully rational, in the sense that they perfectly understand the institutional setting and act strategically, thus taking advantage of the system as much as possible. We could relax both assumptions by assuming that agents assign a region-specific probability p_k to the fact that their elected politician will be able to impact federal policy-making, and thus the federal tax t_f as previously described, whereas with probability $1 - p_k$ he or she is not influencing the federal government's decisions at all. In this latter case the federal tax will consequently be exogenously set at $\hat{t}_f$.[24]

There can be several ways to rationalise such beliefs. Voters may think – or correctly anticipate – that $1 - p_k$ is the probability that a coalition forms, but with their own representative kept out of the formation process.[25] Another possibility could be that voters believe their elected politicians are not the ones who have actual discretion over t_f, but rather some other (technocratic) institutions or political forces over which voters have less control. One more option could be that voters believe the political weight of their region within the coalition is such that their politician is not able to influence t_f, or at least to a lesser degree. This could be the case when their own politicians are less ideally positioned to pull the levers of policy-making because they have less insight in the various practices and rules involved. It could also be that the latter are perceived as potentially stacked in favour of other representatives enjoying more leverage when the rules are set. Lastly, the general opinion could be that once regional representatives are part of the national, higher-level establishment, they will lose interest in their own region, or will be pushed to do so.[26] Clearly, one may think that p_k – as interpreted above – could also vary through time in a more general dynamic setting. In order to keep our robustness check in this section distinct and clear, however, we keep the model static and study how results change when p_k varies.

Conditional on the politician being able to influence t_f, and still solving backwards, results from period 2 in Section 3.1.1 go through in our extended setting here, as voter beliefs do not directly affect actual federal decision-making. However, and moving to period 1 when the federal representatives are elected, the maximisation problem of the median voter in region

---

[24]Notice that at this point it does not matter whether these beliefs originate from voters not fully understanding the game at play, or whether they correctly anticipate that the local politician will not always have a say on the decision of t_f.

[25]We work out this micro-foundation of beliefs in appendix Section A.3, where voters have to make conjectures for the likelihood of their representative being the formateur, who shapes the federal minimum winning coalition. We show that all possible combinations collapse into a formulation identical to Eq. (112), where p_k relates to endogenous federal decision-making and $1 - p_k$ to a setting where t_f is exogenously set ($\hat{t}_f$).

[26]See e.g. Thorlakson (2009) and the references therein.

$k \in K$ now includes the additional constraint that t_f is defined by Eq. (17) with probability p_k, and $t_f = \hat{t}_f$ with probability $1 - p_k$. Here then, Lemma 3 translates into the following lemma.

Lemma 5. *Eqs. (29) to (31) implicitly define the representative in the federal legislature $(\hat{\theta}_\alpha)$, the preferred quantity of the public good (G_α) and the local tax (t_α^1) preferred by the median voter in region α.*

$$\frac{-(1 - p_k)\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) - \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} = -\frac{c'\left(C_\alpha^1\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)} \tag{29}$$

$$\frac{-(1 - p_k)\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) - \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} = -\frac{g'\left(G_\alpha\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)} \tag{30}$$

$$\frac{-p_k \omega_\alpha c'\left(C_\alpha^2\right)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} = \frac{2\eta_\alpha^m \left(\hat{\theta}_\alpha - \theta_\alpha^m\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)}. \tag{31}$$

From Lemma 5, we immediately obtain that Proposition 1 still holds. Eq. (31) is identical to Eq. (23), with the exception that the left hand side of Eq. (31) is multiplied by $p_k \in [0, 1]$, from which we immediately conclude that the equilibrium distortion of θ_k is lower in this case. This result is quite intuitive, since now agents expect the strategic mechanism to be at work only with probability p_k, so that they have less reason to manipulate the federal vote, and reduce distortion as a result. Comparing these distortion across regions subsequently, we obtain Corollary 3. This is the equivalent of Corollary 1, but then in the more general setting allowing for varying voter beliefs.

Corollary 3. *The strategic voting of the median voter – assigning beliefs to the effectiveness of federal representation – is characterised by*

$$\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)} = \frac{p_\alpha}{p_\beta} \frac{\omega_\alpha}{\omega_\beta} \frac{\left(c'(C_\alpha^2)\right)^2 / \eta_\alpha^m}{\left(c'(C_\beta^2)\right)^2 / \eta_\beta^m}. \tag{32}$$

Corollary 3 sheds more light on how beliefs captured by p_k affect our previous results. The more representatives of a certain region are seen as being influential as compared to another, the more voters will distort federal elections in that region. In other words, the more pronounced the beliefs captured by p_α as compared to p_β, the higher the relative mark-up of distortion $\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)}$ in Corollary 3. Under the same assumptions as those made in Section A.2, the same reasoning applies to Lemma 6 and Proposition 6, describing the strategic distortions in the case of three regions.

Next, and combining Eq. (29) with Eq. (30), we arrive at Proposition 4 which is the equivalent of Proposition 2, but then in our beliefs-extended setting here.

33

Proposition 4 (Overborrowing)**.** *In a federation where a cooperative legislature of locally elected representatives bargains over federal grants, but where voters have varying beliefs over the effectiveness of federal representation, local governments face a soft budget constraint. Compared to a situation where the federal government does not co-finance regional public provision, overborrowing is more pronounced.*

Proof. See appendix B. □

It comes at no surprise that – as can be appreciated from the proof at Eqs. (124) to (125) – the distortion in terms of overborrowing is mitigated in this setting, as compared to the result in Section 3.1.2. The intuition here is again that voters are weighing their behaviour against the cases in which distortions are effective, and the cases in which they are not. Again, and under the same assumptions as those made in Section A.2, the same reasoning applies to a setting with three regions. Comparing such a setting to the overborrowing occurring when only two regions make up the federation lastly, we obtain Proposition 5, which is identical to Proposition 7.

Proposition 5 (Overborrowing - 3 regions)**.** *When voters assign varying beliefs to the effectiveness of federal decision-making, and if regions k and z are of the same type (both either net-recipients or both net-contributors), region k over-borrows less than in the 2-region setting. Conversely, if k and z are of opposite types (k is a recipient/contributor and z is a contributor/recipient), then region k over-borrows more than in the 2-region setting.*

Proof. See appendix B. □

A.2   Three-region setting

In this section, we extend the analysis by adding one region to the federal constellation studied above, hence $|R| = 3$. We restrict our focus to the case in which a coalition of two forms: $|K| = 2$. We leave aside the analysis of the equilibrium when the grand coalition forms, because it is isomorphic to what we observe with $|R| = 2$ and hence its study does not bring new insights. We abstract from the discussion or formal analysis of why a given coalition forms between two of the three regions.[27] Following the previous notation and without loss of generality $K = \{\alpha, \beta\} \subset R$. Furthermore, we denote by $z \in R \backslash K$ the region that is not represented in the coalition.

We focus on the case in which the coalition is formed by a net-contributor and a net-recipient,[28] and neglect the trivial case of a coalition of regions of the same kind, which

---

[27]We briefly come back to this point in Section A.1 and Section A.3, where we test the robustness of our model against that.

[28]Region z, despite being left outside the coalition, is part of the federal state and as such they participate

inevitably leads to a corner equilibrium.[29] By construction, if the two regions in the coalition are of opposite type, one of them must be of the same type as region z.

The maximisation problem of the federal government is

$$\max_{t_f} \sum_{k \in K} \left(\omega_k \hat{U}_k - F_k \right) = \sum_{k \in K} \omega_k \left(g\left(G_k\right) + c\left(C_k^1\right) + \hat{\theta}_k c\left(C_k^2\right) - F_k \right), \tag{33}$$

and the first order condition is

$$\tilde{\phi}_f = 0$$
$$\text{with } \tilde{\phi}_f \equiv \omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c'\left(C_\alpha^2\right) + \omega_\beta \hat{\theta}_\beta \Delta_\beta c'\left(C_\beta^2\right). \tag{34}$$

Eq. (34) is the counterpart of Eq. (16). The main difference between them is that with two regions $\Delta_\alpha = -\Delta_\beta$ while in the 3-region case $\Delta_\alpha = -\Delta_\beta - \Delta_z$.

From Eq. (34) follows that

$$\frac{\partial t_f}{\partial G_\alpha} = \frac{\omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c''(C_\alpha^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)}, \tag{35}$$

$$\frac{\partial t_f}{\partial t_\alpha^1} = -Y_\alpha \frac{\omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c''(C_\alpha^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)}, \tag{36}$$

$$\frac{\partial t_f}{\partial \hat{\theta}_\alpha} = -\frac{\omega_\alpha \Delta_\alpha c'(C_\alpha^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)}, \tag{37}$$

Lemma 6 compares results in Eqs. (35) to (37) with the corresponding ones for the 2-region case, that is, Eqs. (18) to (20).

Lemma 6. *When region z (the one out of the coalition) is a net-contributor, the marginal impact (in absolute value) of G_k, t_k^1 and $\hat{\theta}_k$ on t_f is larger in the 3-region setting than in the 2-region one. Viceversa, when z is a net-recipient the marginal effect is smaller.*

Lemma 6 shows that the presence of a third region outside the government affects the strategic behaviour of the ones that belong to the coalition. When region z is a contributor, the two regions in the federal government somehow free-ride on z's contribution. Therefore, the federal government over-reacts to any change in $G_k, t_k'1, \hat{\theta}_k$. On the opposite, the federal coalition anticipates that, if z is a recipient, there is a leakage (or spill-over) effect towards z and, therefore, is more conservative and reacts less to changes in the fundamentals of the model that α and β set in period 1.

---

in the process of collection and redistribution of federal taxes even if they are not participating to the federal decision process.

[29] When both regions in the coalition are net-recipients, they set the largest possible federal tax and extract all surplus from the excluded region. When both regions are net-contributors, they repeal the transfer mechanism and set the transfer to the lowest possible.

Moving to local elections, the objective function of each region k in the government is

$$\max_{t_k^1, G_k, \hat{\theta}_k} U_k = g\left(G_k\right) + c\left(C_k^1\right) + \theta_k^m c\left(C_k^2\right) - \eta_k \left(\hat{\theta}_k - \theta_k^m\right)^2, \tag{38}$$

subject to Eqs. (13), (14) and (34).

The first order conditions for region α are

$$-\frac{\omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)} = -\frac{c'\left(C_\alpha^1\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)}. \tag{39}$$

$$-\frac{\omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)} = -\frac{g'\left(G_\alpha\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)}. \tag{40}$$

$$-\frac{\omega_\beta \left(\Delta_\alpha\right)^2 c'\left(C_\alpha^2\right)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)} = \frac{2\eta_\alpha^m \left(\hat{\theta}_\alpha - \theta_\alpha^m\right)}{\theta_\alpha^m c'\left(C_\alpha^2\right)}. \tag{41}$$

The right hand side of Eqs. (39) to (41) is the marginal rate of substitution respectively between $\{C_\alpha^1, G_\alpha, M_\alpha\}$ and C_α^2. From the previous argument, a decrease in the optimal MRS implies an increase in distortion and finally more borrowing.

Rearranging Eq. (41) and taking its ratio for region α over β, we obtain the corresponding of Eq. (24):

$$\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)} = \frac{\omega_\alpha}{\omega_\beta} \frac{\left(c'(C_\alpha^2)\right)^2/\eta_\alpha^m}{\left(c'(C_\beta^2)\right)^2/\eta_\beta^m} \frac{(\Delta_\alpha)^2}{(\Delta_\beta)^2}, \tag{42}$$

Proposition 6. *In the 3-region minimal coalition, the distortion of $\hat{\theta}$ in region α relative to region β is*

(a) *the same as in the 2-region setting, when region z is neither a recipient nor a contributor* $\left(\Delta_z = 0, \text{ and } (\Delta_\alpha)^2 = (\Delta_\beta)^2\right)$.

(b) *smaller than in the 2-region setting, when regions α and z are either both net-recipients or both net-contributors* $\left((\Delta_\alpha)^2 < (\Delta_\beta)^2\right)$.

(c) *larger than in the 2-region setting, when region α is a recipient (contributor) and region z is a contributor (recipient)* $\left((\Delta_\alpha)^2 > (\Delta_\beta)^2\right)$.

Proof. See appendix B. □

Proposition 6 describes basically three situations. One in which region z (the one that remains out of the coalition) is contributing by the same amount they receive, that is $\Delta_z = 0$. In such case, no distortion appears compared to the 2-region case. This is intuitive: the third

region plays no role in the redistribution game, hence their presence do not affect the strategic behaviour of other players.

The more interesting case is when region z is either a net-recipient of transfers, or a net-contributor. Then, the type of regions does not matter: what is important to compare the level of distortion in the 2-region case with the one in the 3-region setting is whether a region is of the same type as region z or of opposite type. In particular, comparing the ratio of distortions (or mark-ups) $\mu(\theta_\alpha)/\mu(\theta_\beta)$ for the 2- and 3-region settings, such ratio increases if region α is of type opposite to region z and decreases otherwise.

In other words, when α and z are of the same type, the relative distortion of $\hat{\theta}$ by region α as measured by $\frac{\mu(\theta_\alpha)}{\mu(\theta_\beta)}$ is smaller with three regions than with two of them. Clearly, both regions in the coalition face a trade-off between consumption C_α^2 in the second period and distorting $\hat{\theta}$, which is costly through the M component in the utility function. Indeed, more distortion implies a larger grant (smaller contribution) for a net-recipient (net-contributor). However, when regions α and z are of the same type, the cost of distortion is borne by region α alone while the benefit is shared with z.[30] Compared to the two-region case in which region α was able to fully enjoy the increase in consumption, now the environment is less favourable and it justifies a balance in the trade-off with less distortion of $\hat{\theta}_\alpha$.

When α and z are of opposite type, the trade-off faced by region α is the same as in the 2-region setting and hence there is no direct effect of a switch from 2 to 3 regions. However, since α and β are always of opposite type, it must be that z and β are of the same type. Hence, by the previous reasoning, β will distort less then in the 2-region case. As a consequence, the ratio of distortions increases. Region α has the same benefit as stake, but is now facing a softer negotiator. This, of course, may lead to softening of α's best response to β, but the change never matches (in magnitude) the change in β's behaviour.

Figs. 3 to 5 represent the content of Proposition 6 by showing when $\frac{(\Delta_\alpha)^2}{(\Delta_\beta)^2} > 1$ from different space perspectives.

White areas represent combinations of Δ_r such that regions α and z are both of the same type (recipient or contributor) and therefore the distortion of $\hat{\theta}_\alpha$ is smaller in the 3-region setting (case (b) in Proposition 6). Conversely grey areas represent combinations of Δ_r such that regions α and z are of opposite type and therefore the distortion of $\hat{\theta}_\alpha$ is larger in the 3-region setting (case (c) in Proposition 6). Striped areas indicate combinations of Δ_r that are incompatible with the federal budget constraint.

Figs. 3 and 4 mirror each other. This is because by construction $\text{sgn}(\Delta_\alpha) = -\text{sgn}(\Delta_\beta)$. Fig. 5 allows to identify in the $\Delta_\alpha, \Delta_\beta$ space which are the areas (white) in which α and z

---

[30]The extent to which α and z share the benefit from a change in t_f depend on which of the two receives/contributes more.
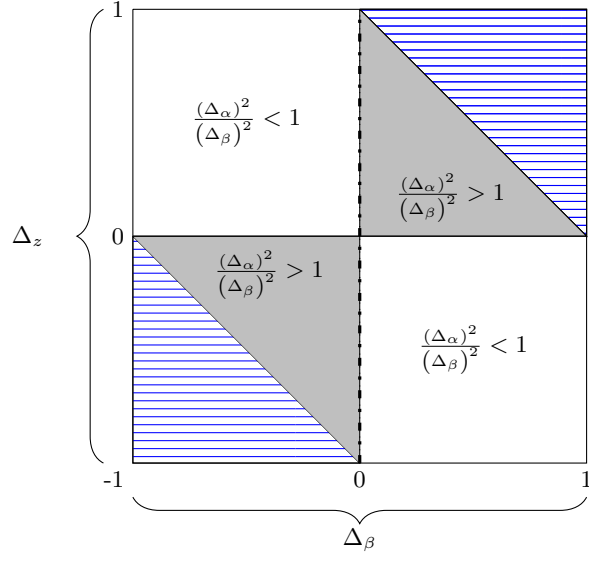
Figure 3: Distortion of $\hat{\theta}_\alpha$ in the $(\beta - z)$ space
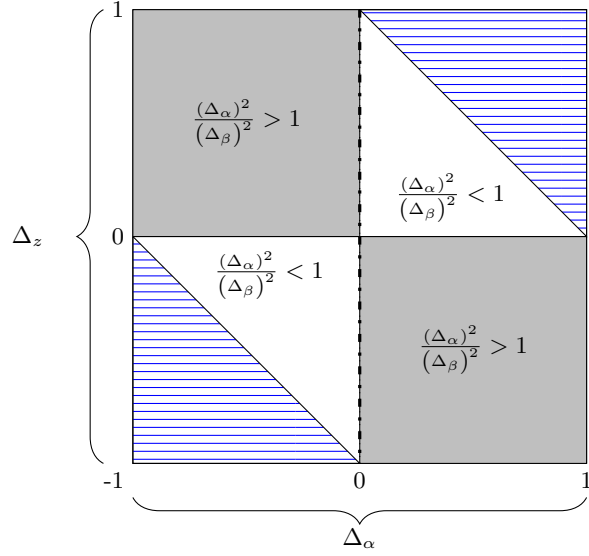
Figure 4: Distortion of $\hat{\theta}_\alpha$ in the $(\alpha - z)$ space

are of the same type and those (grey) in which β and z are of the same type.

Proposition 7. *If regions k and z are of the same type (both either net-recipients or both net-contributors), region k over-borrows less than in the 2-region setting. Conversely, if k and z are of opposite types (k is a recipient/contributor and z is a contributor/recipient), then region k over-borrows more than in the 2-region setting.*

Figure 5: Distortion of $\hat{\theta}_\alpha$ in the $(\alpha - \beta)$ space

Proof. See appendix B. □

The intuition behind Proposition 7 is again that when region α in the coalition is of the same type as the one excl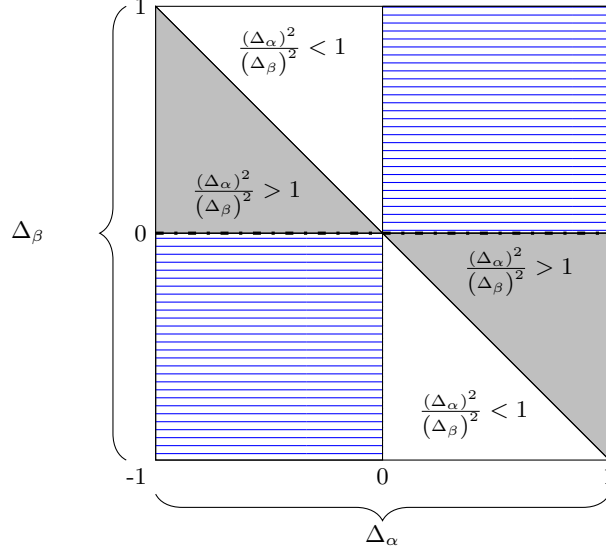uded from the coalition (z), it is as if any benefit of distorting their own behaviour is shared with the excluded region, which makes distortions less profitable. Hence, at the margin, regions stop distorting earlier. Vice-versa, when α and z are of opposite type then region β is playing less aggressive since the cost of an increase in borrowing in region α is shared between β and z. Therefore, it is more profitable to overborrow and at the margin region α ends up borrowing more.

A.3   Coalition formation process

In this section, we add some structure to the 3-region setting of Section A.2, by looking deeper into how a coalition forms. The aim of this extension is twofold. First, we provide a plausible foundation for the 3-region setting in Section A.2, where coalition formation was taken for granted and treated as a black box. Second, we link this section to Section A.1, where voters anticipate that with some probability the (time) preferences of their representative have no direct impact on federal decision making. Indeed, in what follows voters are put in a similar position of uncertainty, as they cannot be sure their own representative will be part of the coalition.

We keep restricting our focus to the minimum-winning coalition, so that $|R| = 3$ and $|K| = 2$. This to demonstrate that equilibria where a coalition of two forms, are indeed

admissible in our setting. Also, whenever a grand-coalition were to form results would be isomorphic to the 2-region setting studied in Section 3.1, hence we would gain little insight by studying such an event. The only difference with respect to Section A.2, therefore, is that we introduce one additional step between federal elections and the decision over t_f: a "formateur" is randomly selected amongst representatives, and is asked to choose a partner to form the federal government. Once this coalition is formed, the representatives of the two regions act cooperatively, as before, and select the federal tax that maximises their joint utility.

We show that, with some probability, the federal tax responds directly to the preferences θ of strategically elected representatives, which is the channel upon which the main model is built. With the complementary probability, federal decisions on taxation do not lend themselves to strategic motives. Hence, results are in line with our previous claim that voters act strategically when it comes to the choice of who is going to represent them at the federal level, as long as they believe the latter will – to a certain extent – be influential in setting the federal tax. Indeed, the more structured model developed here micro-founds one of the interpretations considered in Section A.1, where beliefs were precisely shaped by the probability assigned to the event of being included in the federal coalition.

The timing of the game, hence, becomes the following. In period 1, citizens vote over "local matters" (local taxes, local public good provision and local borrowing) and elect the delegate $\hat{\theta}_r$ to potentially join the federal government. Then, a formateur is randomly selected: each delegate $\hat{\theta}_r$ has the same probability $\frac{1}{3}$ to be selected. To ensure a majority, it suffices for the formateur to choose just one coalition partner. Once this partner is selected, the bargaining process is the same as in Sections A.2 and 3.1, with coalition members cooperatively seeking to maximise joint surplus against the threat of negotiations breaking down.

The type of a region depends on whether it is a contributor or a recipient. Hence, with three regions, it must be that two of them are always of the same type, while the third one is of the opposite one. Regions' type is common knowledge. Hence, at the moment of the vote, citizens know if they are voting in the region of "unique type" or "non-unique type", i.e. whether there is another region of the same type in the federation. We solve those two cases separately, respectively in Sections A.3.1 and A.3.2. In both cases, we show that median voters will be driven to choose the federal representative $\hat{\theta}$ strategically, leading to federal representatives with stronger preferences for consumption in period 2. However, this driving force is mitigated by another one which goes in the opposite direction, and leads to select the median voter as the representative at the federal level. Voters' behaviour will ultimately be the result of a weighted average of the two forces, as depicted (in reduced form) in Section A.1.

We already argued in Section A.2 that a coalition of two regions of a kind allows members

to expropriate the left-out region. When both regions in the coalition are net recipients, they set the largest possible federal tax ($t_f = 1$) and extract all surplus from the excluded region. If both regions are net contributors, they repeal the transfer mechanism and set the federal tax at the lowest possible level ($t_f = 0$).[31] In both cases we arrive at a corner solution, over which voters in the excluded region will have no control. Therefore, when the formateur is of the "non-unique type", their optimal strategy is always to form a coalition with the other region of identical type.

However, a formateur of the "unique type" will have no alternative but to form a coalition with a representative of a region of the opposite type. In this case, the best strategy is to form a coalition with the representative with the lowest preference for consumption in period 2, that is, the one with the lowest $\hat{\theta}$. The formateur's preference to form a coalition with the lowest θ is anticipated by voters and will be reflected in their choice of the representative they elect. The details about this follow below.

A.3.1 Unique type

In this section, we consider the case in which the region's type is opposite to the type of the two other regions. Consider first the possibility that voters' representative is not the formateur, with probability $\frac{2}{3}$. Here then, following the previous argument, the two other regions form a coalition together and expropriate the region of interest, by choosing federal tax $t_f^{\not{E}}$. Utility of the median voter in region r then becomes

$$U_r^m(t_f^{\not{E}}) = g\left(G_r\right) + c\left(C_r^1\right) + \theta_k^m c\left(C_r^2, t_f^{\not{E}}\right) - \eta_r \left(\hat{\theta}_r - \theta_r^m\right)^2. \tag{43}$$

Notice that delegates will have no say over federal spending and, therefore, utility is maximised when $\hat{\theta}_r = \theta_r^m$. That is, when the median voter elects a representative with their same preference over consumption in period 2.

Second, and with the complementary probability $\frac{1}{3}$, the representative is the formateur. Hence, by construction, they are part of the coalition and select, as a partner, the representative with the lowest $\hat{\theta}$. In such case, the two members of the coalition have opposite goals. Expropriating the region outside of the coalition is no longer an option, so they maximise joint surplus as in Section A.2.[32] In a sense coalition partners try to mitigate each other's objectives. The federal tax rate is the result of the weighted utilitarian bargaining solution

---

[31] Clearly, if there are constitutional bounds to the acceptable tax rates, the coalition would opt for these, instead of 0 and 1.

[32] More specifically, the setting for the formateur region is isomorphic to the case when α and z are of the opposite type, while for the region that is not the formateur, the setting corresponds to the one in which α and z are of the same type.

and we denote it by $t_f^*(\hat{\theta}_k)$, with

$$t_f^*(\hat{\theta}_k) = \arg\max_{t_f} \sum_{k \in K} \left(\omega_k \hat{U}_k - F_k \right) = \sum_{k \in K} \omega_k \left(g\left(G_k\right) + c\left(C_k^1\right) + \hat{\theta}_k c\left(C_k^2\right) - F_k \right). \quad (44)$$

The median voter's utility in this case is:

$$U_r^m(t_f^*(\hat{\theta}_r)) = g\left(G_r\right) + c\left(C_r^1\right) + \theta_r^m c\left(C_r^2, t_f^*(\hat{\theta}_r)\right) - \eta_r \left(\hat{\theta}_r - \theta_r^m\right)^2. \quad (45)$$

Hence, when electing the federal representative by maximising Eq. (45), the motive to vote strategically for voters in region r reappears, and is the same as in Section A.2. Indeed, their delegate will now have a clear say over federal spending.

Combining Eq. (43) and Eq. (45) subsequently, and given the probability of becoming the formateur, first period optimisation in region r boils down to

$$\max_{t_r^1, G_r, \hat{\theta}_r} U_r^m(t_f^*(\hat{\theta}_r), \hat{t}_f^{\not{E}}) = \frac{1}{3} U_r^m(t_f^*(\hat{\theta}_r)) + \frac{2}{3} U_r^m(\hat{t}_f^{\not{E}}), \quad (46)$$

subject to

$$C_r^1 = \left(1 - t_r^1\right) Y_r, \quad (47)$$

$$C_r^2 = \left(1 + t_r^1\right) Y_r - G_r + \Delta_r t_f,. \quad (48)$$

Eq. (46) simplifies to

$$\max_{t_r^1, G_r, \hat{\theta}_r} U_r^m = g\left(G_r\right) + c\left(C_r^1\right) + \frac{1}{3}\theta_r^m c\left(C_r^2, t_f^*(\hat{\theta}_r)\right) + \frac{2}{3}\theta_r^m c\left(C_r^2, \hat{t}_f^{\not{E}}\right) - \eta_r \left(\hat{\theta}_r - \theta_r^m\right)^2, \quad (49)$$

which is the isomorphic equivalent of Eq. (112), and consequently also serves as a first example of a possible micro-foundation for the more general case developed in Section A.1.

A.3.2 Not-unique type

Consider now the case in which region r is of a given type and there exist another region of the same type. This brings about three, equally likely, possibilities:

1. The representative of region r becomes the formateur. In this case, the formateur forms a coalition with the representative from the region of identical type. Together they expropriate the region that is left out of the coalition. This leads to a corner solution, similarly to the previous case of expropriation: the only difference is that now our region of interest is the one expropriating. Yet, by choosing a corner solution to expropriate, the choice of t_f does not depend on the type $\hat{\theta}_r$ that the region selects to operate within the federal government. Hence, in this setting, there is no need for strategic manipulation by voters, as their delegate will always get to fully expropriate

the region outside of the coalition. We denote the resulting federal tax by t_f^E, where the expropriating tax will be one of the constitutional boundaries for federal taxation. In such case, the utility of the median voter in region r is:

$$U_r^m(t_f^E) = g(G_r) + c(C_r^1) + \theta_k^m c(C_r^2, t_f^E) - \eta_r \left(\hat{\theta}_r - \theta_r^m\right)^2. \tag{50}$$

2. The formateur is the representative of the region of the same type as our region of interest. Such situation mirrors the previous one. The one region of the same type as our region of interest will invite the region of interest to form a coalition, and together they again expropriate the region left out of the coalition. Once again, voters in region r have no reason to vote strategically, since their delegate always gets to fully expropriate the region outside of the coalition. Eq. (50) is again the utility of the median voter in our region of interest.

3. The formateur is the representative of the unique-type region, hence of the type opposite to our region of interest. Inevitably, the coalition members must have opposing goals. A priori, both regions could be selected by the formateur, yet the formateur has a clear incentive to face the weakest possible opponent, who is represented by the lowest $\hat{\theta}$. The two regions may compete to be part of the coalition, however, competition is limited by the fact that both regions pursue a common goal: to preserve the interests of their (common) type. The benefit of sitting at the table with the formateur is intrinsically determined by the fact of preserving period-2 consumption. There is no ulterior motive to be part of the coalition. As such, compared to a standard competition "à la Bertrand", behaviour is more strategic, which we analyse in the remainder of this section. For now, assume that a representative of type $\bar{\theta}$ is selected to form a coalition. Negotiations will proceed as before, by maximising joint utility, so that the resulting federal tax $\bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})$ is defined as

$$\bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta}) = \arg\max_{t_f} \omega_\phi \left(g(G_\phi) + c(C_\phi^1) + \hat{\theta}_\phi c(C_\phi^2) - F_\phi\right) +$$
$$\bar{\omega}_k \left(g(\overline{G}_k) + c(\overline{C}_k^1) + \bar{\theta} c(\overline{C}^2) - \overline{F}\right) \tag{51}$$

where subscript ϕ denotes the formateur, while overlined variables $(\bar{\omega}, \overline{C}, \overline{G}, \overline{F}, \bar{\theta})$ refer to the region that is chosen to be part of the coalition.

Denote by $i = \{1, 2\}$ the two regions that are of the same kind. Then their maximisation problem boils down to

$$\max_{t_i^1, G_i, \hat{\theta}_i} U_i^m \left(t_f^E, \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})\right) = \frac{2}{3} U_i^m(t_f^E) + \frac{1}{3} U_i^m \left(\bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})\right), \tag{52}$$

subject to

$$C_i^1 = \left(1 - t_i^1\right) Y_i, \tag{53}$$

$$C_i^2 = \left(1 + t_i^1\right) Y_i - G_i + \Delta_i t_f,. \tag{54}$$

where Eq. (52) simplifies to

$$\max_{t_i^1, G_i, \hat{\theta}_i} g\left(G_i\right) + c\left(C_i^1\right) + \frac{2}{3}\theta_i^m c\left(C_i^2, t_f^E\right) + \frac{1}{3}\theta_i^m c\left(C_i^2, \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})\right) - \eta_i\left(\hat{\theta}_i - \theta_i^m\right)^2. \tag{55}$$

The first order condition with respect to $\hat{\theta}_i$ is then $\frac{1}{3}\theta_i^m \frac{\partial c\left(C_i^2, \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})\right)}{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})} \frac{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})}{\partial \hat{\theta}_i} = 2\eta_i\left(\hat{\theta}_i - \theta_i^m\right).$
The solution is $\hat{\theta}_i = \theta_i^m$, whenever $\frac{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})}{\partial \hat{\theta}_i} = 0$. Let's define, whenever $\frac{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})}{\partial \hat{\theta}_i} \neq 0$,

$$\bar{\theta}_i : \frac{1}{3}\theta_i^m \frac{\partial c\left(C_i^2, \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})\right)}{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})} \frac{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})}{\partial \hat{\theta}_i} = 2\eta_i\left(\hat{\theta}_i - \theta_i^m\right) \tag{56}$$

Proposition 8. *When regions 1 and 2 are of the same type, the unique equilibrium for the median voters in both regions is to act strategically and elect $\hat{\theta}_i = \max\{\theta_i^m; \min\{\bar{\theta}_1; \bar{\theta}_2\}\}$, where $i = \{1, 2\}$ and $\bar{\theta}_1$ and $\bar{\theta}_2$ are defined by Eq. (56).*

Proof. See Section B ☐

When a region is of a not-unique type – that is, the region is a net contributor and there exists another contributing region, or the region is a net recipient and there exists another receiving region – two alternatives may materialise, depending on whether or not $\max\{\theta_1^m; \theta_2^m\} < \min\{\bar{\theta}_1; \bar{\theta}_2\}$.

Proposition 8 shows that, if $\theta_i^m \geq \min\{\bar{\theta}_1; \bar{\theta}_2\}$, the elected candidate in region $i = \{1, 2\}$ is $\hat{\theta}_i = \theta_i^m$, hence median voters are acting in a non-strategic way, by selecting a representative of their own type. However, this can only occur in one region at the time, because the condition $\theta_i^m \geq \min\{\bar{\theta}_1; \bar{\theta}_2\}$ cannot be verified simultaneously for both regions. Furthermore, it can only materialise for the region that is left out of the coalition whenever the formateur is of the unique, opposite type.

If $\theta_i^m < \min\{\bar{\theta}_1; \bar{\theta}_2\}$, instead, the elected candidate in region i is $\hat{\theta}_i = \min\{\bar{\theta}_1; \bar{\theta}_2\}$. The existence condition can hold for both regions simultaneously, and it always holds for the region that belongs to the coalition when the formateur is of the unique, opposite type. In such a region, the median voter chooses to distort the type of the elected candidate, by selecting a type $\hat{\theta}_i > \theta_i^m$.

As in Section A.3.1, this equilibrium setting belongs to the family of equilibria described in Section A.1, where the median voter maximises a weighted sum of one element for which the federal tax is orthogonal to the identity of the regional representative and another element for which, instead, the federal tax depends on the identity of the regional representative.

A.4 The bailouts system

The baseline framework described at the top of Section 2 is interpreted, in Section 3, as a setting in which the federal government redistributes across regions through grants. Some minor changes can easily adapt the baseline model, however, to make it more applicable to the scenario where the federal government aids fiscally distressed regions directly, using an ad hoc bailout scheme. This subsection shows how the model is robust to such changes and that the minor adaptations required to interpret the former as a bailout scheme are not affecting the results.

Distinctive characteristics of the Grants and Bailout settings The two settings differ in two aspects. First of all, grants are financed by and distributed to both regions with certainty, whereas bailout interventions will only occur with some probability, when some criteria are met.[33] Second, and most importantly, introducing a system of bailouts often has as its main justification the fact that there is some uncertainty over future regional income. Unlike grants, bailouts are specifically aimed at mitigating the effects of unexpected, negative income shocks by increasing regional gross income ex-post. Therefore, it is crucial to allow for the possibility that income in period 2 in a given region may fall below reasonable levels. The model introduces this aspect through the income shock ϵ_r. However, in the grant setting there is no need to assume income shocks, hence, for the sake of parsimony, we assumed that $\epsilon_r = 0$ and hence that $Y_r^1 = Y_r^2 \equiv Y_r$.

Since bailouts are aimed at rescuing administrations when they are hit in a way that their functioning is compromised, we re-introduce the uncertainty in regional income: in period 2 an income shock may impact a region's purchasing power. Income in period 2 is then $Y_r^2 = Y_r^1 + \epsilon_r$, where $\epsilon_r \in \Re$ is an exogenous random variable. We do not introduce specific assumptions on the distribution function of ϵ_r.

We assume again that $|R| = 2$ and so that the grand coalition forms with $K = \{\alpha, \beta\}$. Furthermore, we assume that the constitution defines the conditions under which a region is rescued. Such conditions could possibly depend on the initial income Y_r^1 in each region, on the realisation of the shocks ϵ_r and on any other parameter, in a way that the total probability that one of the regions is rescued is ρ.[34] When the bailout system is put in motion, the intra-regional transfer is $\Delta_\alpha^f = T(\gamma) \frac{(Y_\beta^2 - Y_\alpha^2)}{2}$, where $T(\gamma) \in [0, 1]$ is an increasing function of the parameter γ which defines the generosity of the transfer scheme, whereas the term $\frac{(Y_\beta^2 - Y_\alpha^2)}{2}$

---

[33]It has to be noticed that in a two-region federation, one region will always be the net receiver and the other the net contributor, both under the grant and bailout systems.

[34]The rescue is granted in order to cope with some unexpected unfavourable events, such as a negative income shock or an idiosyncratic exogenous crisis (hence, a very low or negative realisation of ϵ) or some natural disaster.

ensures at the same time that the money transfer goes from the richer to the poorer region and also ensures that the transfer is ranking invariant (the bailed-out region after the transfer cannot become richer than the contributing region). Hence, according to this system, the rich region will transfer a share of their income to the poor region which depends on the generosity parameter γ. A well known and intuitive example of function $T(\gamma)$ could be $T(\gamma) = A_\gamma$ where A_γ is the Atkinson measure of inequality with parameter γ. Then γ represents the aversion to income inequality (that is, the willingness to redistribute), and the index A_γ defines the share of income that agents are willing to give up in exchange for an egalitarian society.

As in the baseline model, in period 1 voters select G_k, t_k^1 and a representative of type $\hat{\theta}_k$ that will join the federal government.[35] However, the federal government is now in charge of choosing the generosity parameter γ. Such a transfer is activated with probability ρ. Once again, each region of the federation thus marks a constituency where every single voter is a potential candidate, from which one candidate is chosen to represent this respective state in the federal legislature. Once elected, representatives go on to bargain over the rescue transfer. We remain agnostic about the exact shape of T_γ, as long as the function is increasing in γ.

At the beginning of period 2, the federal government (composed of the two regional representatives) sets the parameter γ that indirectly defines $T(\gamma)$. Then the realisation of ϵ occurs and regions set the local tax t_k^2 so to repay the debt.

The utility of a voter is again defined by Eq. (1). The net federal transfer is $\Delta_\alpha^f = T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2}$. Hence, the net-of-transfer disposable income in region α in period 2 is $Y_\alpha^2 + \Delta_\alpha^f$. Such income is partially used to pay the debt back and the remaining is consumed. Therefore, from Eqs. (6) to (9), the variables in the model are linked in the following way:

$$B_\alpha = G_\alpha - t_\alpha^1 Y_\alpha^1 \tag{57}$$

$$t_\alpha^2 = \frac{G_\alpha - t_\alpha^1 Y_\alpha^1}{\left(Y_\alpha^2 + T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2}\right)} \tag{58}$$

$$C_\alpha^1 = \left(1 - t_\alpha^1\right) Y_\alpha^1 \tag{59}$$

$$C_\alpha^2 = (1 + t_\alpha^1)Y_\alpha^1 + \epsilon_\alpha + T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2} - G_\alpha. \tag{60}$$

Clearly, Eq. (1) meets all the requirements in Gans and Smart (1996) and therefore the median voter theorem applies.

We keep the assumption, previously discussed, that the federal government behaves cooperatively, and seeks the weighted utilitarian bargaining solution. The federation is thus governed by a cooperative legislature of regional representatives, counting only two representatives in our case. Once elected, and once the federal coalition is successfully formed,

---

[35] Again, the difference between what is raised through the local tax and what is needed to finance the public good is obtained through borrowing B_k, which must be paid back at the end of period 2.

representatives decide on γ. As a result the federal coalition maximises the sum of the utilities of representatives:

$$\max_{\gamma} \sum_{k \in K} \omega_k E\left(\hat{U}_k - F_k\right) = \sum_{k \in K} \omega_k E\left(g\left(G_k\right) + c\left(C_k^1\right) + \hat{\theta}_k c\left(C_k^2\right) - F_k\right), \tag{61}$$

subject to Eqs. (57) to (60)

where $\hat{U}_k$ represents the utility of the representative elected in region k, given Eq. (1) and the fact that by construction the type of the representative is $\theta_k = \hat{\theta}_k$.

Lemma 7. *The first order condition of the optimisation problem is*

$$\phi_f = 0, \ with$$
$$\phi_f \equiv \rho \omega_\alpha \frac{\partial T(\gamma)}{\partial \gamma} E\left(\hat{\theta}_\alpha c'\left(C_\alpha^2\right)\right) - \rho \omega_\beta \frac{\partial T(\gamma)}{\partial \gamma} E\left(\hat{\theta}_\beta c'\left(C_\beta^2\right)\right). \tag{62}$$

Comparing Eq. (62) with Eq. (17), it is immediate to notice that the only relevant difference appears to be that Eq. (62) is expressed in expected terms, because in this setting income is uncertain. However, the different setting leads exactly to the same result and all further steps and conclusions would hold in this new framework. In other words, the setting is robust to a change in how transfers occur, in order to easily adapt to the bailout context.

In terms of interpretation, though, the uncertainty about future income allows us to interpret the parameter θ as a measure of risk aversion, in the sense that higher θ types prefer more consumption in the risky future, while lower θ types prefer to shift consumption to the present, where income is certain. Therefore, the higher the θ type, the more agents will be interested in insuring themselves against negative shocks. Bailouts, as we modelled them, are thus a kind of insurance that the state provides to all regions. Interpreting bailouts as an insurance, the over-borrowing that we observe in the model has a flavour of the usual moral-hazard reaction of an insured agent.

Appendix B   Proofs

Proof of Lemma 1 and Lemma 2. The federal optimisation problem set out in Section 3.1.1 can be summed up as

$$\max_{t_f} \sum_{k \in K} \omega_k \hat{U}_k = \omega_k \left(g\left(G_k\right) + c\left(C_k^1\right) + \hat{\theta}_k c\left(C_k^2\right)\right), \tag{63}$$

subject to

$$B_k = G_k - t_k^1 Y_k, \tag{64}$$

$$t_k^2 = \frac{G_k - t_k^1 Y_k}{(Y_k + \Delta_k t_f)}, \tag{65}$$

$$C_k^1 = \left(1 - t_k^1\right) Y_k, \tag{66}$$

$$C_k^2 = (1 + t_k^1) Y_k + \Delta_k t_f - G_k. \tag{67}$$

so that we get the following first order condition

$$\omega_\alpha \hat{\theta}_\alpha c' \left(C_\alpha^2\right) \Delta_\alpha + \omega_\beta \hat{\theta}_\beta c' \left(C_\beta^2\right) \Delta_\beta = 0. \tag{68}$$

Since $\Delta_\beta = -\Delta_\alpha$, it follows that

$$\phi_f \equiv \omega_\alpha \hat{\theta}_\alpha c' \left(C_\alpha^2\right) - \omega_\beta \hat{\theta}_\beta c' \left(C_\beta^2\right) = 0. \tag{69}$$

Applying the implicit function theorem, we then obtain

$$\frac{\partial t_f}{\partial \hat{\theta}_\alpha} = -\frac{\frac{\partial \phi_f}{\partial \hat{\theta}_\alpha}}{\frac{\partial \phi_f}{\partial t_f}} = -\frac{\omega_\alpha c'(C_\alpha^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)\Delta_\alpha - \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\Delta_\beta}, \tag{70}$$

or,

$$\frac{\partial t_f}{\partial \hat{\theta}_\alpha} = -\frac{\omega_\alpha c'(C_\alpha^2)}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_\alpha}. \tag{71}$$

Similarly, we obtain

$$\frac{\partial t_f}{\partial G_\alpha} = -\frac{\frac{\partial \phi_f}{\partial G_\alpha}}{\frac{\partial \phi_f}{\partial t_f}} = -\frac{-\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)\Delta_\alpha - \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\Delta_\beta}, \tag{72}$$

or,

$$\frac{\partial t_f}{\partial G_\alpha} = \frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_\alpha}. \tag{73}$$

Lastly, we obtain

$$\frac{\partial t_f}{\partial t_\alpha^1} = -\frac{\frac{\partial \phi_f}{\partial t_\alpha^1}}{\frac{\partial \phi_f}{\partial t_f}} = -\frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)Y_\alpha}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)\Delta_\alpha - \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\Delta_\beta}, \tag{74}$$

or,

$$\frac{\partial t_f}{\partial t_\alpha^1} = -\frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)Y_\alpha}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_\alpha}. \tag{75}$$

The sign of equations Eqs. (71), (73) and (75) follows immediately, when one takes into account that $c''(\cdot) < 0$, while $\omega_k, \hat{\theta}_k, Y_k$ and $c'(\cdot)$ are all positive. $\qquad\square$

48

Proof of Lemma 3. The optimisation problem of the median voter in region $k \in K$ set out in Section 3.1.2 can be summed up as

$$\max_{t_k^1, G_k, \hat{\theta}_k} U_k = g(G_k) + c(C_k^1) + \theta_k^m c(C_k^2) - \eta_k (\hat{\theta}_k - \theta_k^m)^2, \tag{76}$$

subject to Eq. (17) and

$$C_k^1 = (1 - t_k^1) Y_k \tag{77}$$

$$C_k^2 = (1 + t_k^1) Y_k - G_k + \Delta_k t_f \tag{78}$$

so that we get the following first order condition for t_k^1

$$-c'(C_k^1) Y_k + \theta_k^m c'(C_k^2) \left(\Delta_k \frac{\partial t_f}{\partial t_k^1} + Y_k \right) = 0, \tag{79}$$

which, using Eq. (75) for $k = \alpha$ yields

$$\frac{c'(C_\alpha^1)}{c'(C_\alpha^2)} = \frac{\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} \theta_\alpha^m. \tag{80}$$

For G_k, we get the following first order condition

$$g'(G_k) + \theta_k^m c'(C_k^2) \left(\Delta_k \frac{\partial t_f}{\partial G_k} - 1 \right) = 0, \tag{81}$$

which, using Eq. (73) for $k = \alpha$ yields

$$\frac{g'(G_\alpha)}{c'(C_\alpha^2)} = \frac{\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} \theta_\alpha^m. \tag{82}$$

For $\hat{\theta}_k$ finally, we obtain the following first order condition

$$\theta_k^m c'(C_k^2) \left(\Delta_k \frac{\partial t_f}{\partial \hat{\theta}_k} \right) - 2\eta_k (\hat{\theta}_k - \theta_k^m) = 0, \tag{83}$$

which, using Eq. (71) for $k = \alpha$ yields

$$-\theta_\alpha^m \frac{\omega_\alpha c'(C_\alpha^2)^2}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} = 2\eta_\alpha (\hat{\theta}_\alpha - \theta_\alpha^m). \tag{84}$$

□

Proof of Proposition 1. Rewriting Eq. (23), we know that

$$\frac{-\omega_\alpha \theta_\alpha^m}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} = \frac{2\eta_\alpha^m (\hat{\theta}_\alpha - \theta_\alpha^m)}{c'(C_\alpha^2)^2} \quad . \tag{85}$$

Since $c''(C_k^2) < 0$, the left-hand side is always positive and hence a solution can only exist if $\hat{\theta}_k \geq \theta_k^m$. Hence, in equilibrium, the elected federal representative will always have a weakly stronger preference for consumption in period 2 than the median voter. □

Proof of Corollary 1 and Corollary 2. Eq. (24) is simply obtained by combining Eq. (23) computed respectively for region α and β. When there is no federal redistribution, we know that $\Delta_r = 0$ for all regions $r \in R$. The first order condition with respect to $\hat{\theta}_k$, as shown in Eq. (83), is $\theta_k^m c'(C_k^2)\Delta_k \frac{\partial t_f}{\hat{\theta}_k} - 2\eta_k(\hat{\theta}_k - \theta_k^m) = 0$. However, if $\Delta_k = 0$, it reduces to

$$2\eta_k^m \left(\hat{\theta}_k - \theta_k^m \right) = 0. \tag{86}$$

In equilibrium we therefore always have that $\hat{\theta}_k = \theta_k^m$, so that the elected federal representative and the median voter will have identical preferences for consumption in period 2. As a result, there will be no strategic voting. $\square$

Proof of Lemma 4. The optimisation of median voter utility is set out in Section 3.1.2, but then under the scenario without federal co-funding (denoted here by U_k^{NF}), it can be summed up as

$$\max_{t_k^1, G_k, \theta_k^m} U_k^{NF} = g\left(G_k\right) + c\left(C_k^1\right) + \theta_k^m c\left(C_k^2\right) - \eta_k \left(\hat{\theta}_k - \theta_k^m \right)^2, \tag{87}$$

subject to

$$C_k^1 = \left(1 - t_k^1\right) Y_k \tag{88}$$

$$C_k^2 = \left(1 + t_k^1\right) Y_k - G_k \tag{89}$$

so that we get the following first order condition for t_k^1

$$- c'(C_k^1)Y_k + \theta_k^m c'(C_k^2)Y_k = 0, \tag{90}$$

which yields

$$\frac{c'(C_k^1)}{c'(C_k^2)} = \theta_k^m. \tag{91}$$

For G_k, we get the following first order condition

$$g'(G_k) - \theta_k^m c'(C_k^2) = 0, \tag{92}$$

which yields

$$\frac{g'(G_k)}{c'(C_k^2)} = \theta_k^m. \tag{93}$$

For $\hat{\theta}_k$ finally, we obtain the following first order condition

$$- 2\eta_k(\hat{\theta}_k - \theta_k^m) = 0, \tag{94}$$

which implies there will be no strategic voting in this scenario, since $\hat{\theta}_k = \theta_k^m$ as a result. Combining and rewriting Eq. (91) and Eq. (93) then gives us Eq. (26) of Lemma 4. $\square$

Proof of Proposition 2. To ascertain whether regional governments would end up borrowing less in a scenario where the federal government does not intervene, we evaluate the equilibrium characterised by Lemma 4 at the optimal equilibrium outcome G_k^* and t_k^{1*} obtained in Lemma 3, where the federal government did intervene. If welfare is decreasing on the margin at the equilibrium level G_k^*, but increasing at t_k^{1*}, we know that region k has chosen higher public provision and period 1 consumption than it otherwise would have in the alternative scenario without federal intervention. To achieve both outcomes simultaneously, region k would necessarily have needed to borrow more. Using both first order conditions for G_k and t_k^1 derived under Lemma 4, and evaluating these at the equilibrium values implicitly defined by Lemma 3, we thus need to verify whether

$$\frac{\partial U_k^{NF}}{\partial G_k}\bigg|_{G_k=G_k^*} \lessgtr 0, \tag{95}$$

$$\frac{\partial U_k^{NF}}{\partial t_k^1}\bigg|_{t_k^1=t_k^{1*}} \lessgtr 0. \tag{96}$$

Taking into account the maximisation problem set up under Lemma 3, we can write both trade-offs expressed by Eq. (95) and Eq. (96) above as

$$\frac{\partial U_k^{NF}}{\partial G_k}\bigg|_{G_k=G_k^*} \lessgtr 0 \equiv \frac{\partial U_k}{\partial G_k}\bigg|_{G_k=G_k^*}, \tag{97}$$

$$\frac{\partial U_k^{NF}}{\partial t_k^1}\bigg|_{t_k^1=t_k^{1*}} \lessgtr 0 \equiv \frac{\partial U_k}{\partial t_k^1}\bigg|_{t_k^1=t_k^{1*}}. \tag{98}$$

Using Eq. (29), Eq. (30), Eq. (90) and Eq. (92), the evaluation exercise can then be written as

$$\left(\frac{\partial U_k^{NF}}{\partial G_k} - \frac{\partial U_k}{\partial G_k}\right)\bigg|_{G_k=G_k^*} = -\theta_k^m c'(C_k^2)\Delta_k \frac{\partial t_f}{\partial G_k} \lessgtr 0, \tag{99}$$

$$\left(\frac{\partial U_k^{NF}}{\partial t_k^1} - \frac{\partial U_k}{\partial t_k^1}\right)\bigg|_{t_k^1=t_k^{1*}} = \theta_k^m c'(C_k^2)\Delta_k \frac{\partial t_f}{\partial t_k^1} \lessgtr 0, \tag{100}$$

to find, after plugging in Eq. (73) and Eq. (75), that

$$\left(\frac{\partial U_k^{NF}}{\partial G_k} - \frac{\partial U_k}{\partial G_k}\right)\bigg|_{G_k=G_k^*} = -\theta_k^m c'(C_k^2)\left(\Delta_k \frac{\omega_k\hat{\theta}_k c''(C_k^2)}{\left(\omega_\alpha\hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta\hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_k}\right) < 0, \tag{101}$$

$$\left(\frac{\partial U_k^{NF}}{\partial t_k^1} - \frac{\partial U_k}{\partial t_k^1}\right)\bigg|_{t_k^1=t_k^{1*}} = \theta_k^m c'(C_k^2)\left(\Delta_k \frac{\omega_k\hat{\theta}_k c''(C_k^2)Y_k}{\left(\omega_\alpha\hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta\hat{\theta}_\beta c''(C_\beta^2)\right)\Delta_k}\right) > 0, \tag{102}$$

which, as explained above, proves that regional governments will overborrow compared to the setting where the federal government does not intervene. $\qquad\square$

Proof of Proposition 3. In the scenario where the federal government intervenes in regional finance, but where median voters naively cast their vote instead of voting strategically, we know that the elected representative will have the median voter's preference type, so that $\hat{\theta}_k = \theta_k^m$. It therefore suffices to evaluate Eq. (101) and Eq. (102) at the median voter's true preferences, so that

$$\left(\frac{\partial U_k^{NF}}{\partial G_k} - \frac{\partial U_k}{\partial G_k} \bigg|_{\hat{\theta}_k = \theta_k^m} \right) \bigg|_{G_k = G_k^*} = -\theta_k^m c'(C_k^2) \left(\Delta_k \frac{\omega_k \theta_k^m c''(C_k^2)}{\left(\omega_\alpha \theta_\alpha^m c''(C_\alpha^2) + \omega_\beta \theta_\beta^m c''(C_\beta^2) \right) \Delta_k} \right) < 0,$$

(103)

$$\left(\frac{\partial U_k^{NF}}{\partial t_k^1} - \frac{\partial U_k}{\partial t_k^1} \bigg|_{\hat{\theta}_k = \theta_k^m} \right) \bigg|_{t_k^1 = t_k^{1*}} = \theta_k^m c'(C_k^2) \left(\Delta_k \frac{\omega_k \theta_k^m c''(C_k^2) Y_k}{\left(\omega_\alpha \theta_\alpha^m c''(C_\alpha^2) + \omega_\beta \theta_\beta^m c''(C_\beta^2) \right) \Delta_k} \right) > 0,$$

(104)

which shows that more overborrowing as compared to the scenario with naif voters would ensue, since we know that $\hat{\theta}_k > \theta_k^m$ in the equilibrium characterised by Lemma 3.

$\square$

Proof of Lemma 6. We want to compare, for $G_k, \hat{\theta}_k$ and t_k^1, the marginal impact that these have on t_f in the 2-region and 3-region setting. This implies comparing Eqs. (18) to (20) with Eqs. (35) to (37).

Notice that $\frac{\partial t_f}{\partial G_k} > 0, \frac{\partial t_f}{\partial \hat{\theta}_k} > 0, \frac{\partial t_f}{\partial t_k^1} < 0$. Therefore, we will say that the variable has a larger impact (in absolute terms) on t_f:

- for G_k and $\hat{\theta}_k$, when the nominal value of the derivative is larger,

- for t_k^1, when the nominal value of the derivative is smaller.

Comparing (pairwise) the 2- and 3-region derivatives, all three inequalities boil down to the following: the impact in absolute value of each of the variables is larger in the 3-region setting if

$$\frac{1}{\omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \frac{(\Delta_\beta)^2}{\Delta_\alpha} c''(C_\beta^2)} < \frac{1}{\omega_\alpha \hat{\theta}_\alpha \Delta_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \Delta_\alpha c''(C_\beta^2)}.$$

(105)

The equation simplifies to $\Delta_\alpha > \frac{(\Delta_\beta)^2}{\Delta_\alpha}$. We consider separately: a) $\Delta_\alpha > 0$, b) $\Delta_\alpha < 0$.

a) $\Delta_\alpha > \frac{(\Delta_\beta)^2}{\Delta_\alpha}$ translates into $\frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} > 1$. We claim that this occurs if and only if $\Delta_z < 0$.

b) $\Delta_\alpha > \frac{(\Delta_\beta)^2}{\Delta_\alpha}$ translates into $\frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} < 1$. We claim that this occurs if and only if $\Delta_z < 0$.

If the types of α and z are opposite, it means that $|\Delta_\alpha| = |\Delta_\beta| + |\Delta_z|$. Therefore, $\frac{(\Delta_\alpha)^2}{(\Delta_\beta)^2} = \frac{(|\Delta_\beta|+|\Delta_z|)^2}{|\Delta_\beta|^2} > 1$. Hence, we conclude that in case a), for the inequality to be true it must be that α and z are of opposite type. Since, in case a), $\Delta_\alpha > 0$ it must follow that $\Delta_z < 0$.

If the types of α and z are the same, it means that $|\Delta_\alpha| + |\Delta_z| = |\Delta_\beta|$. Therefore, $\frac{(\Delta_\alpha)^2}{(\Delta_\beta)^2} = \frac{|\Delta_\alpha|^2}{(|\Delta_\alpha|+|\Delta_z|)^2} < 1$. Hence, we conclude that in case b), for the inequality to be true it must be that α and z are of the same type. Since, in case b), $\Delta_\alpha < 0$ it must follow that $\Delta_z < 0$.

$\square$

Proof of Proposition 6. From Eq. (42) it immediately appears that the relative distortion (measured by the left hand side) is larger in the 3-region case if and only if $\frac{(\Delta_\alpha)^2}{\left(\Delta_\beta\right)^2} > 1$.

In the proof of Lemma 6 we proved that $\frac{\left(\Delta_\beta\right)^2}{(\Delta_\alpha)^2} > 1$ if the types of α and z are opposite, while $\frac{\left(\Delta_\beta\right)^2}{(\Delta_\alpha)^2} < 1$ when the types of α and z are the same.

Hence, the distortion in the 3-region setting is larger when the types of α and z are opposite, that is: $\mathrm{sgn}(\Delta_\alpha) \neq \mathrm{sgn}(\Delta_z)$. Conversely, the distortion in the 2-region setting is larger when the types of α and z are the same, that is: $\mathrm{sgn}(\Delta_\alpha) = \mathrm{sgn}(\Delta_z)$.

Clearly, if $\Delta_z = 0$, then $(\Delta_\alpha)^2 = (\Delta_\beta)^2$ and the distortion is the same in the two cases. $\square$

Proof of Proposition 7. The marginal rate of substitution $\left|MRS_k^{G,2}(\theta_k^m)\right|$ represents the trade-off between borrowing and future consumption. First notice that $\left|MRS_k^{G,2}(\theta_k^m)\right|$ decreases when consumption of G_k increases (for a given level of C_k^2).

For the 2-region case, we have that: $\left|MRS_\alpha^{G,2}(\theta_\alpha^m)\right| = \frac{\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)+\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}$. For the 3-region case, we have that: $\left|MRS_\alpha^{G,2}(\theta_\alpha^m)\right| = \frac{\omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha (\Delta_\alpha)^2 c''(C_\alpha^2)+\omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)}$.

Hence, comparing the two, if the following inequality is true then the marginal rate is larger in the 2-region setting, meaning that region α borrows less in such setting.

$$\frac{\omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} > \frac{\omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)} \tag{106}$$

$$\frac{1}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} < \frac{(\Delta_\beta)^2}{\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2)} \tag{107}$$

$$\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta \left(\Delta_\beta\right)^2 c''(C_\beta^2) < (\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2))(\Delta_\beta)^2 \tag{108}$$

$$\omega_\alpha \hat{\theta}_\alpha \left(\Delta_\alpha\right)^2 c''(C_\alpha^2) < \omega_\alpha \hat{\theta}_\alpha (\Delta_\beta)^2 c''(C_\alpha^2) \tag{109}$$

$$\hat{\theta}_\alpha c''(C_\alpha^2)((\Delta_\alpha)^2 - (\Delta_\beta)^2) < 0 \qquad (110)$$

Hence

$$(\Delta_\alpha)^2 - (\Delta_\beta)^2 > 0 \qquad (111)$$

Hence, region α borrows more in the 3-region setting when $(\Delta_\alpha)^2 > (\Delta_\beta)^2$.

In the proof of Lemma 6 we proved that $\frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} > 1$ if the types of α and z are opposite, while $\frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} < 1$ when the types of α and z are the same and, finally, that $\frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} = 1$ if $\Delta_z = 0$. $\qquad \square$

Proof of Lemma 5. The maximisation problem of the median voter in region $k \in K$ set out in Section A.1 can be summed up as

$$\max_{t_k^1, G_k, \hat{\theta}_k} U_k = g\left(G_k\right) + c\left(C_k^1\right) + p_k \theta_k^m c\left(C_k^2, t_f^*\right) + (1-p_k)\theta_k^m c\left(C_k^2, \hat{t}_f\right) - \eta_k \left(\hat{\theta}_k - \theta_k^m\right)^2, \quad (112)$$

subject to

$$C_k^1 = \left(1 - t_k^1\right) Y_k \qquad (113)$$

$$C_k^2 = \left(1 + t_k^1\right) Y_k - G_k + \Delta_k t_f \qquad (114)$$

where we use t_f^* to denote the federal tax implicitly defined by Eq. (17), in order to distinguish it from $\hat{t}_f$, which represents the federal tax when exogenously set.

The first order condition for t_k^1 is

$$-c'(C_k^1)Y_k + \theta_k^m c'(C_k^2)\left(\Delta_k p_k \frac{\partial t_f^*}{\partial t_k^1} + Y_k\right) = 0, \qquad (115)$$

which, using Eq. (75) for $k = \alpha$ yields Eq. (29). For G_k, we get the following first order condition

$$g'(G_k) + \theta_k^m c'(C_k^2)\left(\Delta_k p_k \frac{\partial t_f^*}{\partial G_k} - 1\right) = 0, \qquad (116)$$

which, using Eq. (73) for $k = \alpha$ yields Eq. (30). For $\hat{\theta}_k$ finally, we obtain the following first order condition

$$\theta_k^m c'(C_k^2)\left(\Delta_k p_k \frac{\partial t_f^*}{\partial \hat{\theta}_k}\right) - 2\eta_k(\hat{\theta}_k - \theta_k^m) = 0, \qquad (117)$$

which, using Eq. (71) for $k = \alpha$ yields Eq. (31). $\qquad \square$

Proof of Proposition 4. To ascertain whether regional governments, in the setting of Section A.1, would end up borrowing less in a scenario where the federal government does not intervene, we evaluate the equilibrium characterised by Lemma 4 at the optimal equilibrium

outcome G_k^* and t_k^{1*} obtained in Lemma 5, where the federal government did intervene. As in the proof of Proposition 2 above, using both first order conditions for G_k and t_k^1 derived under Lemma 4 and evaluating these at the equilibrium values implicitly defined by Lemma 5, we thus need to verify whether

$$\left. \frac{\partial U_k^{NF}}{\partial G_k} \right|_{G_k=G_k^*} \lessgtr 0, \tag{118}$$

$$\left. \frac{\partial U_k^{NF}}{\partial t_k^1} \right|_{t_k^1=t_k^{1*}} \lessgtr 0. \tag{119}$$

Taking into account the maximisation problem set up under Lemma 5 subsequently, we can write both trade-offs expressed by Eq. (95) and Eq. (96) above as

$$\left. \frac{\partial U_k^{NF}}{\partial G_k} \right|_{G_k=G_k^*} \lessgtr 0 \equiv \left. \frac{\partial U_k}{\partial G_k} \right|_{G_k=G_k^*}, \tag{120}$$

$$\left. \frac{\partial U_k^{NF}}{\partial t_k^1} \right|_{t_k^1=t_k^{1*}} \lessgtr 0 \equiv \left. \frac{\partial U_k}{\partial t_k^1} \right|_{t_k^1=t_k^{1*}}. \tag{121}$$

Using Eq. (21), Eq. (22), Eq. (90) and Eq. (92), the evaluation exercise can then be written as

$$\left. \left(\frac{\partial U_k^{NF}}{\partial G_k} - \frac{\partial U_k}{\partial G_k} \right) \right|_{G_k=G_k^*} = -\theta_k^m c'(C_k^2) \Delta_k p_k \frac{\partial t_f^*}{\partial G_k} \lessgtr 0, \tag{122}$$

$$\left. \left(\frac{\partial U_k^{NF}}{\partial \partial t_k^1} - \frac{\partial U_k}{\partial t_k^1} \right) \right|_{t_k^1=t_k^{1*}} = \theta_k^m c'(C_k^2) \Delta_k p_k \frac{\partial t_f^*}{\partial t_k^1} \lessgtr 0, \tag{123}$$

to find, for $k = \alpha$ and after plugging in Eq. (73) and Eq. (75), that

$$\left. \left(\frac{\partial U_\alpha^{NF}}{\partial G_\alpha} - \frac{\partial U_\alpha}{\partial G_\alpha} \right) \right|_{G_\alpha=G_\alpha^*} = -\theta_\alpha^m c'(C_\alpha^2) \left(\Delta_\alpha p_\alpha \frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2)}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2) \right) \Delta_\alpha} \right) < 0, \tag{124}$$

$$\left. \left(\frac{\partial U_\alpha^{NF}}{\partial \partial t_\alpha^1} - \frac{\partial U_\alpha}{\partial t_\alpha^1} \right) \right|_{t_\alpha^1=t_\alpha^{1*}} = \theta_\alpha^m c'(C_\alpha^2) \left(\Delta_\alpha p_\alpha \frac{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) Y_\alpha}{\left(\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2) \right) \Delta_\alpha} \right) > 0, \tag{125}$$

which, as explained in the proof of Proposition 2, proves that regional governments will overborrow compared to the setting where the federal government does not intervene, an effect which is again mitigated by voter beliefs denoted by p_α. $\qquad \square$

Proof of Proposition 5. In the setting of Section A.1, we still have that the marginal rate of substitution $\left| MRS_k^{G,2}(\theta_k^m) \right|$ represents the trade-off between borrowing and future consumption, which decreases when consumption of G_k increases (for a given level of C_k^2).

Taking into account belief structures of the voters expressed by p_k, we know From Proposition 4 that

$$\left| MRS_\alpha^{G,2}(\theta_\alpha^m) \right| = \frac{(1 - p_\alpha)\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}. \tag{126}$$

For the 3-region case subsequently, including voter beliefs gives us the following expression

$$\left| MRS_\alpha^{G,2}(\theta_\alpha^m) \right| = \frac{(1 - p_\alpha)\omega_\alpha \hat{\theta}_\alpha (\Delta_\alpha)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha (\Delta_\alpha)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}. \tag{127}$$

Now, and as before under Proposition 7, region α will overborrow less in the 2-region setting when – as compared to the 3-region setting – its marginal rate of substitution is larger, or

$$\frac{(1 - p_\alpha)\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)} > \frac{(1 - p_\alpha)\omega_\alpha \hat{\theta}_\alpha (\Delta_\alpha)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}{\omega_\alpha \hat{\theta}_\alpha (\Delta_\alpha)^2 c''(C_\alpha^2) + \omega_\beta \hat{\theta}_\beta (\Delta_\beta)^2 c''(C_\beta^2)}. \tag{128}$$

Then, with $A = \omega_\alpha \hat{\theta}_\alpha c''(C_\alpha)$ and $B = \omega_\beta \hat{\theta}_\beta c''(C_\beta^2)$, Eq. (128) simplifies to

$$\left((1 - p_\alpha)A + B \right)\left((\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right) > \left((1 - p_\alpha)(\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right)(A + B), \tag{129}$$

so that, rewriting, we have that

$$(A + B)\left((\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right) - p_\alpha A \left((\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right) >$$
$$(A + B)\left((\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right) - p_\alpha (\Delta_\alpha)^2 A(A + B), \tag{130}$$

or, with the first term cancelling out on both sides, that

$$- p_\alpha A \left((\Delta_\alpha)^2 A + (\Delta_\beta)^2 B \right) > -p_\alpha (\Delta_\alpha)^2 A(A + B), \tag{131}$$

which can be written as

$$\left(A + \frac{(\Delta_\beta)^2}{(\Delta_\alpha)^2} B \right) > (A + B). \tag{132}$$

The inequality in Eq. (132) then only holds when

$$(\Delta_\alpha)^2 > (\Delta_\beta)^2, \tag{133}$$

hence, region α borrows more in the 3-region setting when $(\Delta_\alpha)^2 > (\Delta_\beta)^2$. For the same reasons as in the proof of Proposition 7, this occurs if region α is a recipient (contributor) and region z is a contributor (recipient). $\qquad\square$

Proof of Lemma 7. Knowing that the only term in Eq. (61) that depends on γ is C_k^2, the maximisation problem is equivalent to

$$\max_\gamma \sum_{k \in K} E\left(\omega_k \hat{\theta}_k c\left((1 + t_k^1)Y_k^1 + \epsilon_k + \Delta_k^f - G_k \right) \right) - F_k. \tag{134}$$

Because bailouts only occur with probability ρ and $\Delta_k^f = T(\gamma)\frac{(Y_{-k}^2 - Y_k^2)}{2}$, Eq. (134) can be further rewritten as

$$
\begin{aligned}
\max_\gamma \ (1-\rho)E\left(\omega_\alpha\hat{\theta}_\alpha c\left((1+t_\alpha^1)Y_\alpha^1 + \epsilon_\alpha - G_\alpha\right) + \omega_\beta\hat{\theta}_\beta c\left((1+t_\beta^1)Y_\beta^1 + \epsilon_\beta - G_\beta\right)\right) \\
+\rho E\left(\omega_\alpha\hat{\theta}_\alpha c\left((1+t_\alpha^1)Y_\alpha^1 + \epsilon_\alpha + T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2} - G_\alpha\right)\right. \\
\left. +\omega_\beta\hat{\theta}_\beta c\left((1+t_\beta^1)Y_\beta^1 + \epsilon_\beta - T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2} - G_\beta\right)\right),
\end{aligned}
\tag{135}
$$

and since the first term does not depend on γ, we can further simplify the equation as

$$
\begin{aligned}
\max_\gamma \ \rho E\left(\omega_\alpha\hat{\theta}_\alpha c\left((1+t_\alpha^1)Y_\alpha^1 + \epsilon_\alpha + T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2} - G_\alpha\right)\right) \\
+\rho E\left(\omega_\beta\hat{\theta}_\beta c\left((1+t_\beta^1)Y_\beta^1 + \epsilon_\beta - T(\gamma)\frac{(Y_\beta^2 - Y_\alpha^2)}{2} - G_\beta\right)\right).
\end{aligned}
\tag{136}
$$

$\square$

Proof of Proposition 8. From the first order condition, it immediately follows that the solution is $\hat{\theta}_i = \theta_i^m$, whenever $\frac{\partial \bar{t}_f^*(\hat{\theta}_\phi, \bar{\theta})}{\partial \hat{\theta}_i} = 0$, and this is true if and only if region i is not part of the coalition, i.e. $\bar{\theta} \neq \hat{\theta}_i$.

Note that (median) voters are aware of the importance of their choice of $\hat{\theta}$, in that it will determine whether their region is part of the coalition when the formateur represents a region of the unique type (i.e. the opposite type compared to both their types).

Suppose, without loss of generality, that $\bar{\theta}_1 < \bar{\theta}_2$. Then, by Proposition 1, either we have $\max\{\theta_1^m; \theta_2^m\} < \bar{\theta}_1 < \bar{\theta}_2$, or $\theta_1^m < \bar{\theta}_1 < \theta_2^m < \bar{\theta}_2$. Finally, notice that, for each region i, preferences are single-peaked around:

- $\bar{\theta}_i$, conditional on being part of the coalition,

- θ_i^m, conditional on being left out of the coalition.

Start from a candidate equilibrium in which region 1 chooses $\hat{\theta}_1 = \bar{\theta}_1$. Then region 2 will be left out of the coalition for any $\hat{\theta}_2 > \bar{\theta}_1$, in which case, conditional on remaining out of the coalition, the best response will always be to select $\hat{\theta}_2$ to be as close as possible to θ_2^m. This implies choosing $\hat{\theta}_2 = \theta_2^m$ if $\theta_2^m > \bar{\theta}_1$. However, if $\theta_2^m \leq \bar{\theta}_1$, the median voter in region 2 may choose $\hat{\theta}_2 = \bar{\theta}_1$ and remain out of the coalition, or may decide to undercut and choose $\hat{\theta}_2 < \bar{\theta}_1$, in order to get closer to θ_2^m. However, by doing so, region 2 would become part of the coalition, in which case they would want to move as close as possible to $\bar{\theta}_2$, which implies to increase $\hat{\theta}_2$, and they can do so up to $\hat{\theta}_2 = \bar{\theta}_1$. Consequently, any deviation above or below $\bar{\theta}_1$ is not profitable.

Suppose, instead, that we start from a candidate equilibrium in which region 2 chooses $\hat{\theta}_2 = \bar{\theta}_2$. Then region 1 will be left out of the coalition for any $\hat{\theta}_1 > \bar{\theta}_2$. In such cases, conditional on remaining out of the coalition, the best response will always be to select $\hat{\theta}_1$ to be as close as possible to θ_1^m. This implies choosing $\hat{\theta}_1 = \bar{\theta}_2$ (remember that $\theta_1^m < \bar{\theta}_1 < \bar{\theta}_2$). Suppose that region 1 decides to move further close to θ_1^m, the moment they move below $\bar{\theta}_2$, region 1 is selected to form the coalition and will as a result want to move as close as possible to $\bar{\theta}_1$. Because $\hat{\theta}_2 = \bar{\theta}_2$, then region 1 can choose $\hat{\theta}_1 = \bar{\theta}_1$ and be part of the coalition, which is what is optimal for them. Then region 1 will choose $\hat{\theta}_1 = \bar{\theta}_1$ and we are back to the situation previously analysed, where we showed that region 2 will choose $\hat{\theta}_2 = \begin{cases} \theta_2^m, & \text{if } \theta_2^m > \bar{\theta}_1 \\ \bar{\theta}_1, & \text{if } \theta_2^m \leq \bar{\theta}_1. \end{cases}$ and this concludes the proof.

$\square$

Appendix C   Additional Empirical Results

Table 4: Summary Statistics

Variable 5-years window	Mean	Std. Dev.	N
DiffEU-Nat	-0.527	4.815	639
DiffEU-Nat, Winsorised .01	-0.503	4.542	639
Extreme (Algan et al.)	0.15	0.358	639
Extreme (Chapel Hill)	0.122	0.328	639
Eurosceptic	0.163	0.369	639
Party Size	9.815	11.972	639
Leading Party	0.116	0.32	639
% Net Contribution/GNI (2000-2015)	0.624	1.175	639
Net Contribution Absolute Values (2000-2015) (Billion Euros)	-0.581	4.041	639
Extreme Left	0.08	0.271	639
Extreme Right	0.07	0.256	639
Variable 3-years window	**Mean**	**Std. Dev.**	**N**
DiffEU-Nat	-0.341	4.301	485
Extreme (Algan et al.)	0.142	0.35	485
Extreme (Chapel Hill)	0.122	0.327	485
Party Size	9.654	12.083	485
Leading Party	0.113	0.317	485
Variable 10-years window	**Mean**	**Std. Dev.**	**N**
DiffEU-Nat	-0.202	3.919	516
Extreme (Algan et al.)	0.141	0.349	516
Extreme (Chapel Hill)	0.11	0.314	516
Party Size	7.924	10.581	516
Leading Party	0.081	0.274	516

Table 5: Extreme Voting at EU vs National Elections: Winsorised Dependent Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme (Algan et. al)	1.248***	1.253***	1.247***	0.729*				
	(0.357)	(0.358)	(0.423)	(0.408)				
Party Size				-0.103***				-0.105***
				(0.0304)				(0.0304)
Leading Party				-0.296				-0.266
				(1.222)				(1.221)
Extreme (Chapel Hill)					1.274***	1.284***	1.188***	0.910**
					(0.399)	(0.401)	(0.450)	(0.442)
Observations	639	639	639	639	639	639	639	639
R-squared	0.010	0.010	0.074	0.165	0.008	0.009	0.072	0.166
Country FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in a party votes share between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy equals to 1 for extreme parties; *Party Size* is a continuous variable measuring a party vote share; *Leading Party* is a dummy set equal to one for the main incumbent party. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure 6: Differences in Voting between EU and National Elections across Moderate and Extreme Parties: 3-years window
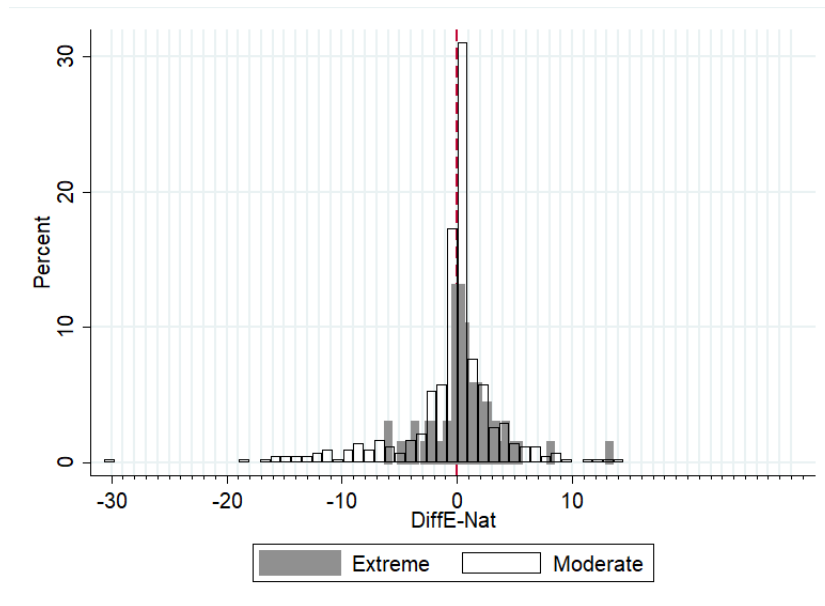
The figure shows the distributions of votes for parties at European and at National elections in the period 1990-2013, differentiating between moderate and extreme parties. The analysis is based on a 3-years window.

Figure 7: Differences in Voting between EU and National Elections across Moderate and Extreme Parties: 10-years window

The figure shows the distributions of votes for parties at European and at National elections in the period 1990-2013, differentiating between moderate and extreme parties. The analysis is based on a 10-years window.

Table 6: Extreme Voting at EU vs National Elections: 3-years time window

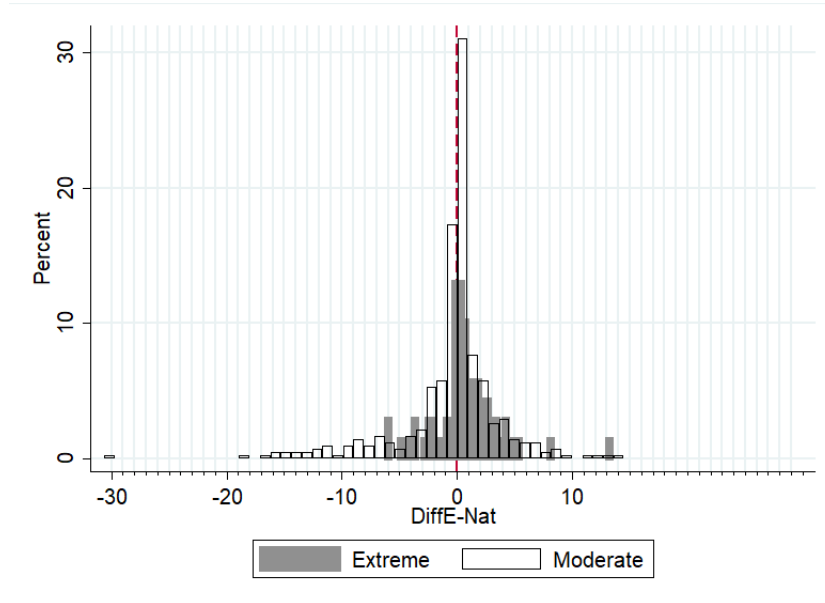|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme (Alg. et al)	1.199***	1.206***	1.581***	0.966**				
	(0.416)	(0.417)	(0.493)	(0.472)				
Party Size				-0.137***				-0.138***
				(0.0349)				(0.0349)
Leading Party				1.348				1.379
				(1.355)				(1.357)
Extreme (Chapel Hill)					1.298***	1.307***	1.613***	1.155**
					(0.462)	(0.461)	(0.528)	(0.509)
Observations	485	485	485	485	485	485	485	485
R-squared	0.009	0.012	0.052	0.152	0.010	0.012	0.051	0.154
Country FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in a party votes share between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy equals to 1 for extreme parties; *Party Size* is a continuous variable measuring a party vote share; *Leading Party* is a dummy equals to one for the main incumbent party. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix D Additional Empirical Results II: A Larger Federation

EU integration was a staggered process which started out from a small set of just six countries, to end up 60 years later as a union of 28 member states. This evolution can be interpreted

Table 7: Extreme Voting at EU vs National Elections: 10-years time window

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme (Alg. et al)	0.977**	0.980**	1.246**	0.912*				
	(0.426)	(0.428)	(0.496)	(0.475)				
Party Size				-0.114***				-0.117***
				(0.0291)				(0.0291)
Leading Party				-0.0518				-0.0214
				(1.403)				(1.403)
Extreme (Chapel Hill)					0.890*	0.897*	1.086*	0.982*
					(0.479)	(0.481)	(0.568)	(0.541)
Observations	516	516	516	516	516	516	516	516
R-squared	0.008	0.009	0.046	0.146	0.005	0.007	0.041	0.145
Country FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in a party votes share between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy equals to 1 for extreme parties; *Party Size* is a continuous variable measuring a party vote share; *Leading Party* is a dummy set equal to one for the main incumbent party. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Extreme Left-Right Voting at EU vs National Elections

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme Left	1.219***	1.218***	1.193**	0.690				
	(0.404)	(0.404)	(0.495)	(0.480)				
Party Size				-0.112***				-0.113***
				(0.0342)				(0.0342)
Leading Party				-0.172				-0.150
				(1.313)				(1.313)
Extreme Right					1.122**	1.134**	1.181*	0.638
					(0.525)	(0.529)	(0.623)	(0.646)
Observations	639	639	639	639	639	639	639	639
R-squared	0.005	0.005	0.066	0.155	0.004	0.004	0.065	0.155
Country FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	YES	YES	YES	NO	YES	YES	YES
Country-Time FE	NO	NO	NO	YES	NO	NO	NO	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in a party votes share between European and National elections); *Extreme left/right* (Algan et al.) is a dummy equals to 1 for extreme left/right parties; *Party Size* is a continuous variable measuring a party vote share; *Leading Party* is a dummy equals to one for the main incumbent party. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

in the light of Section A.1, where we investigate the potential role of beliefs over political effectiveness. More specifically, we allowed agents to assign region-specific probabilities to the fact that their elected politician will be able to impact federal policy-making in a certain period, which in turn affects the intensity of their strategic behaviour. This could occur when voters believe the political weight of their region – within the federal coalition – is such that their politician is more, or less, able to influence federal redistribution. Another

Table 9: Extreme Voting at EU vs National Elections: Net Contributors and Net Receivers (Absolute Contribution)

	(1)	(2)	(3)	(4)
Extreme (Algan et al.)	2.183**	1.861**		
	(0.888)	(0.827)		
Extreme Alg*2^{nd}thirdile-Abs.	-1.845	-2.327**		
	(1.175)	(1.174)		
Extreme Alg*3^{rd}thirdile-Abs.	-0.717	-0.992		
	(1.087)	(1.069)		
Extreme (Chapel Hill)			2.184**	1.862**
			(0.888)	(0.827)
Extreme CH*2^{nd}thirdile-Abs.			-2.589**	-2.635**
			(1.209)	(1.218)
Extreme CH*3^{rd}thirdile-Abs.			-0.382	-0.282
			(1.033)	(0.995)
Observations	639	639	639	639
R-squared	0.073	0.161	0.074	0.163
Country FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Country-Time FE	NO	YES	NO	YES
Controls	NO	YES	NO	YES

Note: The Table reports OLS coefficients and Robust Standard errors in brackets. The dependent variable is *DiffEU-Nat* (i.e. the difference in a party votes share between European and National elections); *Extreme* (Algan et al. or Chapel Hill) is a dummy equals to 1 for extreme parties; *Tertiles* split countries in three groups based on their level of net EU transfers (Absolute values). The omitted category is the first tertiles (most net contributors). *** p<0.01, ** p<0.05, * p<0.1.

possibility would be when voters perceive the chances to be included in the ruling coalition as becoming slimmer, the more the set of member states expands. Applied to our setting of EU enlargement, both interpretations imply that voters will behave less strategically since they expect the influence of their representatives to decline. However, it could also be that EU political processes are perceived as potentially stacked in favour of the initial member states, as they enjoyed quite some leverage when the initial rules were set. Knowing the ins and outs of EU policy making therefore, the perceived probability for these representatives to steer EU budget decisions could come out reinforced post-expansion. Consequently, we expect the effects of EU enlargement to be ambiguous.

To bring this prediction to the data, we consider the 2004 discontinuity when the EU moved from 15 to 25 (and then 27 in 2007) countries: the biggest expansion in European history, which, indeed, took place in the middle of our sample period. We can, then, evaluate whether strategic voting in countries already in the EU before 2004 has changed after the 2004 enlargement. Fig. 8 reports the average values of our dependent variable (*DiffEU-Nat*) for extreme and Eurosceptic parties before/after 2004. The left panel does not show any variation before/after 2004, in terms of relative electoral success of extreme parties at the EU elections. Conversely, we observe a relative higher support for Eurosceptic parties at the EU elections after 2004. This provides some support for the second hypothesis, whereby voters

support even more extreme representatives after a federal expansion, reacting to an increased distance between local and federal preferences.

Figure 8: Relative Change Before/After 2004 in Voting for Extreme/Eurosceptic Parties