Risk-sharing with a central authority: Free-riding, lack of commitment, and bail-outs*

Matthias Kredler[†] & Rubén Veiga-Duarte[‡]

May 23, 2023

Find the latest version of the paper here

Abstract

Agents with independent risks (*regions*) often create unions and delegate redistributive power to a central institution (*center*) that provides risk-sharing through costly transfers. However, there is the risk that regions free-ride on each other; this risk may be exacerbated when the center cannot commit to future policies. We study a differential game of two regions that make savings decisions and a benevolent center that sets transfers but lacks commitment. One region always ends up bankrupt and the center provides a *bailout*: as the poor regions enters bankruptcy, there is an upward jump in transfers to the poor region that coincides with a downward jump in the poor region's consumption. Delegation to a center is Pareto-improving provided that the center's welfare weights reflect initial wealth differences between regions. However, once asymmetries in regions' wealth and the center's welfare weights become large, dynamics become unstable, thus hastening impoverishment and bail-outs.

*We thank Andrés Erosa, Jose Vicente Rodriguez Mora, Almuth Scholl and seminar/workshop audiences at MacCaLM Edinburgh, ENTER Jamboree Tilburg, SAE Simposio Alicante, Univ. of Bristol, LMU Munich and Autónoma Madrid for useful comments. Matthias Kredler acknowledges research funding by the Fundación Ramón Areces; Spanish Agencia Estatal de Investigación (grants RYC2018-025230-I and PID2019-107134RB-I00 / AEI / 10.13039/501100011033), MCIN/AEI /10.13039/501100011033-CEX2021-001181-M (María de Maeztu) and Comunidad de Madrid, grant EPUC3M11 (V PRICIT). Rubén Veiga acknowledges PhD funding by Spanish Ministerio de Economía y Competitividad (Ref. BES-2016-077176). An older version of this paper was part of Rubén Veiga PhD thesis *Essays on Risk-sharing, Default and Fixed-Term Contracts*. The views expressed in this paper are those of the authors and do not necessarily represent the views of Banco de España or the Eurosystem.

[†]Department of Economics at Universidad Carlos III de Madrid. Email: mkredler@eco.uc3m.es

[‡]Banco de España. Email: ruben.veiga@bde.es

1 Introduction

There is a number of economic situations in which agents face independent risks, giving rise to welfare gains from risk-sharing through transfers from lucky to unlucky agents. However, the prospect of such transfers usually entails moral hazard: agents may under-invest, or provide low effort, when promised insurance transfers. In the face of such situations, agents often create unions and delegate re-distributive power to a central institution (*center*). Some examples are i) the central government in fiscal federation of regions (our main application), ii) the chief in a village of farmers, iii) the household head in a family, or iv) the headquarter of a holding firm. While federal countries will be our main application in this article, it should be obvious how to interpret our results in other contexts.

Within federal states, and also within political unions of countries such as the European Union, there exists a tension between fiscal decentralization and sustainable borrowing. When regions have fiscal independence, they have incentives to over-borrow and to free-ride on the other members of the union. After all, the center will be tempted to bail out regions in poor financial health. Recognizing these perverse incentives, centers often announce that there won't be future bail-outs and create rules that are supposed to rule them out. However, when push comes to shove, such pre-commitments are often scrapped and regions in trouble are eventually rescued.

In this paper, we provide a dynamic model of federations that captures these interactions. We study a differential game of two regions that make savings decisions and a central authority that sets (costly) transfers between the regions. We reduce the dimensionality of the state space to the share of wealth held by the first region, exploiting homogeneity. We characterize how different transfer schedules distort the regions' savings decisions and lead to inefficiency in Markov-Perfect Equilibria.

We find that one region always ends up bankrupt and the center provides what resembles a *bailout* (our main result): as the poor region enters bankruptcy, there is an upward jump in transfers to the poor region that coincides with a downward jump in the poor region's consumption. This allocation is inefficient, but it is the only stable arrangement close to the poor region's constraint. We provide a novel cost-benefit interpretation of the value-matching conditions at the constraint in a *limit game*. The limit-game analysis shows that for the poor region, a discontinuous consumption path is optimal since the center's bail-out adds an additional disincentive to save, making earlier bankruptcy optimal. For the center, an upward jump in transfers at the bail-out is optimal for the following reason: Before the bail-out, the center takes the poor region's consumption as given; once the poor region hits the constraint, however, the center controls the poor region's consumption, thus giving it an additional motive to tax the rich region and prop up the poor region's consumption. In the center's eyes, the poor region consumes too much just before the bail-out, providing a further motive to hasten the bail-out by reducing transfers.

Our setting provides a theory of fiscal delegation. In a numerical exercise, we ask under which conditions regions would prefer to delegate power to a center over staying in autarky. We find that for any distribution of initial wealth between regions, there exists a range of center's welfare weights that make it optimal for *both* regions to join the union.

The more unequal the initial wealth distribution, the more the center's weights have to tilt towards the rich/large region. This is against conventional wisdom, since it suggests that even in very asymmetric situations delegation to a center can be Pareto-improving. However, we also find that a tilted center is detrimental to stability in the following sense. When the center's weights on regions are equitable, there are stable dynamics on a large interval of the state space and wealth shares tend to equalize. However, this interval with stable dynamics disappears once the center's weight tilts too much in favor of one region, thus making impoverishment and a quick bail-outs more likely. Also, the inefficiency at the bail-out is exacerbated: we find that both i) percentage decrease in the poor region's consumption and ii) the percentage increase in transfers at the bail-out are higher the lower the center's weight on the small/poor region is.

We carry out our analysis in continuous time, which has the main advantage that we can use first-order conditions for the timing decision of the bailout. It should be mentioned here that the usefulness of continuous-time games has been challenged by Simon and Stinchcombe (1989); in a nutshell, the reason is that there is no notion of a ”last period”, e.g. when constructing trigger strategies. Since we consider Markovian strategies as a function of a joint state, the same line of defense applies as in Sannikov (2007), who argued that the pitfalls of continuous-time games are circumvented by studying public perfect equilibria (in which players' strategies are functions of the public history of the game).

The literature in political economy has studied dynamic models of federations, but has focused on the allocation of federal spending to different goods rather than on regions' savings decisions. Prominent examples are Battaglini and Coate (2007) and Battaglini and Coate (2008), who study a dynamic political-economy model in which a central legislature chooses expenditures on a productive public good versus inefficient pork-barrel policies.

Furthermore, a recent literature on borrowing crises in currency unions relates to our work. Our theory shares with this literature the prediction that regions have incentives to under-save. Whereas this literature focuses on monetary policy (which we are silent upon), our work puts the center's commitment problem center-stage. Aguiar, Amador, Farhi, and Gopinath (2015) find, as we do, that regions under-save, providing a rationale for debt ceilings. Similar results were obtained earlier by Chari and Kehoe (2007) and Chari and Kehoe (2008). Farhi and Werning (2017) show that there is a role for fiscal intervention by the center (as there is in our setting), but they do not investigate lack of commitment.

The literature on endogenous sovereign default has recently shown interest in bail-outs, specifically in IMF interventions. Fink and Scholl (2016) find that IMF bailouts have avoided sovereign defaults in the short-run but at the cost of increased sovereign risk in the future. They model bail-outs as an exogenously-given arrangement, whereas we endogenize the center's decision. Also, we point out that in their model, a *bailout* means an alternative source of financing with some conditionality attached, whereas we define it as a jump discontinuity in transfers. Other papers on this front include Roch and Uhlig (2018) and Boz (2011).

Finally, our model can be seen as is yet another instance of a second-best risk-sharing mechanism. The seminal contributions in this literature are Thomas and Worrall (1988) and Kocherlakota (1996), who showed that lack of commitment entails incomplete insurance. Ábrahám and Laczó (2018) extend the basic setting by a public storage technology, thus studying a similar environment to ours. We find that delegation to a center provides more insurance than is possible in Ábrahám and Laczó (2018): Our center can provide insurance to an agent who is stuck with zero output forever, whereas limited commitment rules out insurance in this situation (since the rich agent would leave the contract). Similarly related is Attanasio and Pavoni (2011), who study market insurance under hidden effort and hidden savings. Our results, as theirs, imply excess smoothness of consumption (i.e. agents' consumption does not react one-for-one to permanent-income shocks). However, in their setting storage is not used by agents in equilibrium, whereas in our setting it is. Finally, our bail-out result, accompanied by an inefficient consumption path of the poor region, is echoed in the altruism models of Barczyk and Kredler (2014a) and Barczyk and Kredler (2014b), who study a model of two agents with altruistic preferences that provide transfers to each other.

The rest of the article is structured as follows. Section 2 introduces our model and Section 3 establishes some preliminary results that make the model tractable. Section 4 studies the case of costly redistribution and establishes the main result on bail-outs. Section 5 studies the case in which redistribution is costless. Section 6 concludes, discussing potential extensions to quantitative models.

2 Model

2.1 Physical environment

We consider an economy with three agents in continuous time, $t \in [0, \infty)$. There are two *regions* and a *center*.[1] We denote variables concerning Region 1 by plain letters (a_t, c_t etc.) and variables concerning Region 2 with primes (a'_t, c'_t etc.).

Laws of motion for regions' assets. Each region possesses a stock of a productive resource (the *asset*), denoted by a_t and a'_t respectively.[2] Assets grow at an exogenous rate $\rho > 0$. At each t, the region decides on its flow consumption, $c_t \geq 0$, and is required to make a transfer flow, τ_t, which the center sets. The transfer τ_t can be positive, in which case it amounts to a tax, or negative, in which case it amounts to a subsidy. Furthermore, there are Brownian shocks to the asset, giving rise to the flow budget constraints

$$\mathrm{d}a_t = (\rho a_t - c_t - \tau_t)\mathrm{d}t + s(a_t, a'_t)\mathrm{d}B_t, \quad [2.1]$$

$$\mathrm{d}a'_t = (\rho a'_t - c'_t - \tau'_t)\mathrm{d}t + s(a'_t, a_t)\mathrm{d}B'_t, \quad [2.2]$$

where B_t and B'_t are uncorrelated Wiener Processes. The function $s(\cdot, \cdot)$, specified below, determines the volatility of the shocks. The regions are subject to the natural borrowing

[1] Our language here always refers to the example of a fiscal federation; it should be clear how to apply the model to other contexts, such as those mentioned in Section 1.

[2] In reality, we interpret a_t as the entirety of the region's productive resources plus assets minus debt.

limits

$$a_t \geq 0, \qquad a'_t \geq 0 \qquad \text{for all } t. \tag{2.3}$$

Volatility. We assume that volatility is a proportional to a geometric average between the region's own wealth and total wealth in the economy, $A_t \equiv a_t + a'_t$, i.e.

$$s(a, a') = \sigma\sqrt{a}\sqrt{a + a'}, \tag{2.4}$$

where $\sigma > 0$ is a parameter. This specification has two desirable features. First, it removes portfolio effects in the social planner's problem: A planner facing this type of shocks has no preference over how total wealth is allocated between the two regions. This is because aggregate risk only depends on total wealth A_t in the economy, but not on how it is distributed, since $\mathrm{Var}(\mathrm{d}A_t) = \mathrm{Var}\left(\sigma\sqrt{a_t}\sqrt{A_t}\mathrm{d}B_t + \sigma\sqrt{a'_t}\sqrt{A_t}\mathrm{d}B'_t\right) = \sigma^2 A_t^2 \mathrm{d}t.$. A second advantage of our specification is that it removes any systematic drift that shocks have on wealth shares of the two regions, as will be evident from Eq. [3.4].[3]

Redistribution by the center. To simplify our analysis, we assume that the center cannot save; it only re-distributes resources between the two regions. Thus at each t, one out of τ_t and τ'_t is positive and the other negative. To allow for costs to this re-distributive process, we specify the center's flow budget constraint as

$$X\left(\frac{\tau_t}{Y_t}\right)Y_t + X\left(\frac{\tau'_t}{Y_t}\right)Y_t = 0, \tag{2.5}$$

where $Y_t \equiv \rho A_t$ denotes total output in the economy and where $X(\cdot)$ is a twice differentiable function that satisfies $X(\tilde{\tau}) = \tilde{\tau}$ for all $\tilde{\tau} \leq 0$, $X'(0) = 1$, and $X''(\tilde{\tau}) \leq 0$ for all $\tilde{\tau} > 0$. The form of the budget constraint [2.5] entails that re-distributing the same fraction of output costs the same share of output irrespective of the size of the economy.

In our main application (fiscal federations), it is reasonable to assume that redistribution is costly, as there are dead-weight losses to taxation (i.e. a Laffer Curve) and political costs to re-distribution between regions. Our baseline assumption, invoked in Section 4, is thus

Assumption 1 (Costly redistribution)**.** $X''(\tilde{\tau}) < 0$ *for all* $\tilde{\tau} > 0$ *and* $\lim_{\tilde{\tau}\to\infty} X'(\tilde{\tau}) = 0$.

Section 5 briefly summarizes our results under the alternative assumption that transfers are costless; the mathematicals details for this case are given in the Appendix 5.[4]

[3] We have also considered generalized geometric averages, i.e. $s(a, a') = \sigma a^{\xi}(a + a')^{1-\xi}$ for $\xi \in [0, 1]$. The case $\xi = 1$ corresponds to (a_t, a'_t) following geometric Brownian Motions. Many of our results still go through. However, the construction of *smooth equilibria* mentioned in Section 5 fails since these rely on the center being indifferent about the distribution of wealth.

[4] We allow for mass transfers (i.e. a transfer that generate a discontinuity in the asset's trajectory) when redistribution is costless. Standard arguments imply that the center should not use mass transfers when redistribution is costly.

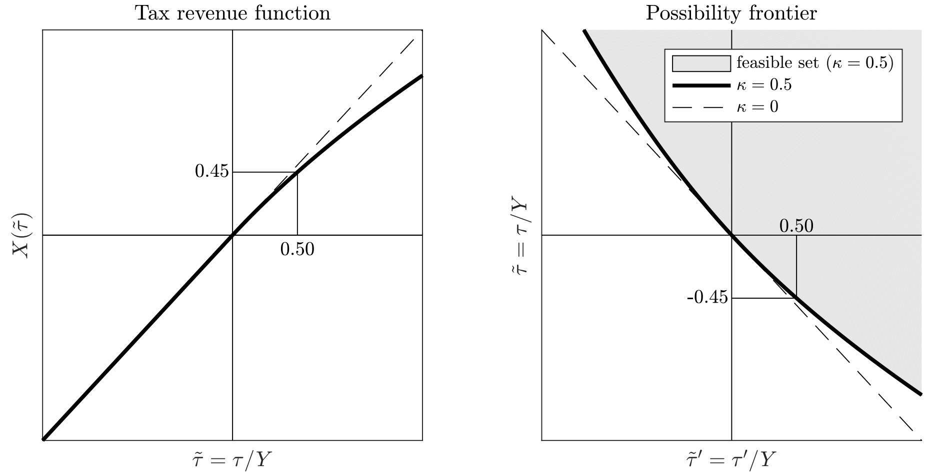

Figure 1: Tax distortions

In our numerical examples, we use the following parametric specification:

$$X(\tilde{\tau}) = \begin{cases} \tilde{\tau} & \text{if } \tilde{\tau} \leq 0, \\ \frac{(1+\alpha\tilde{\tau})^{1-\kappa}-1}{\alpha(1-\kappa)} & \text{if } \tilde{\tau} > 0. \end{cases} \qquad [2.6]$$

Here, $\kappa \in [0, 1)$ governs the curvature of $X(\cdot)$ and $\alpha > 0$ determines how fast revenues decrease when taxes increase. Figure 1 plots the function $X(\tilde{\tau})$ and shows the possibility frontier that $X(\cdot)$ induces for different tax-subsidy combinations. The plots are in terms of tax shares $\tilde{\tau} \equiv \tau/Y$ and $\tilde{\tau} \equiv \tau'/Y$, i.e. they depict transfers as a fraction of aggregate output. For $\kappa = 0$, there are no distortions and taxes τ_t are transformed one-for-one into subsidies to the other region, i.e. $-\tau_t' = \tau_t$, corresponding to the 45-degree lines in the graphs. For $\kappa > 0$, Ass. 1 holds and resources are lost in the redistributive process; the higher taxes are, the lower is the marginal revenue, $X'(\tilde{\tau}) = (1+\alpha\tilde{\tau})^{-\kappa}$. In the example in the graph, imposing a tax of 50% of total output on a region results leaves 45% of output to give to the other region. The decreasing marginal return to taxation is also reflected in the convex shape of the possibility frontier in the right graph.

Preferences. All agents maximize expected utility and discount at the common rate ρ.[5] Regions' instantaneous utility over consumption is assumed to be $u(c) = \ln(c)$.[6] [7] The center has instantaneous utility

$$u^c(c, c') = \mu u(c) + (1 - \mu)u(c'), \qquad [2.7]$$

[5] For parsimony, we assume that the return to assets equals agents' discount rates. Our results also go through when specifying a return $r \neq \rho$: Equilibrium strategies do not change, since substitution and income effects cancel when preferences are logarithmic and income is linear in the state variable.

[6] The choice of log-utility is not essential for our analysis; what *is* essential is that utility is (i) increasing, (ii) concave, (iii) homothetic, and that (iv) it satisfies the Inada condition $\lim_{c \to 0} u'(c) = \infty$.

[7] For simplicity, we assume that each region consists of a measure 1 of households. It is easy to accommodate regions of varying size, re-parameterizing the weight $\mu = [\mu_N N]/[\mu_N N + (1 - \mu_N)(1 - N)]$, where N and N' are the measure of households in each region and μ_N is the weight that planner puts on Region 1 household.

where $\mu \in [0,1]$ is an exogenously-given weight the center places on Region 1. While we specify the center's criterion in an ad-hoc fashion, some political-economy models with probabilistic voting endogenously generate decision-making according to a utilitarian welfare criterion as in [2.7], see the overview by Persson and Tabellini (2002), Section 9.

Timing and information structure. We assume the following timing protocol within an instant $[t, t + \mathrm{d}t)$:

1. (a) All agents observe current asset positions (a_t, a'_t).
   (b) The center sets transfers (τ_t, τ'_t) subject to its budget constraint [2.5]. The center cannot tax broke regions, i.e. we require that $\tau_t \leq 0$ if $a_t = 0$ and $\tau'_t \leq 0$ if $a'_t = 0$. Both τ_t and τ'_t are public information.
2. (a) Regions' interim wealth is then: $\hat{a}_t = a_t - \tau_t \mathrm{d}t$ and $\hat{a}'_t = a'_t - \tau'_t \mathrm{d}t$.
   (b) Regions choose consumption rates c_t and c'_t (observable only by themselves), subject to the requirement that their consumption expenditures over the interval do not exceed interim wealth:

$$c_t \mathrm{d}t \in [0, \hat{a}_t] \quad , \quad c'_t \mathrm{d}t \in [0, \hat{a}'_t]. \qquad [2.8]$$

3. Finally, flow utility is collected by all agents, returns to the assets accrue, and wealth shocks are realized, giving rise to the following asset positions in the next instant:

$$a_{t+\mathrm{d}t} = a_t + (\rho a_t - c_t - \tau_t)\mathrm{d}t + s(a_t, a'_t)\epsilon_t \sqrt{\mathrm{d}t},$$
$$a'_{t+\mathrm{d}t} = a'_t + (\rho a'_t - c'_t - \tau'_t)\mathrm{d}t + s(a'_t, a_t)\epsilon'_t \sqrt{\mathrm{d}t}, \qquad [2.9]$$

   where $\epsilon_t, \epsilon'_t \sim N(0,1)$ are standard normal shocks that are independent across regions and time. We assume that ϵ_t is only observable to Region 1 and that ϵ'_t is only observable to Region 2.

We study the limit of the game as $\mathrm{d}t$ approaches zero. It is important to note here that as $\mathrm{d}t \downarrow 0$, a region will only face a restriction on its consumption flow once its beginning-of-period asset stock a_t is exactly zero. A region with a positive asset stock $a_t > 0$ that faces a fixed transfer flow $\tau_t = \bar{\tau} > 0$, can choose the consumption flow c_t arbitrarily large as we let $\mathrm{d}t \to 0$, as can readily be seen from [2.8].[8] If $a_t = 0$, however, the center can restrict Region 1's consumption flow, i.e. it must hold that $c_t \mathrm{d}t \leq -\tau_t \mathrm{d}t$. This point in the state space will be of particular interest since it is the only state in which the center can exert direct control on a region's consumption. In summary, the restrictions on the agents' actions implied by our timing protocol are:

$$\begin{aligned} &\text{If } a_t = 0: && \tau_t \leq 0 && \text{and} && c_t \leq -\tau_t, \\ &\text{if } a'_t = 0: && \tau'_t \leq 0 && \text{and} && c'_t \leq -\tau'_t. \end{aligned} \qquad [2.10]$$

[8] This is not true any more if the center removes a lump sum from the stock a_t, as can occur when redistribution is costless; in this case the term $\tau_t \mathrm{d}t$ does not vanish.

2.2 Agents' problems and equilibrium definition

Given the information structure just laid out, the center faces an inference problem that is familiar from principal-agent settings with moral hazard: A bad outcome $a_{t+\mathrm{d}t}$ for Region 1 may have come about by reckless behavior (Region 1 choosing a large c_t), but may just as well be the result of bad luck (a low realization of $\mathrm{d}B_t$). It is thus natural to assume that the center can only condition its actions on the current asset positions (a_t, a_t') in the first stage. Formally, we restrict attention to Markov-perfect equilibria, i.e. equilibria in which the center can only condition its strategy on the payoff-relevant state (a_t, a_t').[9] Also, since regions' feasible sets are affected by the center's transfers once they are broke, region's policies must be made contingent both on (a_t, a_t') *and* on the center's action in the first stage, i.e. c and c' may in general be functions of $(a_t, a_t', \tau_t, \tau_t')$. In Section 3.2 we derive results that allow us to restrict attention to (a_t, a_t') when both regions have positive wealth.

We are now prepared to write down the problem faced by the regions and the center:

Region's problem. Given transfer policies $\{\tau(a_t, a_t'), \tau'(a_t, a_t')\}$ by the center and a consumption policy $c'(a_t, a_t'; \tau_t, \tau_t')$ by Region 2, Region 1 chooses a consumption rule $c(a_t, a_t'; \tau_t, \tau_t')$ to solve:

$$v(a_0, a_0'; \tau_0, \tau_0') = \max_{c(\cdot)} \quad \mathbb{E}_0 \int_0^\infty e^{-\rho t} u\left(c(a_t, a_t'; \tau_t, \tau_t')\right) \mathrm{d}t \qquad [2.11]$$
$$\text{s.t.} \quad [2.1], [2.2], \text{ and } [2.10], , \qquad a_0, a_0' \text{ given.}$$

First, note here that we do not impose the borrowing limit $a_t \geq 0$ as a separate state constraint since it is implied by the policy constraint [2.10]. Second, note that we require the region to specify an optimal policy in response to *any* possible policy (τ_t, τ_t') that the center may choose in the first stage of the game in each instant, i.e. we require Region 1 to specify its best response in each subgame in Stage 2 of the instantaneous game; this will turn out to be key to have well-defined best-response problem for the center. Region 2's problem is analogous to [2.11], making the obvious adjustments.

Center's problem. Given regions' policy rules $\{c(a, a'; \tau, \tau'), c'(a, a'; \tau, \tau')\}$, the center chooses policies $\{\tau(a, a'), \tau'(a, a')\}$ to solve

$$w(a_0, a_0') = \max_{\{\tau(\cdot), \tau'(\cdot)\}} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u^c\big(c(a_t, a_t'; \tau_t, \tau_t'), c'(a_t, a_t'; \tau_t, \tau_t')\big) \mathrm{d}t \qquad [2.12]$$
$$\text{s.t.} \quad [2.1], [2.2], [2.5] \text{ and } [2.10], \qquad a_0, a_0' \text{ given.}$$

Equilibrium definition. The definition of equilibrium is then:

Definition 1. *A* Markov-perfect equilibrium *(MPE) is a set of value functions* $\{v, v'; w\}$ *and policy functions* $\{c, c'; \tau, \tau'\}$ *such that:*

[9] We note here that our results also characterize the Markov-perfect equilibria for the full-information environment, i.e. one where the center is able to observe regions' past consumption decisions.

1. *Given federal policies* $\{\tau, \tau'\}$ *and Region 2's policy* c', $\{c, v\}$ *solve Region 1's problem,* [2.11].

2. *Given* $\{\tau, \tau'\}$ *and* c, $\{c', v'\}$ *solve Region 2's problem, i.e. the analog of* [2.11] *for Region 2.*

3. *Given regions' policies* $\{c', c\}$, $\{\tau, \tau', w\}$ *solve the federal government's problem,* [2.12].

In general, we restrict attention to equilibria with value functions that are smooth (at least twice differentiable) on the interior of the state space, which is natural since Brownian Motion is a strong smoother. We do allow for discontinuous policies at the borrowing constraints, however, where the shock volatility is zero.

3 Preliminary results

In this section, we derive some preliminary results that are essential to make the problem tractable. The first result allows us to reduce the state to one variable under a mild homogeneity assumption on strategies, invoking homothetic utility and linear returns to assets. The second result allows us to focus on the Stage-1 state (a, a') instead of interim assets $(\hat{a}, \hat{a}')$ in Stage 2 of the instantaneous game, at least in the interior of the state space. We then proceed to study two important benchmark allocations (efficiency and wealth-pooling), showing that efficient allocations cannot be supported generically, even when the center can commit to policies.

3.1 Exploiting homogeneity

Change of variables. We first define a new pair of state variables,

$$A_t = a_t + a'_t \quad , \quad P_t = \frac{a_t}{a_t + a'_t}. \tag{3.1}$$

Here, A_t is total wealth in the economy and P_t is the wealth share of Region 1. Furthermore, we express policies as shares out of total wealth, defining

$$C_t = c_t/A_t, \quad C'_t = c'_t/A_t, \qquad T_t = \tau_t/A_t, \quad T'_t = \tau'_t/A_t. \tag{3.2}$$

Applying the Itô Rule to the mapping defined in [3.1] and invoking the definitions in [3.2], the law of motion for the new state variables becomes

$$\frac{\mathrm{d}A_t}{A_t} = (\rho - C_t - C'_t - T_t - T'_t)\mathrm{d}t + \sigma \underbrace{\left[\sqrt{P_t}\mathrm{d}B_t + \sqrt{1-P_t}\mathrm{d}B'_t\right]}_{\equiv \mathrm{d}B_t^A}, \tag{3.3}$$

$$\mathrm{d}P_t = \left[P_t(C'_t + T'_t) - (1-P_t)(C_t + T_t)\right]\mathrm{d}t + \sigma\sqrt{P_t(1-P_t)}\underbrace{\left[\sqrt{1-P_t}\mathrm{d}B_t - \sqrt{P_t}\mathrm{d}B'_t\right]}_{\equiv \mathrm{d}B_t^P}. \tag{3.4}$$

It is worthwhile to pause here to gain a better grasp of these equations. We first comment on the drift terms (those involving $\mathrm{d}t$): i) The growth rate of total wealth, $\mathrm{d}A_t/A_t$, is reduced one-for-one by the consumption shares of both regions, C_t and C'_t, as well as by the deadweight loss of taxation, $T_t + T'_t$. ii) A high consumption share or a high tax rate on Region 2 (C'_t, T'_t) increase Region 1's wealth share (P_t), whereas the opposite is true for Region 1's consumption share and tax rate (C_t/ T_t). iii) Terms relating to the second derivatives of the function $P(a, a')$ cancel out due to our particular choice for the volatility function $s(\cdot, \cdot)$, thus corroborating our previous claim that shocks do not introduce any systematic drift in regions' wealth shares.

We now turn to the shocks in [3.3] and [3.4], i.e. terms involving $\mathrm{d}B_t$ and $\mathrm{d}B_t$. We note that when $P \to 0$, shocks to the rich region –in this case Region 2– mainly move total wealth A_t, but not so much the wealth share P_t. Shocks to the poor region, however, have a stronger impact on P_t than on A_t.[10] Finally, we note that the change of variables transforms the shocks $(\mathrm{d}B_t, \mathrm{d}B'_t)$ from (a, a')-space to a new pair of orthogonal shocks $(\mathrm{d}B^A_t, \mathrm{d}B^P_t)$ in (A, P)-space, as defined in [3.3] and [3.4]. [11] We use a shock process $\mathrm{d}B^P_t$ as an input into our numerical solutions, once we have reduced the state space to P, which is what we now turn to.

Taking a step back and inspecting the right-hand sides of [3.3] and [3.4], we see that both the growth rate of total wealth, $\mathrm{d}A_t/A_t$, and the change in the wealth share, $\mathrm{d}P_t$, are independent of total wealth, A_t, as long as the policies $\{C_t, C'_t, T_t, T'_t\}$ are. This forms the basis of our dimension reduction strategy and motivates:

Definition 2 (A-linear strategies)**.** *$c(\cdot), c(\cdot), \tau(\cdot), \tau'(\cdot)$ are A-linear strategies if there exist functions $C(\cdot), C'(\cdot), T(\cdot), T'(\cdot) : [0, 1] \to \mathbb{R}^+_0$ such that $c(a, a') = C(P)A, \tau(a, a') = T(P)A$, etc. $\forall(a, a')$.*

Intuitively, A-linearity requires that large and small economies look "proportionally alike". Strategies, when expressed as shares of total wealth A_t, should only depend on a region's wealth share, but not on the total size of the economy. We deem this assumption natural, more so in the context of homothetic preferences and the linear production technology.[12] [13]

3.2 Recursive characterization

In this subsection, we recursively characterize agents' problems and derive the Hamilton-Jacobi-Bellman (HJB) and Euler equations. We do so using the transformed state vari-

[10]The same is true for consumption decisions: Since the poor region has lower wealth, its consumption decision has a stronger impact on the wealth ratio.

[11]It can easily be verified that the processes B^A_t and B^P_t are independent Wiener Processes: (i) They have zero drift, $\mathbb{E}_t[\mathrm{d}B^A_t] = \mathbb{E}_t[\mathrm{d}B^P_t] = 0$, (ii) they have unit variance, $(\mathrm{d}B^A_t)^2 = (\mathrm{d}B^P_t)^2 = \mathrm{d}t$, and (iii) their covariance is zero, $(\mathrm{d}B^A_t)(\mathrm{d}B^P_t) = 0$.

[12]However, this restriction is not without loss of generality: It rules out equilibria in which the risk-sharing arrangement changes in nature when A_t rises or falls. An example of a strategy profile that we rule out is as follows: Both regions choose an efficient consumption rate while $A_t \geq \bar{A}$, but they give up cooperation and choose an inefficiently high consumption rate once they observe $A_t < \bar{A}$.

[13]Also note that since otal output, ρA_t, is proportional total wealth, the policies C_t, C'_t etc. map naturally to GDP shares such as consumption/GDP, debt/GDP etc.

ables (A_t, P_t) and invoking A-linear strategies.

Region's HJB. Bellman's Principle for Region 1 at an interior point $P \in (0,1)$, over an infinitesimal amount of time $\mathrm{d}t$, is[14]

$$\tilde{V}(P,A) = \max_{C} \left\{ \ln(CA)\mathrm{d}t + e^{-\rho \mathrm{d}t}\mathbb{E}_t \tilde{V}(P_{t+\mathrm{d}t}, A_{t+\mathrm{d}t}) \right\}, \qquad [3.5]$$

where $\tilde{V}(P,A)$ is the value function in terms of the new state variables and where we denote partial derivatives by subscripts. We now take advantage of the logarithmic form of utility to guess the following form of the value:

$$\tilde{V}(A,P) = \tfrac{1}{\rho} + \tfrac{1}{\rho}\ln(A) + V(P), \qquad [3.6]$$

where $V : [0,1] \to \mathbb{R}$ is a function that is to be determined. [15] We now use this guess and write the continuation value to a first order in [3.5]. Invoking A-linear strategies, the terms in A drop out (since we guessed the correct functional form) and we find the following simplified HJB that depends exclusively on P:[16]

$$\begin{aligned}\rho V(P) = \max_{C \in \mathbb{C}(P)} \Bigg\{ & \underbrace{\ln(C)}_{\text{flow utility}} \underbrace{- [C + C'(P) + T(P) + T'(P)]\frac{1}{\rho}}_{\text{run-down of common resource } A} \\ & + \Big(\underbrace{-(1-P)[T(P) + C] + P[T'(P) + C'(P)]}_{\equiv \dot{P}_t \text{ (drift of } P\text{): distributional concerns}} \Big) V_P(P) \Bigg\} \\ & + \underbrace{\sigma^2 P(1-P)}_{\text{volatility of } P\text{: distributional shocks}} \quad V_{PP}(P) \quad \underbrace{-\frac{\sigma^2}{2\rho}}_{\text{shocks to common resource } A}\end{aligned} \qquad [3.7]$$

where we define $\mathbb{C}(P) = [0, -T(0)]$ if $P = 0$ and $\mathbb{C}(P) = [0, \infty)$ otherwise. It is worthwhile to pause here for a moment to grasp the meaning of the different terms in this equation.

We start by discussing Region 1's consumption choice, C, which plays out inside the max-operator. The effects of C are threefold. First, there is the straightforward effect that consuming more gives higher flow utility, $\ln(C)$. Second, an increase in C reduces the aggregate resource A, which is valued marginally at $1/\rho$. We see that Region 2's consumption and the dead-weight loss from transfers, $T + T'$, are accounted for as well in this term. Third, higher consumption decreases Region 1's wealth share, as captured by i) the drift of P and ii) the marginal valuation of this wealth share, $V_P(P)$. Taking the derivative with respect to C inside the max-operator, we obtain the following first-order condition (FOC):

$$\frac{1}{C(P,T)} \geq \frac{1}{\rho} + (1-P)V_P(P), \qquad \text{with equality if } P > 0. \qquad [3.8]$$

[14] We neglect for now the constrained case, $P = 0$, but include it again in the HJB [3.7]. The adjustments that have to be made in this case should be obvious.

[15] This guess can be arrived at naturally using A-linear strategies for all players to form the integral $\tilde{V}(A_0, P_0) = \mathbb{E}_0 \int_0^\infty e^{-\rho t}\ln(C(P_t)A_t)\mathrm{d}t$ and simplifying.

[16] The corresponding equation for Region 2 can be obtained making the obvious adjustments .

It says that the marginal utility from increasing consumption should equal the marginal cost of doing so, which, in turn, is given by marginal value of the common resource, $1/\rho$, plus the marginal cost of a decrease in the wealth share. Note that if $P = 0$, the FOC may hold as an inequality, in which case Region 1 is constrained and consumes the subsidy $-T(0)$ received from the center. Denoting by $C_{unc} = \left[\frac{1}{\rho} + (1-P)V_P(P)\right]^{-1}$ the unconstrained-optimal consumption rule that obtains from solving the FOC [3.8] with equality, Region 1's best response in the consumption stage is thus given by the following simple rule:

$$C(P,T) = \begin{cases} C_{unc}(P) & \text{if } P > 0, \\ \min\{C_{unc}(0), -T(0)\} & \text{if } P = 0, \end{cases} \qquad [3.9]$$

This result shows that Region 1's consumption is independent of the center's transfer whenever $P > 0$, which simplifies computation of equilibrium enormously. The economic reason is that policies of the other players do not change Region 1's marginal value of savings to a first order.[17]

We now turn to discussing the terms in the last row of the HJB [3.7], which capture the two orthogonal risk components. The first term isolates shocks to the wealth distribution: The more risk-averse the region is regarding the wealth distribution (i.e. the higher V_{PP}), the lower the region's flow value, ρV, will be. The last term isolates shocks to total wealth; since regions are risk-averse this effect is always negative.

Center's HJB. We now turn to the recursive representation for the center's problem. Denoting the center's value function by W and following the same steps as for the regions' problem, the center's HJB is found to be (suppressing again function arguments)

$$\begin{aligned} \rho W = \max_{T,T'} \Big\{ & \mu \ln(C) + (1-\mu)\ln\Big(C'\Big) - \Big[C + C' + T + T'\Big]\frac{1}{\rho} \\ & + \Big(-(1-P)[T + C] + P[T' + C']\Big) W_P \Big\} + \sigma^2 P(1-P)W_{PP} - \frac{\sigma^2}{2\rho}, \\ \text{s.t.} \quad & X(T/\rho) + X(T'/\rho) = 0, \qquad T \le 0 \text{ if } P = 0, \qquad T' \le 0 \text{ if } P = 1, \end{aligned} \qquad [3.10]$$

where regions' consumption rates C and C' are determined by best responses in the consumption stage, Eq. [3.9].

In the center's HJB [3.10], we recognize the corresponding terms from Region 1's HJB [3.7], however now both regions' flow utilities enter into the center's value. The center takes a region's contemporaneous consumption decisions as given while the region has positive wealth. Once a region is broke, however, the center's marginal calculus changes since it can restrict the region's consumption: for example, the center effectively sets Region 1's consumption for small enough subsidies, $T < C_{unc}(0)$. Furthermore, given our assumption on shocks, the last term of the HJB (which stems from volatility of A_t) is invariant in P and thus there is no portfolio effect entering the center's redistribution decision. We only state the HJB here and analyze the first-order conditions separately for the different cases (constrained versus unconstrained regions) in the following sections.

[17] This is, in fact, a pervasive feature of savings games with Brownian noise, see also Barczyk and Kredler (2014b).

3.3 Two benchmarks: Efficient allocations and wealth pooling

Before turning to equilibria of the full game, we study two interesting benchmarks that can be solved in closed form and will aid our understanding. Proofs for this section and further results are provided in the Appendix, Sections F and G.

Efficiency. In the case in which redistribution is costless, the efficient allocations can be tightly characterized using a social planner's problem. In a nutshell, given the volatility specification [2.4] the planner is indifferent when it comes to deciding in which region's asset to invest, thus transfers T, T' and the evolution of P_t are indeterminate. However, efficiency imposes tight restrictions on consumption rules and on the evolution of aggregate wealth, which we summarize in the following proposition:

Proposition 1 (Efficient allocations under costless redistribution)**.** *Assume that there are no distortions from taxation, i.e. $X(\tilde{\tau}) = \tilde{\tau}$. Then there is a continuum of Pareto-efficient allocations, parameterized by $\hat{\mu} \in [0,1]$. For given $\hat{\mu}$, the efficient consumption paths satisfy $C^{eff}(\hat{\mu}) = \hat{\mu}\rho$ and $C'^{eff}(\hat{\mu}) = (1-\hat{\mu})\rho$. In all efficient allocations, total wealth is driftless, i.e. $\dot{A}_t^{eff} = 0 \; \forall t$.*

When redistribution is costly (under Ass. 1), however, the planner's problem is more complicated, since then the efficient allocation can no longer be independent of P_t. Intuitively, the social planner would like to avoid transfers since they involve dead-weight loss, but is forced to use transfers once a region has run out of wealth.

In general, however, the efficient consumption allocation cannot be sustained in this environment; this is true even if redistribution is costless and even if the center can commit to future transfer policies. This is because there is tension between two forces: Efficiency requires that consumption shares C and C' be constant over time. In other words, regions should be insured against shocks to their wealth shares ($\mathrm{d}B_t^P$) and their consumption should only vary in response to aggregate shocks ($\mathrm{d}B_t^A$). However, this obviously kills savings incentives for regions and makes them over-consume.

Proposition 2 (Generic inefficiency)**.** *Suppose that $\sigma > 0$ and restrict attention to bounded transfer flows by the center. Then the consumption allocation $C_t = C^{eff}(\hat{\mu})$ and $C'_t = C'^{eff}(\hat{\mu})$ for all t cannot be supported as an equilibrium whenever $\hat{\mu} \in (0,1)$.*

There are two exceptions to this non-existence result that are illustrative, yet somewhat obvious: i) If $\mu = 0$ or $\mu = 1$, then the center can implement its preferred allocation by immediately expropriating the less-liked region. ii) If there are no shocks –and thus no gains from risk-sharing–, then a center *with commitment* can implement any efficient allocation by an initial mass transfer and committing to zero transfers thereafter. This exception, however, crucially hinges on commitment.

Wealth pooling. Second, consider the following modified *wealth-pooling* (WP) environment. There is no center, i.e. we set $T_t = T'_t = 0$ for all t. Both regions consume out of the common resource, A_t^{wp}, which evolves according to

$$\mathrm{d}A_t^{wp} = (\rho - C_t - C'_t)A_t^{wp}\mathrm{d}t + \sigma^2(A_t^{wp})^2\mathrm{d}W_t,$$

where W_t is a standard Wiener Process. We note here that this modified environment is a standard game of strategic resource extraction.

Proposition 3 (Wealth pooling)**.** *In the wealth-pooling environment, regions' equilibrium strategies are $C^{wp} = C'^{wp} = \rho$ and the common resource is run down at the (inefficiently high) rate $\dot{A}_t^{wp}/A_t^{wp} = -\rho$.*

Inefficiency arises due to a classical tragedy of the commons: regions fail to take into account the negative externality they exert on each other by consuming out of the common resource. The WP allocation is interesting because it gives us an upper bound on what a region may reasonably want to consume out of the common resource. This can be seen by from Region 1's FOC [3.8]. Region 1 chooses C^{wp} either (i) if indifferent with regard to its wealth share, i.e. $V_P = 0$, or (ii) if it owns all wealth in the economy, i.e. $P = 1$. In turn, Region 1 chooses a lower rate, $C < C^{wp}$, whenever it prefers to increase its wealth share, i.e. if $V_P > 0$ and $P < 1$.

4 Costly redistribution: bail-out equilibria

This section analyzes (A-linear) MPEs assuming that redistribution is costly, which is our baseline Assumption 1. We start by solving for the center's best response, first on the interior and then at the boundaries of the state space. We then describe the equilibrium using a numerical example. This equilibrium features *bail-outs*, which we define as an upward jump discontinuity in subsidies to the poor region when it hits the constraint. The sharp increase in subsidies coincides with a sharp drop in consumption of the poor region. Finally, we introduce a *limit game* that gives a marginal-cost-benefit interpretation of agents' value-matching condition at the constraint and state our main result on bail-outs.

4.1 Center's problem

We will only discuss the aspects of the center's problem here that are economically most relevant; we refer the reader to Appendix A for the details.

Unconstrained problem. First, consider the center's problem for a fixed $P \in (0, 1)$ on the interior of the state space. Leaving out all terms in the Hamiltonian in [3.10] that the center cannot affect, the center's problem reduces to

$$\max_{T,T'} \Big\{ -(\underbrace{T + T'}_{\text{dead-weight loss}})^{\frac{1}{\rho}} + \underbrace{[PT' - (1-P)T]\, W_P(P)}_{\text{gains from re-distribution}} \Big\} \qquad [4.1]$$
$$\text{s.t. } X(T/\rho) + X(T'/\rho) \geq 0; \qquad P, W_P(P) \text{ given,}$$

To understand the trade-off that the center faces, consider a situation in which $W_P(P) > 0$, i.e. the center wants to re-distribute towards Region 1. The second term in the criterion encodes the center's gains from increasing P, which it weighs against the costs of redistribution (the first term). This cost is the dead-weight loss from taxation, $T + T'$, valued at the marginal value of common resources, $1/\rho$. Since the marginal dead-weight loss increases as the center redistributes more, the optimal T' is increasing in W_P; see Appendix A for the detailed solution.

Constrained problem. We now shift our attention to the boundary $P = 0$, where Region 1 may be borrowing-constrained.[18] In this case, the center can effectively control the broke region's consumption, which alters the center's trade-off completely.

We will first solve an auxiliary (*constrained*) problem at $P = 0$, pretending that Region 1 always consumes the entire transfer provided by the center and saves nothing; this problem is convenient to solve and is guaranteed to have an interior solution. Then, we combine the auxiliary problem with the unconstrained problem [4.1] to solve the center's actual problem at $P = 0$. For ease of interpretation, we write the center's problem in the variable $S \equiv -T(0)$, which is the subsidy to the broke region and thus positive. From the center's HJB [3.10] (omitting again terms the center cannot affect), the center's constrained problem is

$$\max_{S\geq 0}\left\{\mu\ln(S) + (1-\mu)\ln(C^{wp}) - \left[C^{wp} + \tilde{T}(S)\right]\frac{1}{\rho}\right\} \qquad [4.2]$$

where we define the tax revenue needed to cover the subsidy from the center's constraint in problem [4.1] by the function $\tilde{T}(S) = \rho X^{-1}(S/\rho)$.[19] Since Region 1 does not save by assumption, the aggregate resource decreases by $C^{wp} + T'$. Taking the derivative with respect to S in [4.2], one obtains the FOC that pins down the optimal constrained subsidy S_{con}:[20]

$$\underbrace{\frac{1}{\rho}}_{\text{MC of taxing}} = \underbrace{\frac{\mu}{S_{con}}}_{\text{center's MU from R.1 consuming } S} \underbrace{X'\left(\tilde{T}(S_{con})/\rho\right)}_{\text{marginal tax revenue}}. \qquad [4.3]$$

The left-hand side captures the marginal cost of taxing Region 2, which is given by the marginal cost of resources. This marginal cost equals the marginal benefit, which consists in the marginal utility that the center obtains from Region 1 consuming the marginal tax revenue.

Center's problem at $P = 0$. The unconstrained and constrained problems above can be combined to characterize the solution to the center's true problem at $P = 0$. In Appendix A, we write this combined problem in the single choice variable S. Recall now that Region 1's best response in the consumption stage is to consume the entire subsidy as long as $S \leq C_{unc}(0)$, but to save any additional subsidy beyond this level. Thus, the center's payoff equals that of the constrained problem for $S \leq C_{unc}(0)$, while it equals that of the unconstrained problem for $S > C_{unc}(0)$. The payoff function is concave on the two parts and continuous at $S = C_{unc}(0)$, which allows a simple characterization of

[18]The analysis for the point $P = 1$ is analogous, making the obvious adjustments.

[19]For the sake of exposition, we have maintained the terms in Region 2's consumption in the center's problem although the center cannot affect the optimal choice $C'(0) = C^{wp}$.

[20]Note that the center's criterion in [4.2] is strictly concave in S, since concavity of $X(\cdot)$ implies that convexity of $X^{-1}(\cdot)$ and thus convexity of $\tilde{T}(\cdot)$. Thus the FOC is necessary and sufficient.

the center's optimal strategy. In our numerical solutions, we find that the center's optimal subsidy always occurs on the constrained part.

4.2 Equilibrium dynamics

We solve numerically for the MPE of the economy using a standard Markov-chain approximation on a grid for $P \in [0,1]$ using a centered differencing scheme. We iterate the game backwards in time until policy functions converge, finding a unique equilibrium from a variety of final guesses for policies. The equilibrium is found in less than a second on a standard computer. For details on the computation and additional results, see Appendix E.

To parameterize our baseline economy, we choose a standard discount rate $\rho = 0.04$ (one time unit corresponding to a year). We set $\mu = 0.5$, meaning that the center does not favor any of the two regions.The shock volatilityis set to $\sigma = 0.1$. We adopt the specification [2.6] for $X(\cdot)$, setting the curvature parameter to $\kappa = 1.5$. Given this value, we then choose $\alpha = 8.3$ such that the subsidy to a constrained broke region (which is the maximal subsidy in equilibrium) is 10% of aggregate output. To get a sense on how large the costs to redistribution are, observe that our parameterization of $X(\cdot)$ that the maximal dead-weight loss (occurring again at the constraint) is 13.2% of aggregate GDP.

Figure 2 shows the equilibrium outcomes. As expected, regions' value and policy functions are strictly increasing in their wealth share. The center, however, always prefers to balance the wealth distribution; it enjoys maximal utility at wealth parity, where regions' utilities are equalized and the dead-weight loss from redistribution is zero. Zooming in at the boundaries, we can see the bail-out at play: The broke region's consumption jumps down and the center's subsidy jumps up at the constraint; we will analyze this point in detail in the next section.

Remarkably, the model generates non-linear dynamics, as can be seen from the drift, $\dot{P}$, in the lower right panel. There is a large stable region with (stochastic) steady state $P = 1/2$ in the interior of the state space. In this stable region, the center chooses a redistributive tax policy that is successful in counteracting the impoverishing dynamics that are induced by regions' free-riding behavior (see the decomposition of the drift into consumption and transfer components in the lower right panel).[21]However, once wealth inequality is high, excessive consumption by the impoverished region generate a drift towards bankruptcy, generating two impoverishment intervals close to the boundaries. In terms of efficiency, notice that depletion of total resources (dA/A) is lowest, and closest to the efficient benchmark $\dot{A}^{eff} = 0$, at $P = 1/2$, but increases towards the boundaries (bottom-left panel).[22] In what follows, we will characterize the bail-out at the boundary in detail.

[21]To give a sense of the stability around $P = 1/2$, we calculate that it takes around 150 years in expectation for one region to end up in bankruptcy from this point, see Appendix J for details.

[22]This discontinuity in asset depletion in the frontier is smaller that the jump in consumption of the poor region, as the discontinuous increase in transfers exacerbates the deadweight loss.

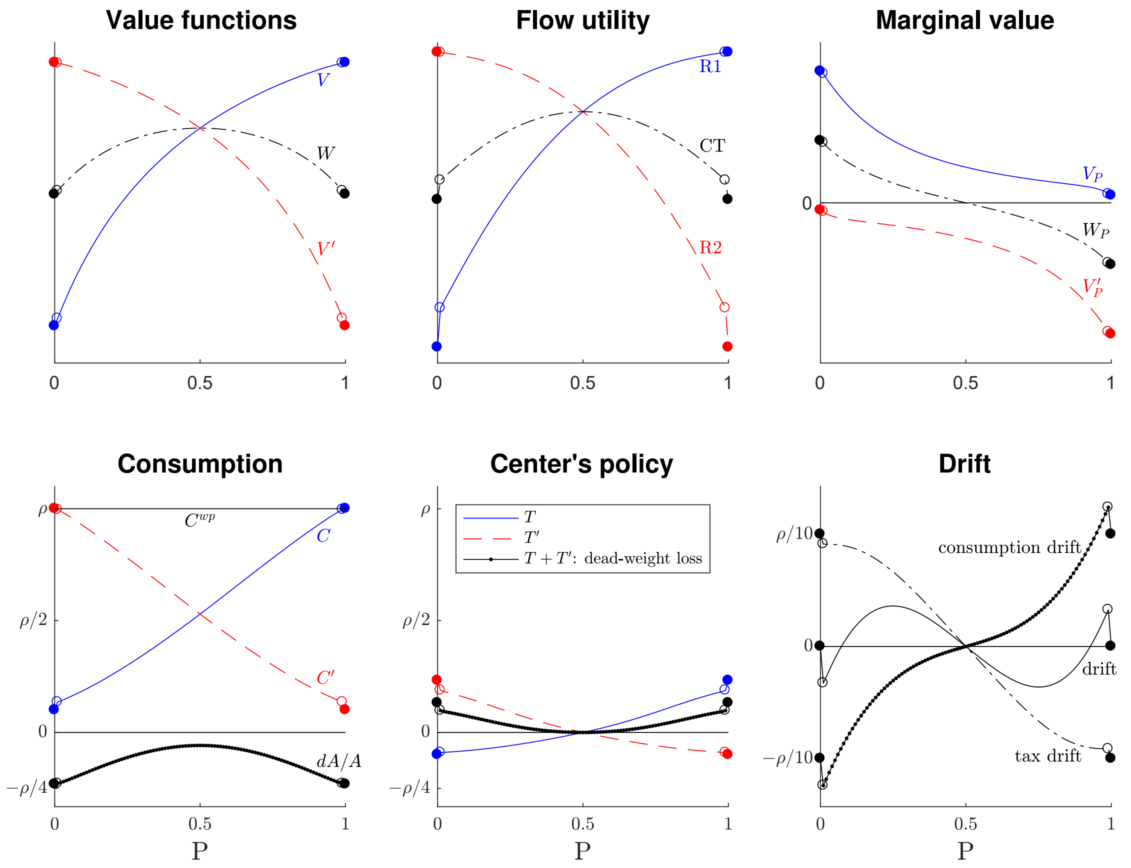

Figure 2: Equilibrium with bail-outs under costly redistribution.

Model parameters: $\rho = 0.04$, $\sigma = 0.1$, $\mu = 0.5$, $\kappa = 1.5$, $\alpha = 8.3$. Grid size: $N = 101$. The two circles at the boundaries depict a variable's value on the second-to-last grid point (empty circle) and the last grid point (solid circle), indicating continuity of the variable at the boundary. *Flow utility* includes losses from run-down of aggregate wealth, e.g. for Region 1 we plot $\ln(C) - (C + C' + T + T')/\rho$. Panel *Drift* decomposes the *drift* $\dot{P}$, defined in Eq. [3.7], into terms concerning consumption (*consumption drift*: $-(1-P)C + PC'$ and terms in transfers (*tax drift*: $-(1-P)T + PT'$). The equilibrium with $\mu = 0.4$ is shown in Figure 5.

4.3 Bail-outs

The previous analysis shows that both the poor region's and the center's problem change in nature at the boundary $P = 0$: the poor region suddenly faces a constraint on consumption, and the center can suddenly dictate the poor region's consumption, at least on some range. This explains why optimal policies can display jump discontinuities at $P = 0$, even if value functions must be continuous at this point. Mathematically, equilibrium requires that all agents' HJBs be fulfilled both at $P = 0$ itself and in the limit as $P \downarrow 0$; continuity of the value functions then implies value-matching conditions for all players in the limit as $P \downarrow 0$.[23] In this section, we show that these value-matching conditions

[23] There are no smooth-pasting conditions on the value-function derivative since the constrained interval, i.e. the set $\{P : P = 0\}$ is degenerate, thus no value function derivative exists within this regime.

can be understood economically as optimality conditions in a *limit game* played between Region 1 and the center when Region 1's wealth is almost run down. In this limit game, agents optimally time the point at which the boundary is hit, taking as given the policies of the other players; this concept is novel to the best of our knowledge.

The section proceeds in three steps. We first define the limit game. Second, we show that agents' optimality conditions in the limit game imply the value-matching conditions, thus providing us with a marginal-cost-benefit interpretation for the latter. Third, we characterize the best responses of the poor region and the center and summarize our main result on bail-outs in a proposition.

4.3.1 Defining the limit game

To define the limit game, we first derive how the time interval until the constraint is hit (*time-to-broke*) is affected by agents' decisions. We then analyze the center's and Region 1's best responses and show that they imply the value-matching conditions when we convert the marginal calculus from quantities (C, S) to a timing decision.

Time-to-broke. Consider the game between Region 1 and the center when only a small quantity of wealth ΔP is left for Region 1.[24] The choice variables in the game are thus consumption C for Region 1 and the subsidy $S = -T$ the center pays to Region 1. We limit our analysis to policies that lead the economy to $P = 0$.[25] The drift converges to $\dot{P}^{lim}(C, S) \equiv S - C$ as $P \to 0$ by [3.4]. Let us denote by $b(C, S)\Delta P$ the expected time that is left until Region 1 is broke given a pair of policies (C, S), for fixed $P = \Delta P$ (small). Since the drift dominates as the volatility of the process approaches zero at $P = 0$, this *time-to-broke* can be approximated by[26]

$$b(C, S)\Delta P = \frac{\Delta P}{-\dot{P}^{lim}(C, S)} = \frac{\Delta P}{C - S} \qquad \text{for any } C > S. \qquad [4.4]$$

From this equation, we obtain the *time-to-broke* function $b(C, S) = 1/(C - S)$ as the inverse of the absolute drift. As is intuitive, higher consumption by the poor region hastens bankruptcy, whereas higher subsidies delay it.

[24]Note that Region 2 will optimally choose $C' = C^{wp}$ in the limit by its FOC [3.8], thus we can fix its policy.

[25]The boundary behavior of the stochastic process P_t at $P = 0$ is anything but trivial since the volatility of the process approaches zero at the boundary in a square-root shape. In particular, it is not guaranteed that the system reach the boundary in finite time. In Appendix I, we show that two cases of boundary behavior are possible. First, for low drifts ($\lim_{P\to 0} -T(P) - C(P) < \sigma^2/2$), the boundary is hit almost surely in the limit and the expected exit time converges to zero as $P \to 0$. Second, for high-enough drifts ($\lim_{P\to 0} -T(P) - C(P) \geq \sigma^2/2$) the probability of hitting the boundary is zero. Importantly, our assumptions thus allow for the possibility that the poor region stays solvent forever; however, this does not occur in the equilibrium we find.

[26]Section I in the Appendix derives the precise expressions for the expected exit times, which also contain terms in the volatility. However, we present a heuristic argument here, which is simpler and also leads to the correct value-matching conditions.

Towards a definition of the limit game. Now, consider Region 1's problem, fixing a subsidy S by the government. We limit attention to consumption rates $C \geq S + \epsilon$ to ensure that the economy converges to $P = 0$, where $\epsilon > 0$ is a (small) constant. Taking as given the value $V(0)$ that Region 1 obtains when broke, its limit problem is

$$V(\Delta P) = \max_{C \geq S+\epsilon} \left\{ U^{lim}(C,S) b(C,S) \Delta P + e^{-\rho b(C,S)\Delta P} V(0) \right\}, \text{ where}$$
$$U^{lim}(C,S) = \ln(C) - (C + C^{wp} + \tilde{T}(S) - S)\frac{1}{\rho}. \qquad [4.5]$$

This is reminiscent of the Bellman principle [3.5], $b(C,S)\Delta P$ playing the role of $\mathrm{d}t$ and instead ΔP being fixed. The second term in the definition of U^{lim} has the same interpretation as in the HJB [3.7], capturing the run-down of common resource A.

We now write Region 1's payoff, i.e. the term in brackets in [4.5], as a function of C and S. Approximating $e^{-\rho b(C,S)\Delta P} \simeq 1 - \rho b(C,S)\Delta P$ and as $\Delta P \to 0$, Region 1's limit objective function is

$$v^{lim}(C,S) \equiv \lim_{\Delta P \to 0} \frac{V(\Delta P) - V(0)}{\Delta P} = \frac{U^{lim}(C,S) - \rho V(0)}{C - S}, \qquad \text{for } C \geq S + \epsilon. \quad [4.6]$$

This says that Region 1's payoff in the limit game is proportional to i) time to broke and ii) the differential of flow utility before broke over the flow value once broke. Note here that v^{lim} is identical to the derivative of the value function at $P = 0$, which is intuitive: Region 1 sets C to maximize the value just before being broke, taking as given $V(0)$, thus maximizing the slope of the value function.

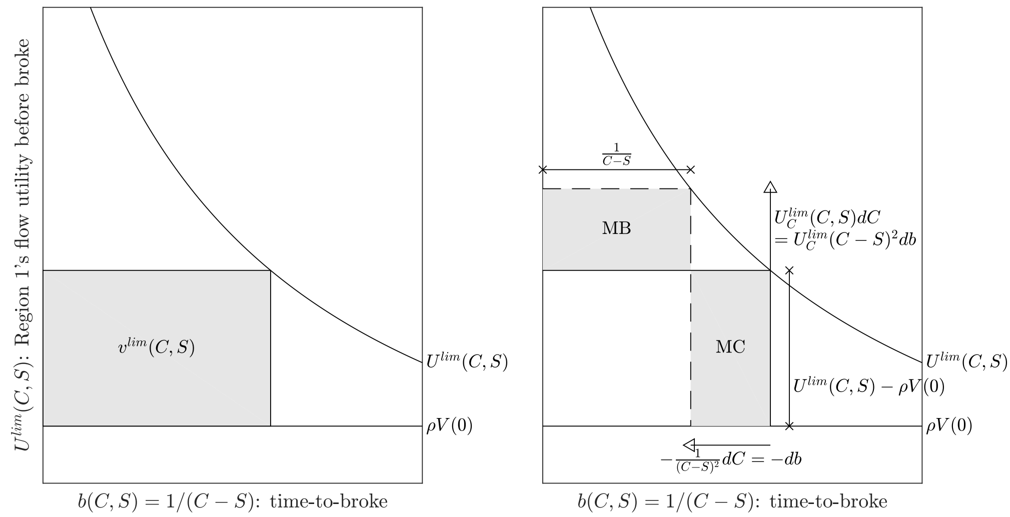

Figure 3: Illustration of Region 1's value-matching condition in terms of marginal cost and marginal benefit

Region 1's limit payoff v^{lim} is graphically represented in the left panel of Figure 3. For a fixed S, the hyperbola in the graph traces out all combinations of flow utility,

$U^{lim}(C, S)$, and time-to-broke, $b = 1/(C - S)$, that obtain when Region 1 varies C. Higher C increases the flow value U^{lim} that Region 1 receives above the bankruptcy value $\rho V(0)$ –the height of the shaded rectangle–, but shortens time-to-broke b (the width of the rectangle). Region 1 optimally chooses C to maximize the rectangle under the curve $U^{lim}(C, \cdot)$, i.e. the flow-value differential to bankruptcy multiplied by how long this differential is enjoyed.

We now turn to the center's problem. Following the same steps as for Region 1, the center's limit objective function is

$$w^{lim}(C, S) \equiv \lim_{\Delta P \to 0} \frac{W(\Delta P) - W(0)}{\Delta P} = \frac{U^{c,lim}(C, S) - \rho W(0)}{C - S} \qquad \text{for } S \leq C - \varepsilon, \tag{4.7}$$

where again we bound away S from C to have a well-defined problem. Here, the center's flow utility is defined as

$$U^{c,lim}(C, S) = \mu \ln(C) + (1 - \mu)\ln(C)^{wp} - \left[C + C^{wp} - S + \tilde{T}(S)\right]\tfrac{1}{\rho}. \tag{4.8}$$

We then define the limit game and its equilibrium as:

Definition 3 (Limit game)**.** *The* limit game *at $P = 0$ is given by the pair of problems* $\max_C v^{lim}(C, S)$ *for Region 1 and* $\max_S w^{lim}(C, S)$ *for the center. An* equilibrium *of the limit game is a pair of* best responses (C_{lim}, S_{lim}) *satisfying* $C_{lim} = \arg\max_C v^{lim}(C, S_{lim})$ *and* $S_{lim} = \arg\max_S w^{lim}(C_{lim}, S)$.

4.3.2 Value-matching conditions

We now proceed to show that an equilibrium of the limit game implies the value-matching conditions at $P = 0$. Lemma 1 in the Appendix B formalizes this equivalence.

Region 1. We first explain Region 1's marginal-cost-benefit analysis, depicted in the right panel of Figure 3. Given a fixed S, Region 1's first-order condition (FOC) for C is obtained by taking the derivative with respect to C in Eq. [4.6] as

$$\underbrace{\left[\frac{U_C^{lim}(C, S)}{C - S}\right]}_{MB} = \underbrace{\left[\frac{U^{lim}(C, S) - \rho V(0)}{(C - S)^2}\right]}_{MC}. \tag{4.9}$$

The left-hand side of this equation gives the marginal benefit (MB) of increasing consumption by dC, which is represented by the area MB in the figure. MB equals the marginal increase in flow utility, U_C^{lim}, multiplied by the time over which this benefit accrues. The right-hand side of [4.9] represents the marginal cost (MC) of increasing consumption by dC, represented by the area MC. This marginal cost consists of giving up the utility differential $U^{lim} - \rho V(0)$ over a time interval $|db| = |(C - S)^2 dC|$, whose length is determined by how much time-to-broke decreases due to dC. In the optimum, it must be that MC=MB, i.e. that the areas have the same size.

Now, multiply Eq. [4.9] by $(C-S)^2$ and invoke Region 1's FOC for C to obtain Region 1's value-matching condition at $P=0$:[27]

$$\underbrace{(C-S)U_C^{lim}(C,S) - \left[U^{lim}(C,S) - \rho V(0)\right]}_{\equiv NMB(C,S)} = 0, \tag{4.10}$$

where the left-hand side defines Region 1's *net marginal benefit* (NMB). We note here that when multiplying by $(C-S)^2$ we have changed the units in the marginal calculus: Since $db/dC = -(C-S)^2$, the value-matching condition [4.10] corresponds to Region 1 making infinitesimal changes to time-to-broke, b. This may also be seen from the right panel in Fig. 3. If Region 1 shortens time-to-broke by a marginal unit, $-db$, it incurs a marginal cost of $[U^{lim} - \rho V(0)]db$, corresponding to the term in brackets in the value-matching condition [4.10]. The marginal decrease of time-to-broke, $-db$, translates into an increase in consumption of $dC = (C-S)^2 db$ that is enjoyed over a time interval of length $b = 1/(C-S)$, corresponding to the product $(C-S)U_C^{lim}$ in [4.10] .

Center. We now proceed analogously for the center. Taking the derivative of w^{lim} with respect to S in [4.7] and setting equal to zero, we find

$$\underbrace{\left[\frac{U^{c,lim} - \rho W(0)}{(C-S)^2}\right]}_{MB^c} = \underbrace{\left[\frac{-U_S^{c,lim}(C,S)}{C-S}\right]}_{MC^c}. \tag{4.11}$$

The marginal cost of an increase dS in the subsidy (right-hand side) is given by the marginal increase in the deadweight loss, $-U_S^{c,lim} = (\tilde{T}'(S)-1)/\rho \geq 0$, which occurs over time-to-broke $1/(C-S)$. The marginal benefit of increasing S is that it delays the bail-out (i.e. it decreases time-to-broke) by $dS/(C-S)^2$, enabling the center to enjoy the utility differential $U^{c,lim} - \rho W(0)$ over the bail-out for longer (note that this differential must be positive, otherwise there is no trade-off). To obtain the center's value-matching condition at $P=0$, again multiply Eq. [4.11] by $(C-S)^2$, invoke the center's FOC and re-arrange:

$$\underbrace{U^{c,lim}(C,S) - \rho W(0) + (C-S)U_s^{c,lim}(C,S)}_{\equiv NMB^c(C,S)} = 0, \tag{4.12}$$

which now has the interpretation that the center's net marginal benefit (NMB^c) of *increasing* time-to-broke by a marginal unit db must be zero in the optimum.[28]

4.3.3 Best responses and equilibrium

Armed with the value-matching conditions, we are now prepared to characterize the best responses of the impoverished region and the center in the limit game. We find two

[27] To see that this is the value-matching condition at $P=0$, take the limit $P \to 0$ in Region 1's HJB, Eq. [3.7], and note that $U_C^{lim}(C,S) = 1/C - 1/\rho = \lim_{P\to 0} V_P(P)$ by Region 1's first-order condition [3.8].

[28] To see that [4.12] is the value-matching condition, take limits of the center's HJB [3.10] as $P \to 0$ and observe that the center's FOC [A.1] for the subsidy, when written in terms of S and taking the limit $P \to 0$, becomes $W_P = \left[\tilde{T}'(S) - 1\right]\frac{1}{\rho}$.

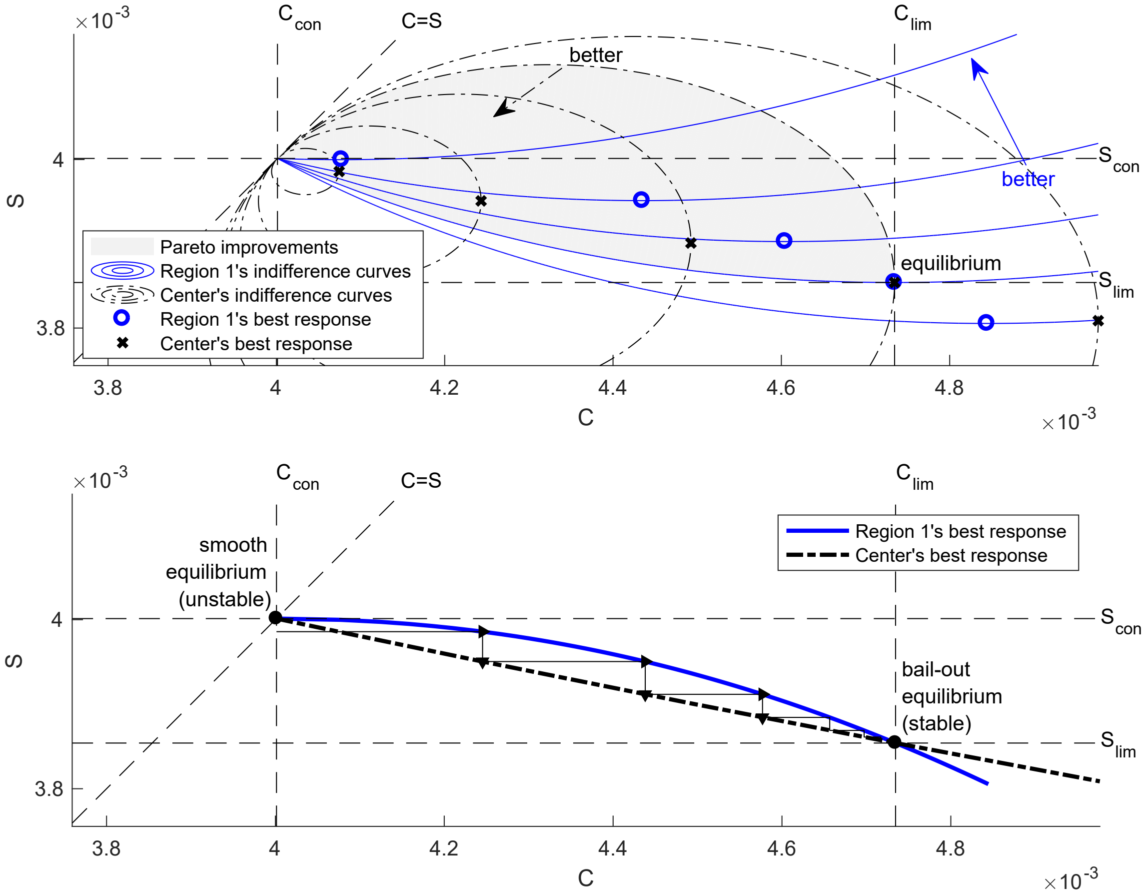

Figure 4: Limit game between Center and Region 1 at $P = 0$.

Model parameters as in Fig. 2. C- and S-axis are drawn on same scale. The upper panel shows level lines of $v^{lim}(C, S)$, i.e. Region 1's indifference curves, and level lines of $w^{lim}(C, S)$, i.e. the center's indifference curves in the limit game, where $V(0)$ and $W(0)$ are values of playing the constrained allocation (C_{con}, S_{con}) forever at $P = 0$. The shaded area indicates Pareto improvements over the equilibrium (C_{lim}, S_{lim}). Lower panel plots best responses, which coincide with allocations that satisfy agents' value-matching conditions.

equilibrium candidates: smooth paths and the (non-smooth) bail-out equilibrium. We show that only the bail-out equilibrium is stable. This section contains all arguments of the proof of our main result, Prop. 4, giving the economic intuition for each step; we provide a concise mathematical proof in Appendix D.

Region 1's best response. Figure 4 depicts preferences (upper panel) and best responses (lower panel) in the limit game, fixing the policies at $P = 0$ to the constrained levels ($C(0) = C_{con}$ and $S(0) = S_{con}$). In the upper panel, the spiral-shaped solid lines are Region 1's indifference curves. Region 1 prefers policies towards the north, where subsidies are high. For fixed S, Region 1's best response occurs at the tangency of the iso-S lines with its indifference curves (circles in the upper panel, solid line in the lower panel). We see that whenever the center sets a subsidy $S < S_{con}$ below the bail-out level, Region 1 responds with a downward-jumping consumption trajectory, i.e. $C > C_{con}$. We will now explain why this is the case.

We first establish that Region 1's net marginal benefit (NMB) in Eq. [4.10] is decreasing in C, as one would expect. Since marginal utility of consumption is decreasing, we have

$$NMB_C(C,S) = (C-S)U^{lim}_{CC}(C,S) = -\frac{C-S}{C^2} < 0 \qquad \text{for all } C > S. \tag{4.13}$$

We now show that whenever the center sets a subsidy below the bailout level, i.e. $S < S_{con}$, this makes Region 1's NMB at $C = C_{con}$ positive. Again from [4.10], we obtain that the marginal effect of a decrease in S on NMB is

$$-NMB_S(C,S) = U^{lim}_C(C,S) + U^{lim}_S(C,S) = \frac{1}{C} - \frac{\tilde{T}'(S)}{\rho}. \tag{4.14}$$

Lowering S has two opposing effects on Region 1's NMB. First, a lower S hastens the bail-out and makes it more costly for Region 1 (in terms of C) to delay the bail-out, thus making marginal-utility considerations more important in Region 1's marginal calculus. This effect is captured by the term U^{lim}_C in [4.14], which is positive. A second effect (U^{lim}_S) goes in the opposite direction, but turns out be weaker: A lower subsidy S implies a lower deadweight loss from taxation, which increases Region 1's pre-bailout utility and thus makes Region 1 want to delay the bail-out (and thus decrease C). We will now show that the first effect is dominant, unless Region 1's and the center's preferences are perfectly aligned.

Note that we can read the last equality in Eq. [4.14] as follows: Lowering S increases Region 1's NMB if and only if $1/C > \tilde{T}'(S)/\rho$, i.e. if and only if Region 1 benefits from an increase in C that is financed by taxing Region 2. But now, the center's FOC in the bail-out, see Eq. [4.2], is

$$\frac{\mu}{C_{con}} - \frac{\tilde{T}'(S_{con})}{\rho} = 0.$$

So if $\mu < 1$ (i.e. if the center places at least some weight on Region 2), this directly implies that $-NMB_S > 0$ in Eq. [4.14] for $C = C_{con}$ and all $S \leq S_{con}$. Thus, $MNB(C_{con}, S) > 0$ for all $S \leq S_{con}$, which implies that Region 1's best response to $S < S_{con}$ is some $C > C_{con}$ (since $NMB_C < 0$, as was established before). As is intuitive, Region 1 has an incentive to increase its consumption before the bail-out since in its eyes the bail-out subsidy is too low. Interestingly, there must be a conflict of interest between the center and Region 1 (non-identical preferences, $\mu < 1$),) for the reasoning to go through.

Center's best response. Now, turn to the center. In Fig. 4, the center's indifference curves in the limit game (the thin dash-dotted lines) trace out circles that move ever closer to the bail-out configuration (C_{con}, S_{con}) as utility levels increase. This is not surprising, since the center is in charge in the bail-out and sets both S and C at its preferred level. The center's best response occurs where the highest indifference curve is reached at the point of tangency with the iso-C lines. We see that whenever Region 1 over-consumes before the bail-out and sets $C > C_{con}$, the center reacts with an upward jump of the subsidy upon the bail-out, i.e. $S < S_{con}$. We now show and explain why this is the case.

From [4.12], we see that the center's net marginal benefit decreases in S, as we expect:

$$NMB^c_s(C,S) = (C-S)U^{c,lim}_{ss}(C,S) = -(C-S)\frac{\tilde{T}''(S)}{\rho} < 0 \qquad \text{for all } C > S, \quad [4.15]$$

since the marginal tax required to additional subsides is increasing, i.e. $\tilde{T}''(S) > 0$. An increase in Region 1's consumption C affects the center's NMB as follows:

$$NMB^c_c(C,S) = U^{c,lim}_c(C,S) + U^{c,lim}_s(C,S) = \frac{\mu}{C} - \frac{\tilde{T}'(S)}{\rho}. \quad [4.16]$$

There are again two effects of C on the center's NMB, the first being positive (at least for C close to C_{con}) and the second being negative. As for the first effect, note that if $\mu > 0$ we have

$$U^{c,lim}_c(C_{con}, S_{con}) = \frac{\mu}{C_{con}} - \frac{1}{\rho} > \frac{\mu}{C_{con}} - \frac{\tilde{T}'(S_{con})}{\rho} = 0. \quad [4.17]$$

Here, the inequality follows from $\tilde{T}'(S) > 1$ for $S_{con} > 0$ (for which it is required that $\mu > 0$) and the last equality is the center's FOC in the bail-out. Since $U^{c,lim}_c(\cdot)$ is continuous, [4.17] implies that $U^{c,lim}_c > 0$ for C close to C_{lim}, meaning that the effect is initially positive.[29] Intuitively, the center gains utility from Region 1 increasing its consumption marginally above the bail-out level since this consumption is taken out of Region 1's wealth and thus entails no deadweight loss. Higher pre-bailout utility gives the center incentives to delay the bail-out, i.e. to increase S.

However, there is a second, negative effect, captured by the term $U^{c,lim}_s$. Higher C hastens the bail-out and makes it more costly for the center (in terms of S) to delay the bail-out, making marginal-utility considerations more important in the center's marginal calculus. These considerations tell the center to decrease S to minimize the dead-weight loss, since $U^{c,lim}_s = -[\tilde{T}'(S) - 1]/\rho < 0$.

Similar to what occurred for Region 1, observe in [4.16] that a hike in C increases the center's NMB if and only if the center would benefit from an increase in C that is financed taxing Region 2, i.e. if and only if $\mu/C > \tilde{T}'(S)$). Now, when fixing $S = S_{con}$ and increasing C above C_{con}, the center's bail-out FOC implies that $NMB^c_c(C, S_{con}) = \frac{\mu}{C} - \frac{\tilde{T}'(S_{con})}{\rho} < \frac{\mu}{C_{con}} - \frac{\tilde{T}'(S_{con})}{\rho} = 0$. This implies that the center best-responds setting $S < S_{con}$ to any $C > C_{con}$, since $-NMB^c_S < 0$. Intuitively, if Region 1 overconsumes in the eyes of the center ($C > C_{con}$), then the center has an incentive to minimize tax distortions and hasten the bail-out by lowering the subsidy below the bail-out level ($S_{lim} < S_{con}$).

Equilibrium. The unique intersection of the two best responses in Fig. 4 occurs at (C_{lim}, S_{lim}), which is the Nash equilibrium of the limit game. We see that a lens opens northwest of the equilibrium with allocations that Pareto-dominate the Nash equilibrium.[30] This is not surprising, since any efficient allocation should satisfy smoothness

[29] The effect then turns negative for values $C > \mu\rho$.

[30] Note that in fact *any* allocation in the limit game is Pareto-dominated by some allocation closer to (C_{con}, S_{con}).

of consumption paths and tax distortions. The existence of the Pareto lens tells us that both parties would benefit from binding agreements that i) lower Region 1's pre-bailout consumption and ii) increase the center's pre-bailout subsidies. However, our model tells us that any such pre-bailout agreements would be feeble (as indeed they are in reality) since at least one party is prone to deviate from the agreement. Consider, for example, agreements on the center's best response north-east of the bail-out equilibrium; under these, Region 1 would have incentives to unilaterally deviate and increase its consumption. On the other hand, agreements in the Pareto lens that lie on Region 1's best response would be immune to consumption deviations by the poor region, but the center would be tempted to tax (and re-distribute) below the agreed-upon level. Once the bail-out occurs, efficiency is restored: The poor region being borrowing-constrained gives the center the leverage to control the poor region's consumption and thus the incentive to tolerate a higher tax burden on the rich region.

Instability of smooth paths. Figure 4 shows that a bail-out with discontinuous C- and S-paths is one possible equilibrium outcome at $P = 0$. However, the figure also insinuates that there is, of course, the possibility that both agents' policies are continuous, In fact, such smooth paths are consistent with value-matching.[31] In our numerical simulations, however, we never find such *smooth equilibria*. The reason is that the smooth equilibrium is unstable. Note in the lower panel of Figure 4 that if the center sets a subsidy slightly below S_{con}, then Region 1 reacts with a large increase in consumption. The center, in turn, reacts with a further decrease of the subsidy, the dynamics leading towards the stable bail-out equilibrium. In fact, Region 1's best response has infinite slope at the smooth equilibrium, whereas the slope of the center's response is finite. This can be shown taking a second-order Taylor expansion of Region 1's value-matching condition [4.10], which shows that along Region 1's best response it holds that

$$\begin{aligned} &NMB(C_{con} + \mathrm{d}C, S_{con} - \mathrm{d}S) \\ &\quad \simeq \underbrace{NMB_C(C_{con}, S_{con})}_{=0} \mathrm{d}C - NMB_S(C_{con}, S_{con})\mathrm{d}S + \tfrac{1}{2}NMB_{CC}(C_{con}, S_{con})\mathrm{d}C^2 \\ &\quad - NMB_{CS}(C_{con}, S_{con})\mathrm{d}C\mathrm{d}S + \tfrac{1}{2}NMB_{SS}\mathrm{d}S^2 = 0. \end{aligned} \tag{4.18}$$

This equation implies that if the center deviates below S_{con} by a $\mathrm{d}S$ of order Δ^2, then Region 1 responds in the limit by a $\mathrm{d}C$ that is of order Δ.[32] In Appendix C, we show that in turn the *center*'s best response is approximated, to a second order, by a line with finite slope.

Main result. We summarize the above discussion in the following proposition, which is our main result. The logic of the proof is fully contained in the previous analysis; however, we also provide a more concise mathematical proof in the appendix.

[31] But note that smooth paths are not an equilibrium of the limit game since payoff is ill-defined if $\dot{P} = 0$.

[32] The highest-order term in $\mathrm{d}S$ is the one in $\mathrm{d}S^1$. The highest-order term in $\mathrm{d}C$ is the one in $\mathrm{d}C^2$, which implies the claim.

Proposition 4 (Bail-Outs)**.** *Assume that redistribution is costly (Ass. 1) and that* $\mu \in (0,1)$*. In any A-linear MPE in which* $P = 0$ *is locally approached, i.e.* $\dot{P}_{lim} \leq 0$*, and in which Region 1 consumes at least what the center would assign to it when broke, i.e.* $C(0) \geq S_{con}$*, we have* $\dot{P}(0) = 0$ *and* $S(0) = C(0) = S_{con}$*. Also, one of the following must occur:*

1. *(non-smooth paths) We have* $\dot{P}_{lim} < 0$*. Then* (C_{lim}, S_{lim}) *are an equilibrium of the limit game. In the limit game, a bail-out is the best response to a consumption drop by Region 1, i.e. a strategy such that*

$$C_{lim} = \lim_{P \to 0} C(P) > C(0) = S_{con}, \quad [4.19]$$

has the center best-responding by

$$S_{lim} = \lim_{P \to 0} S(P) < S(0) = S_{con}. \quad [4.20]$$

Vice versa, a downward drop in consumption is Region 1's best response to a bail-out, i.e. given that the center plays a strategy satisfying [4.20]*, Region 1's best response is such that* [4.19] *holds.*

2. *(smooth paths) We have* $\dot{P}_{lim} = 0$ *and* $C_{lim} = S_{lim} = S_{con}$*. However, this allocation is unstable in the limit game in the following sense. A second-order decrease of the subsidy to* $S_{con} - \Delta^2$*, for* Δ *small, entails a first-order increase in Region 1's consumption as a best response (*$C_{con} + \mathcal{O}(\Delta)$*), to which the center would react*

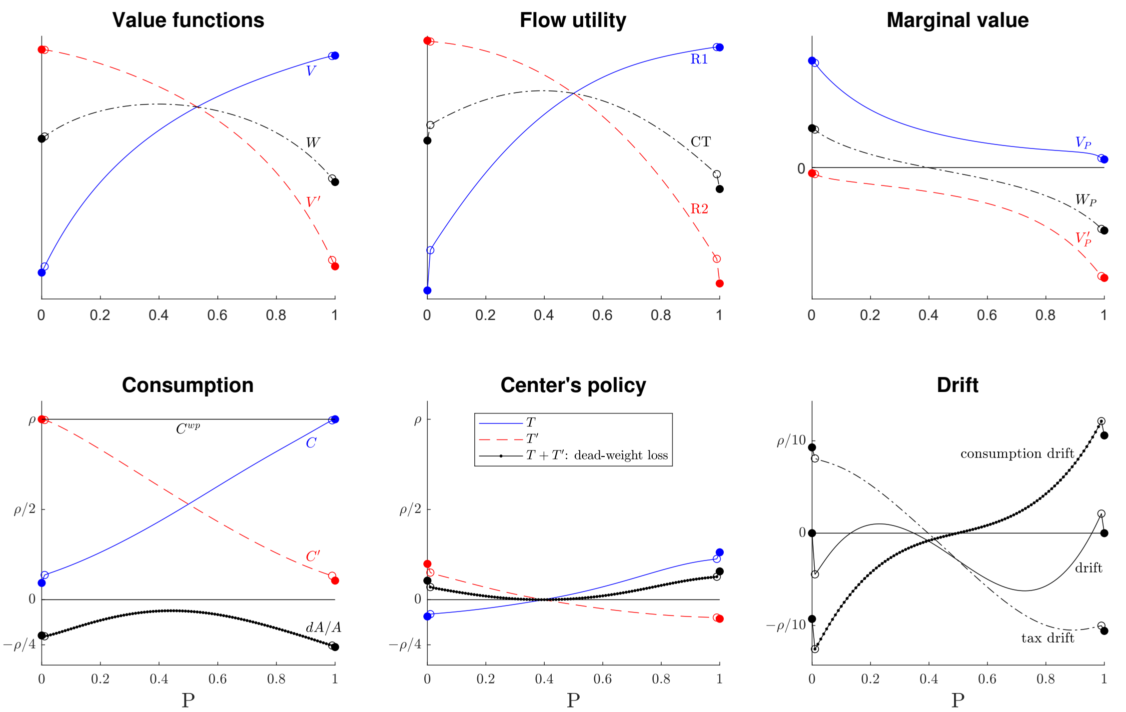

Figure 5: Equilibrium with bail-outs with $\mu = 0.4$. Rest of parameters as in Fig. 2

Dynamics. Fig. 5, right panel, reveals important changes to the dynamics with respect to the symmetric case: i) the interior stable steady state moves towards $P = 0$, ii) the impoverishment interval of Region 1 grows, whereas iii) the impoverishment interval of the favored Region 2 shrinks. Fig. 6 (bottom-center) shows that the stable central steady state moves left faster than μ as the weight on Region 1 decreases and eventually disappears entirely once the center favors one region too much. In this sense, our model predicts that political balance between regions lowers the likelihood of bail-outs and favors stability.

Bail-outs. Unsurprisingly, the center provides a more generous bail-out to the favored region than to the less-liked region (Fig. 5, center panel, and Fig. 6, top right), which causes the depletion of resources ($-\dot{A}$) to be most severe when the favored region is bailed out (Fig. 6, top center). Interestingly, the size of the jump discontinuities at the bail-out, measured in percentage terms, decrease as the center puts more weight on a region and the interests of the two players become more aligned (Fig. 6, bottom right).

Free-riding. Returning now to the stable region at the interior of the state space, the left-bottom panel of Fig. 6 shows that the maximum value of the center is attained when taxes are zero and the deadweight loss vanishes, which occurs close the the $P = \mu$ line. The same panel shows that close to this point free-riding of regions, as measured by asset depletion $-\dot{A}$, is also minimized. The top-center panel of Fig. 6 shows that the economy that gets closest to efficient resource use ($\dot{A} = 0$) is the one in which the center is balanced ($\mu = 1/2$), again confirming that balance between regions' political power is desirable. We conjecture that this relative efficiency is due to the fact that regions expect transfers to be low for a long time; Appendix J derives a Kolmogrov Backward Equation for the expected time to the first bail-out and shows that it is maximized (at about 150 years) when $\mu = 1/2 = P$, confirming this intuition.

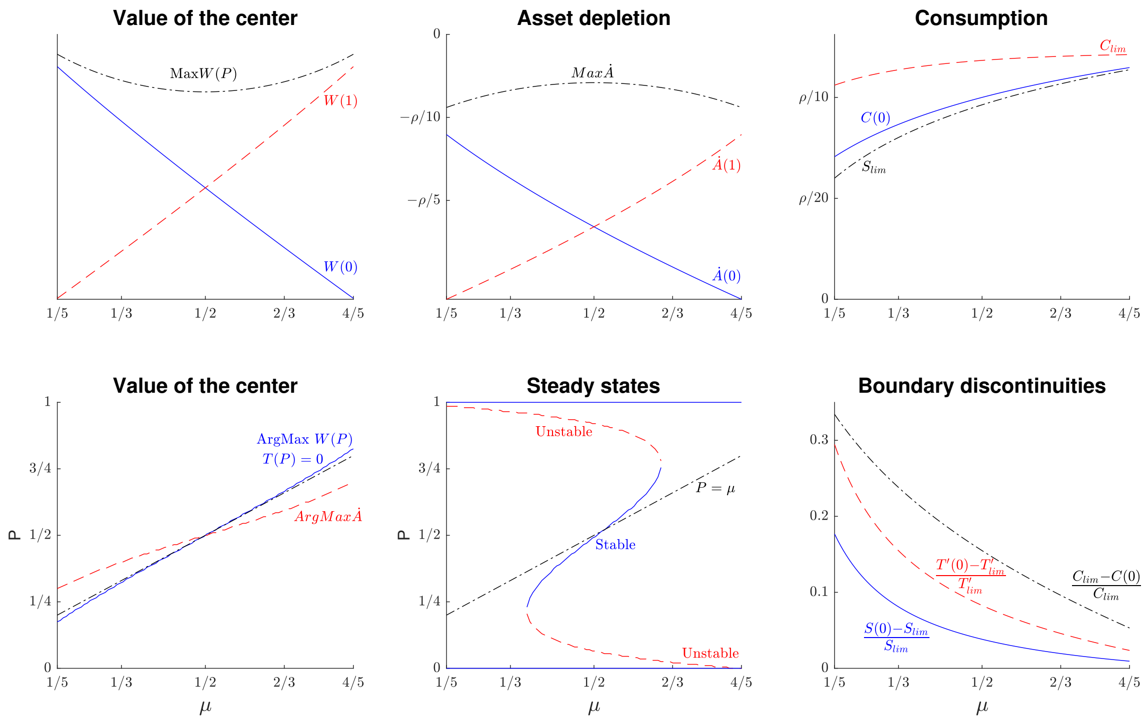

Figure 6: Comparative statics in μ. Model parameters as in Fig. 2, varying $\mu \in [0.2, 0.8]$.

4.5 Voluntary participation

Finally, it is interesting to ask under which conditions it is optimal for regions to delegate authority to a central authority. In other words, we ask: How large are the gains in risk-sharing that a central authority in our setting can offer compared to the costs implied by the loss of sovereignty? To answer this question, we compare the value of participating in a union with weight μ (which we will interpret as Region 1's political power) and the value in an environment with the center is committed to zero transfers forever (independence thereafter).[33]

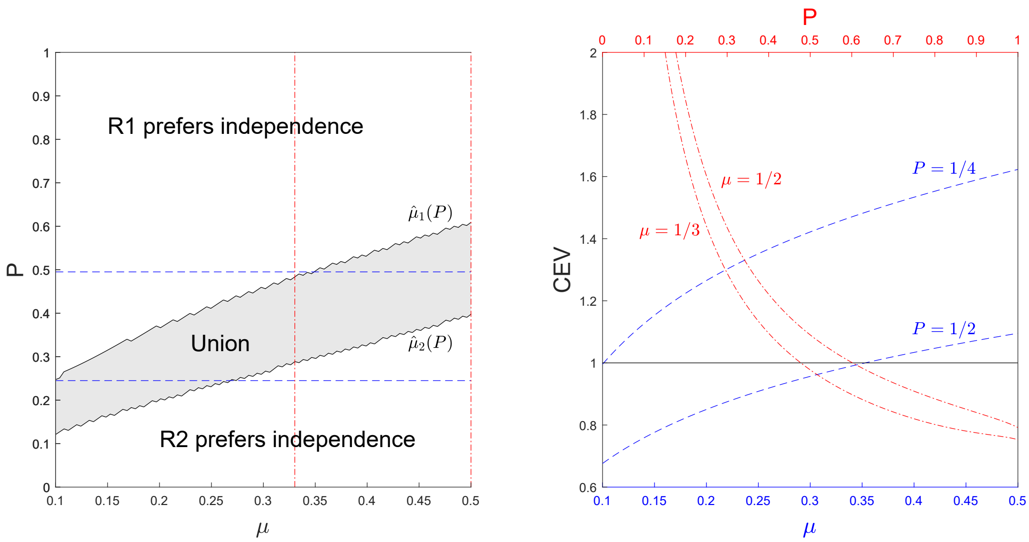

Figure 7: Voluntary participation by P and μ. Model parameters as in Fig. 2. In the *independence* scenario, regions have access to emergency resources of $10^{-6}C_{WP}$ when broke.

The left panel of Fig. 7 shows for which values of μ an agreement between the two regions is possible for fixed values of initial relative wealth P. Given initial wealth share P, Region 1 is willing to participate in a union if given political power above the threshold $\hat{\mu}_1(P)$, which is increasing in P. Similarly, Region 2 (with initial wealth $1-P$) is only willing to join the union if Region 1's political power is below $\hat{\mu}_2(P)$. The existence of a wide shaded region (in which a union is preferable for both regions) shows that the risk-sharing agreement is very valuable under our numerical specification.

The right panel of Fig. 7 quantifies Region 1's benefit of joining the union, showing the relative increase in consumption under independence (for all time and all states of the world) needed to match the welfare that Region 1 obtains by joining the union. We trace out this consumption equivalent variation (CEV) following the dashed blue and red

[33] We note here that independence does not mean that the two regions' problems are entirely separate in this environment. Under our volatility specification, [2.4], the other region's wealth influences the shock variance and thus the other region's wealth remains a state. Also, for computational reasons, we add a small exogenous endowment at the boundaries, ensuring that regions can have positive consumption when broke.

lines in the left panel. The following results stand out: First, CEVs are large, indicating the high value of the risk-sharing agreement implied by delegation to the center, despite the bail-out inefficiency. Second, political power (higher μ) is highly valued by regions. Third, and very interestingly, the gains from joining the union for a poor region (e.g. $P = 1/4$) are very large, the CEV exceeding 1.3 when approaching a political weight of $\mu = 1/4$ (which Region 2 would still accept). These gains by far exceed the gains of equally rich regions ($P = 1/2$) entering a symmetric risk-sharing agreement ($\mu = 1/2$), which are only about a third as large (CEV$\simeq 1.1$). However, once the wealth distribution is so imbalanced that an early bail-out becomes very likely, it becomes hard to convince the rich region to join the union, even when offering it large political power.

5 Costless redistribution: expropriation and indeterminacy

Assuming costly redistribution is reasonable in our main application of a fiscal federation, but it looses appeal in applications where resources can be transferred without major costs between two parties (say in family or in a firm). In Appendix K, we study a setting with costless redistribution, i.e. $X(\tilde{\tau}) = \tilde{\tau}$.

We find two kinds of equilibria. In the first, the center immediately expropriates the less-liked region, using a large lump-sum transfer. In a second type of equilibrium, bankruptcy can be avoided. However, this type of equilibrium is fragile: It exists only in the knife-edge case that the center puts exactly the same welfare weight on both regions ($\mu = 1/2$). We show that there exists a continuum of such equilibria, varying in the generosity of transfers to Region 1.

Our results thus indicate that delegating power to a center with a cheap transfer technology can be dangerous for politically weak agents, as the center has incentives to expropriate weak agents in order to control their actions.

6 Conclusions and outlook

This paper has provided a tractable setting to study the interaction of two regions and a center in a fiscal federation. The simple set-up here allowed us to reduce the dimensionality of the game's state space to one, thus allowing for tight characterizations of non-classical behavior at the boundaries. Our main finding is that inefficient bail-outs –sudden drops in consumption, paired with a sudden increase in transfers– occur. Despite these inefficiencies, there is still ample room for welfare gains from risk-sharing; our setting provides a theory for when economic agents find it optimal to delegate power to a central authority.

The techniques used here make the model amenable to extension; the most natural extension is arguably a mean-reverting productivity state for regions, say in the form of a Poisson process on a finite set of states. We conjecture that this extended model would feature temporary bail-outs: A poor region with low productivity would enter a bail-out regime when hitting the borrowing constraint, but leave this the regime once it switches to a higher productivity.[34] The numerical techniques used here (Markov-Chain Approximation Methods) carry over, being both easy to implement and highly efficient. This opens the door for a more serious quantitative model of fiscal federations.

A Center's problem

Unconstrained problem. Consider the center's problem at $P \in (0,1)$ in [4.1]. We note that the objective function in the problem is linear and the set of admissible controls is strictly convex, since $X(\cdot)$ is strictly convex. Hence the first-order conditions (FOCs) are both necessary and sufficient:

$$W_P > 0: \qquad X'(T'_{unc}/\rho) = B(P, W_P), \qquad \text{[A.1]}$$

$$W_P \le 0: \qquad X'(T_{unc}/\rho) = \frac{1}{B(P, W_P)}, \qquad \text{[A.2]}$$

where $B(P, W_P) = \frac{1/\rho - PW_P}{1/\rho + (1-P)W_P}$. Since $X'(\cdot)$ is a decreasing function, these first-order conditions uniquely pin down the tax on the region that the center wants to distribute away from.[35] The transfer to the other region can then be obtained from the government's budget constraint. The comparative statics are as expected (focus on the case $W_P > 0$): The higher W_P, the more the center wants to re-distribute to Region 1, and the lower $B(P, W_P)$ is. In the optimum, the center tolerates a lower marginal revenue from taxing Region 2, corresponding to a higher tax on Region 2 (T'_{unc}), see [A.2].

Under the parametric specification [2.6] for the tax-revenue function $X(\cdot)$ that we use in the numerical simulations, we obtain the following closed-form solutions for transfers:

$$W_P > 0: \quad T'_{unc} = \frac{\rho}{\alpha}\big[B(P, W_P)^{-1/\kappa} - 1\big], \quad -T_{unc} = \frac{\rho}{\alpha(1-\kappa)}\big[B(P, W_P)^{-\frac{1-\kappa}{\kappa}} - 1\big];$$

$$W_P \le 0: \quad T_{unc} = \frac{\rho}{\alpha}\big[B(P, W_P)^{1/\kappa} - 1\big], \quad -T'_{unc} = \frac{\rho}{\alpha(1-\kappa)}\big[B(P, W_P)^{\frac{1-\kappa}{\kappa}} - 1\big]. \qquad \text{[A.3]}$$

We will now derive these solutions for the case $W_P > 0$; the case $W_P \le 0$ can be solved in an analogous manner. Given the particular form of the tax-revenue function in [2.6], the

[34] In a similar vein, Barczyk and Kredler (2014b) showed how to extend a stylized one-dimensional model of two altruistic savers to a four-dimensional setting, in which the state consists of both agents' wealth and income.

[35] Note that it is in principle possible that $B(P, W_P) \le 0$. Then, if the marginal revenue from taxation is always positive, i.e. $\lim_{\tilde{\tau}\to\infty} X'(\tilde{\tau}) \ge 0$ (which is indeed the case for the specification we choose in our numerical solutions, where this limit is zero), a maximizer does not exist – it becomes unboundedly large. This corresponds to a situation in which the center is so keen to re-distribute towards Region 1 that the center would even take away output from Region 2 output and destroy it; in our numerical simulations this situation never occurs. Existence of a bounded maximizer can be guaranteed when specifying $X(\cdot)$ such that the $\lim_{\tilde{\tau}\to\infty} X'(\tilde{\tau}) = -\infty$; this could be rationalized as a Laffer Curve under which the negative effect of taxation on output becomes very strong at some point. Numerically, we experimented with a quadratic Laffer curve, i.e. the specification $X(\tilde{\tau}) = \tilde{\tau} - \beta\tilde{\tau}^2$ with $\beta > 0$. However, this specification led to less-stable solutions, presumably since the curvature was too weak.

first-order condition [A.1] implies that for the optimal tax T'_{unc} on Region 2 we have

$$1 + \frac{\alpha T'_{unc}}{\rho} = B(P, W_P)^{-1/\kappa}. \tag{A.4}$$

Solving for T'_{unc} immediately gives us the solution for T'_{unc} for the case $W_P > 0$ in Eq. [A.3]. To solve for T_{unc}, note that the government budget constraint tells us that $-T_{unc}/\rho = X(T'_{unc}/\rho)$ and thus

$$-T_{unc}/\rho = \frac{\left(1 + \alpha T'_{unc}/\rho\right)^{1-\kappa} - 1}{\alpha(1-\kappa)}. \tag{A.5}$$

Now, use Eq. [A.4] to find the closed form solution for T_{unc} in the case $W_P > 0$ in [A.3].

Center's combined problem at $P = 0$. Under Ass. 1, the constraint problem has a unique (positive) solution that can be found substituting the budget constraint $S = \rho X(T'/\rho)$ into the FOC [4.3].

Writing the center's payoff at $P = 0$ as a function of S and taking into account the poor region's response in the consumption stage from [3.9], we obtain that the center aims to maximize the following Hamiltonian) function $H_0(S)$:

$$\begin{aligned} H_0(S) =& \mu \ln(\min\{S, C_{unc}(0)\}) + (1-\mu)\ln(C^{wp}) \\ & - [C^{wp} + \min\{S, C_{unc}(0)\} + \tilde{T}'(S) - S]\tfrac{1}{\rho} + \underbrace{\max\{0, S - C_{unc}(0)\}}_{=\dot{P}} V_P(0), \end{aligned} \tag{A.6}$$

where C_{unc} is the solution to [3.8] and C^{wp} is wealth pooling consumption ($= \rho$). At the point $S = C_{unc}(0)$, where the poor region starts saving, the function H_0 generically has a kink (but is continuous); the left and right derivative at this point are

$$\nabla^- H_0(C_{unc}) = \frac{\mu}{C_{unc}} + D, \qquad \nabla^+ H_0(C_{unc}) = V_P(0) + 1/\rho + D,$$

where the term $D = (d\tilde{T}'/dS)/\rho$ is the same on both sides of the kink. Thus the function H_0 may kink upward, downward, or be smooth at $S = C_{unc}$, depending on the sign of $V_P(0) + 1/\rho - \mu/C_{unc}$. In general, H_0 may attain its global maximum at a local maximum on the range $S \in (0, C_{unc})$, at the point $S = C_{unc}$ (in case there is an upward kink), or at a local maximum on the range $S > C_{unc}$. If $\mu > 0$, the corner $S = 0$ cannot be optimal by the Inada condition on the utility function.

B Equivalence of limit-game and value-matching conditions

Lemma 1 (Limit game and value-matching conditions)**.** *Consider a profile of A-linear policies and value functions, $\mathcal{P} = \{V, V', W; C, C', T, T'\}$. Denote $C_{lim} = \lim_{P\to 0} C(P)$, $S_{lim} = -\lim_{P\to 0} T(P)$ etc. Assume that $C_{lim} > S_{lim}$ and that all agents' policies satisfy their FOCs approaching $P = 0$, i.e. $1/C_{lim} = 1/\rho + V_P(0)$, $1/C'_{lim} = 1/\rho$ and $X'(-S_{lim}/\rho) = 1/(1 + \rho W_P(0))$. Then the following two statements are equivalent:*

1. *(C_{lim}, S_{lim}) is an equilibrium of the limit game given boundary values $V(0)$ and $W(0)$.*

2. *$\mathcal{P}$ satisfies value-matching at $P = 0$, i.e. $V(0) = V_{lim}$, $W(0) = W_{lim}$, and $\mathcal{P}$ satisfies all players' HJBs in the limit as $P \to 0$.*

> **Proof:**
> We will show that Statements 1 and 2 in the lemma are equivalent to the pair of equations [4.10] and [4.12] holding.
>
> First, note that both players' payoffs in the limit game are strictly concave on the range range that the limit game considers ($C > S$), as evidenced by Eq. [4.13] and [4.15]. Thus the limit-game FOCs [4.9] for Region 1 and [4.11] for the center are necessary and sufficient for a limit-game equilibrium, which shows that Statement 2 is equivalent to [4.10] and [4.12] holding.
>
> Second, Statement 2 implies [4.10] and [4.12] when invoking the limiting FOCs $1/C_{lim} = 1/'\rho + V_P(0)$ and $X'(-S_{lim}/\rho) = 1/(1 + \rho W_P(0))$. The converse also holds, since we restrict attention tor profiles for which the limiting FOCs to hold.
>
> ■

C Approximation of limit-game responses

Lemma 2 (Limit-game best responses)**.** *Given the allocation (C_{con}, S_{con}) at $P = 0$, second-order accurate approximations for $(C_{con} + dC, S_{con} - dS)$ in the limit game yield:*

1. *Region 1's best response is approximated by*

$$dC = -dS + \sqrt{2C_{con}^2 \left(\frac{1}{C_{con}} - \frac{\tilde{T}'(S_{con})}{\rho} \right) dS + \left(1 + \frac{\tilde{T}''(S_{con}) C_{con}^2}{\rho} \right) dS^2}, \tag{C.1}$$

   thus $dC \propto \sqrt{dS}$ for small dS and the best response has infinite slope at (C_{con}, S_{con}).

2. *The center's best response is approximated by*

$$dS = \left(\sqrt{1 + \frac{\mu\rho}{C_{con}^2 \tilde{T}''(S_{con})}} - 1 \right) dC \tag{C.2}$$

   and is locally linear with finite negative slope.

Proof:
Region 1's best response. Calculating the derivatives in the second-order expansion [4.18] for Region 1's best response and multiplying by $-C_{con}^2$ yields

$$\tfrac{1}{2}\mathrm{d}C^2 + \mathrm{d}S\mathrm{d}C + C_{con}^2\left[\frac{\tilde{T}'(S_{con})}{\rho} - \frac{1}{C_{con}}\right]\mathrm{d}S - \tfrac{1}{2}\frac{\tilde{T}''(S_{con})}{\rho}C_{con}^2\mathrm{d}S^2 = 0.$$

For given $\mathrm{d}S$, solve for $\mathrm{d}C$ using the quadratic formula to obtain [C.1].[36] $\mathrm{d}C \propto \sqrt{\mathrm{d}S}$ follows since below the square root in [C.1] the term in $\mathrm{d}S$ dominates, from which the claim on the infinite slope immediately follows.

Center's best response. A second-order Taylor expansion of $NMB^c(\cdot)$ gives[37]

$$\begin{aligned}NMB^c(C_{con}+\mathrm{d}C, S_{con}-\mathrm{d}S) \simeq\ & \underbrace{NMB_C^c(C_{con},S_{con})}_{=0}\mathrm{d}C - \underbrace{NMB_S^c(C_{con},S_{con})}_{=0}\mathrm{d}S \\ & + \tfrac{1}{2}NMB_{CC}^c\mathrm{d}C^2 - NMB_{SC}^c\mathrm{d}C\mathrm{d}S + \tfrac{1}{2}NMB_{SS}^c\mathrm{d}S^2 \\ =& -\tfrac{1}{2}\frac{\mu}{C_{con}^2}\mathrm{d}C^2 + \frac{\tilde{T}''(S_{con})}{\rho}\mathrm{d}C\mathrm{d}S + \tfrac{1}{2}\frac{\tilde{T}''(S_{con})}{\rho}\mathrm{d}S^2 = 0.\end{aligned} \quad [C.3]$$

Calculating the derivatives and multiplying by $\rho/\tilde{T}''(S_{con})$ yields the quadratic form

$$\tfrac{1}{2}\mathrm{d}S^2 + \mathrm{d}C\mathrm{d}S - \tfrac{1}{2}\underbrace{\left[\frac{\mu\rho}{C_{con}^2\tilde{T}''(S_{con})}\right]}_{\equiv Z_{con}}\mathrm{d}C^2 = 0, \quad [C.4]$$

where $Z_{con} > 0$ since $\tilde{T}'' > 0$. Solving for $\mathrm{d}S$ given $\mathrm{d}C$ yields[38]

$$\mathrm{d}S = \left(\sqrt{1+Z_{con}} - 1\right)\mathrm{d}C, \quad [C.5]$$

Eq. C.5 together with $Z_{con} > 0$ implies the claims about the finite slope and linearity. ■

D Proof of Prop. 4 (main result on bail-outs)

We will first show that $S(0) = C(0) = S_{con}$. $\dot{P}_{lim} \leq 0$ reveals that the center is not interested in creating a positive drift at zero, i.e. it implies that the center's payoff H_0 in the problem [A.6] is decreasing for $S \geq C_{unc}(0)$ and so that the center optimally sets $S(0) \leq C_{unc}(0) = C_{lim}$. Now, since we assumed $C(0) \geq S_{con}$, it must be that $C_{unc}(0) \geq S_{con}$ by Region 1's optimality at $P = 0$, see [3.9]. It then follows that on the range $S \in [0, C_{unc}(0)]$, the center's payoff H_0 is uniquely maximized at $S = S_{con}$, see the center's constrained problem [4.2]. From this we conclude that $S(0) = S_{con} = C(0)$ and thus $\dot{P}(0) = 0$.

1. Non-smooth paths: We first cover the case when $\dot{P}_{lim} < 0$. C_{lim} and S_{lim} are an equilibrium of the limit game by Lemma 1.

We will now show that setting $C_{lim} > S_{con}$ is a best response to a subsidy $S_{lim} < S_{con}$. Region 1's best response is characterized by $NMB(C_{lim}, S_{lim}) = 0$, see [4.12]. Note that

$$-NMB_S(C_{con}, S) = \frac{1}{C_{con}} - \tilde{T}'(S) > \frac{\mu}{C_{con}} - \tilde{T}'(S_{con}) = 0 \qquad \text{for } S > S_{con}, \tag{D.1}$$

where the inequality follows from the assumption $\mu < 1$ and the fact that $\tilde{T}(\cdot)$ is convex; the last equality follows from the center's first-order condition in the constrained problem at $P = 0$, Eq. [4.3]. It thus follows that $NMB(C_{con}, S) > 0$ for $S < S_{con}$. Since $NMB_C < 0$ by Eq. [4.13], it follows that Region 1's best response satisfies $C_{lim} > C_{con} = S_{con}$.

Finally, we show that the center best-responds by setting $S_{lim} < S_{con}$ to any $C_{lim} > S_{con}$ in the limit game. A best response satisfies $NMB^c(C_{lim}, S_{lim})$, see Eq. [4.12]. Now, $NMB^c(C, S_{con}) < 0$ for any $C > S_{con}$ since

$$NMB^c_C(C, S_{con}) = \frac{\mu}{C} - \frac{\tilde{T}'(S_{con})}{\rho} < 0, \tag{D.2}$$

where the inequality follows again from the center's FOC [4.3]. Since $NMB^c_S < 0$ by Eq. [4.15], this implies that the center's best response satisfies $S_{lim} < S_{con}$.

2. Smooth paths: Since we assumed $\dot{P}_{lim} \leq 0$, the only case that remains to cover is $\dot{P}_{lim} = 0$. $\dot{P}_{lim} = 0$ immediately implies $S_{lim} = C_{lim}$ by the law of motion [3.4]. From optimality of Region 1's consumption at $P = 0$, it follows that $C_{lim} = C_{unc}(0) \geq C(0)$, see [3.9]. We will now rule out that $C_{lim} > C(0)$. Since $\dot{P}_{lim} = 0$, this would imply that $S_{lim} > S(0)$. So $S_{lim} > S(0)$ attains the maximum in the center's unconstrained problem [4.1] at $P = 0$; but note that this means that the center should also prefer S_{lim} over $S(0)$ at zero, since both options lie on the unconstrained part of the center's problem. This is a contradiction and we conclude that $C_{lim} = C(0 = S_{lim} = S(0)$.

The claims about instability follow directly from the approximations of the best responses in Lemma 2 together with the lower panel of Fig. 4. ■

E Computational appendix

We solve for the equilibrium using the Markov-chain approximation method.[39] We use an equally spaced grid for P with 101 grid points. The method approximates the diffusion for P by a trinomial process, i.e. a Markov chain on the grid for P that either stays at the same grid point or jumps to one of the two adjacent states. The transition probabilities are chosen such that the first and second moments of the innovations are matched to the underlying diffusion. The method is equivalent to an explicit finite-difference method with centered differencing for both first and second derivatives.

[39] See Barczyk and Kredler (2014b) and their computational appendix for an introduction to the the Markov-chain approximation methods with an application to a continuous-time savings problem.

To obtain a final guess for the value functions, we use a consumption function that is linear in P with end points $C(0) = C_{con}$ (where C_{con} comes from the center's constrained problem [4.2]) and $C(1) = C^{wp} = \rho$. We choose a linear consumption function constructed for Region 2 with $C'(0) = C^{wp}$ and $C'(1) = C'_{con}$, where C'_{con} comes from the center's constrained solution at $P = 1$. We choose the tax on the rich region be linearly connecting $T = 0$ at $P = 0.5$ with T_{con} at $P = 1$ (the tax on the rich region in the center's constrained solution at $P = 1$), or with T'_{con} at $P = 0$. We then back out the corresponding subsidy to the poor region from the government budget constraint. Given these policy guesses, we calculate the value functions for all agents that result from these policies being played forever, which can be efficiently solved for by solving a sparse system of equations.

We then go backward in time from the final guess, choosing the maximal time increment that satisfies the Courant-Friedrich-Lewy stability condition (which simply says that no jump probability of the Markov chain can be negative). In line with our specification of the approximating chain, we calculate centered difference quotients to approximate the first value-function derivatives within the state space. At $P = 0$, we use the upward quotient and at $P = 1$ the downward quotient. Given these derivatives, we obtain agents' optimal policies for all interior grid points from the FOC [3.8] of Region 1, the one for Region 2 and [A.1] for the Center.

At the boundary $P = 0$, we proceed as follows (the computations at $P = 1$ are analogous). We obtain $C_{unc}(0)$ from Eq. [3.9], which completely characterizes R1's response in the consumption stage. Region 2's policy is always $C'(0) = \rho$. To obtain the center's policy, following the results in Appendix A, we first calculate the payoffs H_0 defined in Eq. [A.6] for the following three scenarios:

1. Obtain the S_{con} that solves the constrained problem [4.2] by solving Eq. [4.3] with a root-finding routine. This step can be done outside the time loop. If $S_{con} \leq C_{unc}$, this is a local maximum and we obtain its payoff $H_{con} = H_0(S_{con})$. If $S_{con} > C_{unc}$, then the center is better off giving a subsidy on the range $S \geq C_{unc}$ and we assign $H_{con} = -\infty$.

2. Obtain the payoff $H_0(C_{unc})$ at the kink.

3. Obtain the S_{unc} that solves the problem when Region 1 is unconstrained from Eq. [A.3]. If $S_{unc} \geq C_{unc}$, this is a local maximum and we assign $H_{unc} = H_0(S_{unc})$. If $S_{unc} < C_{unc}$, the center will be better off giving a subsidy on the range $S \leq C_{unc}$ and we assign $H_{unc} = -\infty$.

Finally, pick the maximum out of H_{con}, $H_0(C_{unc})$, and H_{unc} and the associated maximizer as the center's equilibrium policy $S(0)$. Then back out the corresponding tax on Region 2, $T'(0)$, from $\tilde{T}(S) = \rho X^{-1}(S/\rho)$.

Given agents' policies, calculate the drift at each grid point from Eq. [3.4]. Given the drift and the volatility, see Eq. [2.4], the jump probabilities for the approximating Markov chain are calculated and all value functions are updated.

Go backward in time until a convergence criterion for policies is met.

F Efficient allocations and (non)-existence of efficient equilibria

Proof of Proposition 1
Consider a social planner who puts weight θ on Region 1, weight θ' on Region 2, and weight $(1-\theta-\theta')$ on the center, where $\theta, \theta' \in [0,1]$ and $\theta+\theta' \leq 1$. Thus the planner's criterion is

$$\mathbb{E}_0 \int_0^\infty e^{-\rho t}\Big\{\big[\underbrace{\theta+(1-\theta-\theta')\mu}_{\equiv\hat{\mu}}\big]\ln(c_t)+\big[\underbrace{\theta'+(1-\theta-\theta')(1-\mu)}_{=(1-\hat{\mu})}\big]\ln(c_t')\Big\}\mathrm{d}t. \quad \text{[F.1]}$$

Note that we have introduced here the parameter $\hat{\mu}$ from the proposition. Clearly, any pair of permissible weights (θ, θ') maps to a unique value for $\hat{\mu}$ and we have $\hat{\mu} \in [0,1]$. Also, any value of $\hat{\mu} \in [0,1]$ is associated to some permissible planning weights, as may be readily seen by setting $\theta = \hat{\mu}$ and $\theta' = 1-\hat{\mu}$.

We now turn to the resource constraint in the planner's problem. By Eq. [3.3], aggregate wealth evolves according to

$$\mathrm{d}A_t = (\rho - C_t - C_t')A_t\mathrm{d}t + \sigma A_t\underbrace{[\sqrt{P_t}\mathrm{d}B_t + \sqrt{1-P_t}\mathrm{d}B_t']}_{\equiv \mathrm{d}W_t^A}, \quad \text{[F.2]}$$

where we have written consumption strategies out of total wealth (i.e. $c_t = C_tA_t$) and where the term $T+T'$ cancels since $X(T) = T$ by Ass. 2. Now, observe that the shock to aggregate wealth defined in the last equation, $\mathrm{d}W_t^A$, is a standard Wiener Process: We have $\mathbb{E}_t\mathrm{d}W_t^A = 0$ and $(\mathrm{d}W_t^A)^2 = (P_t + 1 - P_t)\mathrm{d}t = \mathrm{d}t$. It is thus inessential which path $\{P_t\}$ the planner chooses; the stochastic properties of the resulting paths $\{A_t\}$ for aggregate wealth will be the same and depend solely on the consumption paths $\{C_t, C_t'\}$.

Maximizing [F.1] subject to the resource constraint [F.2] is a standard Merton-type consumption-savings problem. The associated HJB is

$$\rho W^{\hat{\mu}}(A) = \ln(A) + \max_{C,C'}\Big\{\hat{\mu}\ln(C)+(1-\hat{\mu})\ln(C')+(\rho-C-C')AW_A^{\hat{\mu}}(A)\Big\}+\sigma^2A^2W_{AA}^{\hat{\mu}}(A),$$

where $W^{\hat{\mu}}(A)$ is a $\hat{\mu}$-planner's value of having on hand aggregate wealth A. We now guess the functional form $W^{\hat{\mu}} = \alpha + \beta\ln(A)$; from Eq. [3.6], it is also clear that we must have $\beta = 1/\rho$. Using this guess in the HJB, we verify that it has the correct form and find the optimal consumption rates given in the proposition, $C^{eff}(\hat{\mu}) = \hat{\mu}$ and $C'^{eff}(\hat{\mu}) = (1-\hat{\mu})\rho$. The claim on the drift of A_t then follows directly from the law of motion [F.2]. ■

Proof of Theorem 2(Generic Inefficiency):
By contradiction. Suppose that the efficient allocation with weight $\hat{\mu}$ was implemented in an equilibrium. By Prop. 1, this implies that $C_t = \hat{\mu}\rho$ and $C'_t = (1-\hat{\mu})\rho$ for all t with probability one.

We will first rule out that the economy spends time at the bounds of the state space and then turn to the interior of the state space. Suppose that the economy spent a positive amount of time at $P_t = 1$ with positive probability. Then, by Region 1's FOC [3.8] and since we assumed $\hat{\mu} \in (0,1)$, we have $C(1) = \rho > \hat{\mu}\rho$, which violates efficiency. In the same fashion, the economy spending time at $P_t = 0$ can be ruled out by Region 2's FOC.

Consider now, without loss of generality, the case in which $P_{\bar{t}} \in (0,1)$ for some time $\bar{t} \geq 0$. Since $T(\cdot)$ and $T'(\cdot)$ are bounded and $\sigma > 0$, the path $\{P_t\}_{t=\bar{t}}^{\infty}$ will cover some neighborhood $\mathcal{P} = (P_{\bar{t}} - \epsilon, P_{\bar{t}} + \epsilon)$ around the initial state, see the law of motion [3.4]. Clearly, the consumption and value functions in $\mathcal{P}$ must then be constant and equal to the efficient levels in this neighborhood. Now, this means $V_P(P) = 0$ for all $P \in \mathcal{P}$, which in implies that the optimal policy is $C^*(P) = \rho$ for all $P \in \mathcal{P}$ by the FOC for consumption, Eq. [3.8]. But this again gives a contradiction, since $C^*(P) = \rho > \hat{\mu}\rho$. ■

(Non)-existence of efficient equilibria It is important to note here that inefficiency hinges crucially on the existence of shocks and the lack of commitment. If there is no uncertainty, the center can indeed achieve efficiency by committing not to transfer any resources ever. Another exception to the non-existence result for efficient equilibria occurs at the extremes of the Pareto Frontier, where only one region consumes.

G Wealth pooling

Proof of Proposition 3
The HJB for Region 1, taking as given a strategy $C'(A)$ by Region 2, is

$$\rho\tilde{V}^{wp}(A) = \ln(A) + \max_{C\geq 0}\left\{\ln(C) + [\rho - C - C'(A)]A\tilde{V}^{wp}_A\right\} + \sigma^2 A^2 \tilde{V}^{wp}_{AA}(A).$$

Guessing that the value function is of the form $\tilde{V}^{wp} = \alpha + \ln(A)/\rho$, where α is a coefficient that is to be determined, we find the optimal consumption rule $C^{wp} = C'^{wp} = \rho$ given in Proposition 3. It can then be verified that if Region 2 follows the same strategy and sets $C'(A) = \rho$ and that Region 1's HJB holds. We have thus found a symmetric equilibrium with value function

$$\tilde{V}^{wp}(A) = \frac{1}{\rho} + \frac{\ln(A)}{\rho} + \underbrace{\frac{\ln(\rho)}{\rho} - \frac{2\rho + \sigma^2}{\rho^2}}_{\equiv V^{wp}}. \qquad \text{[G.1]}$$

Here, the constant V^{wp} allows comparison to the equilibrium value function $V(P)$, see again Eq. [3.6]. Region 2's value is identical to Region 1's. ■

A question that arises naturally is if the WP consumption allocation can be supported as an equilibrium in the original game, i.e. when regions' property rights are intact. Indeed, we show that the WP consumption allocation *cannot* be supported as a MPE, i.e. in the case when the center chooses transfers optimally and lacks commitment.[40] The reason is that the center would use its power to discipline a region that has hit the borrowing constraint, at which point WP breaks down.

Proposition 5 (No MPE with WP)**.** *The WP strategies $C(P) = C'(P) = C^{wp}$ for all P cannot be supported as a MPE when the center lacks commitment.*

Proof:
Suppose, by way of contradiction, that there was a MPE in which regions play $C(P) = C'(P) = C^{wp}$ for all P. Suppose that $\mu > 0$ (otherwise, consider the equivalent argument for Region 2). Then, consider the center's problem at $P = 0$. Since consumption is constant, it must be that the center's value function satisfies $W_P = 0$. Now, from the center's HJB [3.10] it follows that the center's optimal strategy is to set $T(0) = -\mu\rho$ once Region 1 is broke, which obliges Region 1 to set $C(0) = \mu\rho < \rho = C^{wp}$. Thus the equilibrium is not WP – a contradiction. ■

H Comparative statics for volatility

In Fig. 8 we explore how the size of the shocks σ affects the equilibrium. As seen in the right panel, when uncertainty is sufficiently high, the stable equilibrium at $P = 1/2$ vanishes.

As expected, the value for the center decreases as uncertainty increases (left panel). Maybe surprisingly, the asset depletion decreases with σ. This suggest that regions have more incentives to save when wealth is evenly split. The reason is that the stakes are much higher when the stable dynamics around $P = 1/2$ disappear and regions are eager to avoid a situation in which the drift push them towards bankruptcy.

[40] However, it can be proven that the WP allocation *can* be supported as an equilibrium in the game between the regions if the center can commit to future (non-optimal) policies. For this, transfers to the broke region at the constraint have to be high enough such that C^{wp} is feasible for the broke region.

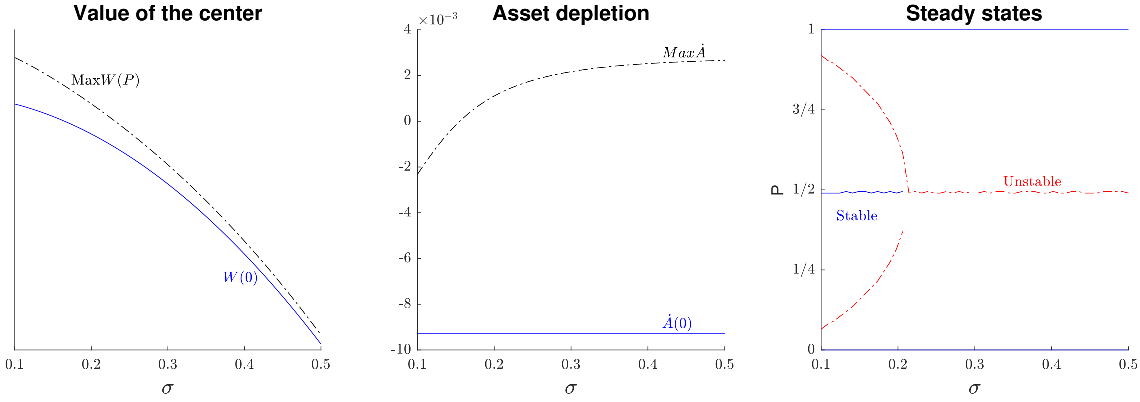

Figure 8: Comparative statics (σ). Model parameters as in Fig. 2 with $\sigma \in [0.1, 0.5]$.

I Boundary behavior

It turns out that it is quite complicated to determine if the process P_t ever reaches the boundary $P = 0$ or not. In this appendix, we derive a condition on the limiting drift $\dot{P}_{lim} = \lim_{P \to 0} \dot{P}(P)$ that determines if $P = 0$ is reached or not.

In general, we will consider a diffusion process P_t on the interval $(0, 1)$ induced by agents' policies according to the law of motion [3.4]. We will consider subsets of the state space, intervals (a, b) with $0 \leq a < b < 1$, and study how long it takes the process to exit the interval and at which bound the process exits, i.e. we let $P_0 \in (a, b)$. In the following, denote by $T_{(a,b)}$ the first (random) exit time of the process P_t from the interval (a, b), which is possibly infinite.

We first consider a simple process with constant drift and an approximating function for volatility and derive the condition on the drift; we then approximate the law of motion in the neighborhood of zero by this simple process. The following lemma follows Helland (1996), who provides a nice overview of the mathematical results that are available for the boundary behavior of one-dimensional diffusion processes (thus it is essential that we have reduced the dimensionality of the problem to 1). Intuitively, they rely on space and time transformations that are judiciously chosen so that the transformed process is a standard Wiener Process.[41] We reformulate the proof here for our convenience, explaining the key steps in more detail than the original text, which is quite dense; consult the original text and the sources cited therein, such as the classic probability textbook Breiman (1992), to fill in the details.

Lemma 3. *Let* $\mathrm{d}P_t = \mu \mathrm{d}t + \sigma\sqrt{P_t}\mathrm{d}W_t$ *be a diffusion on the interval* $(0, b)$ *with initial condition* $P_0 \in (0, b)$*. Let* T *be the random (possible infinite) exit time from the interval. Define the expected time to exit the interval from the point* P_0 *as* $S(P_0) = \mathbb{E}_0[\int_0^T \mathrm{d}t]$*. Then:*

1. *If* $\mu \geq \sigma^2/2$*, then the boundary* $P = 0$ *is inaccessible, i.e. for all* $P_0 \in (0, b)$ *we*

[41] See also the online lecture notes "Tricks for exit times and probabilities" by Matthias Kredler for a (hopefully) intuitive explanation of these methods that is based on approximating discrete Markov chains.

have

$$Prob(\textit{exit at } 0|P_0) = 0.$$

2. *If, however, $\mu < \sigma^2/2$, then $P = 0$ is accessible. Specifically:*
   (a) *$S(P_0)$ is finite for all $P_0 \in (0, b)$ and as $P_0 \to 0$, $S(P_0) \to 0$.*
   (b) *The exit probabilities from the interval $(0, b)$ are*

$$Prob(\textit{exit at } b|P_0) = (P_0/b)^{1-2\mu/\sigma^2},$$
$$Prob(\textit{exit at } 0|P_0) = 1 - (P_0/b)^{1-2\mu/\sigma^2},$$

and thus the probability to exit at 0 approaches 1 as $P_0 \to 0$.

Proof of Lemma 3

Space transformation: We will first look for a monotone transformation of space (P) that converts the process P_t into a martingale (or brings it to "natural scale", in the language of the mathematical literature on diffusions). Precisely, we look for an increasing, twice differentiable function $v(\cdot)$ such that the transformed process $U_t \equiv v(P_t)$ has the representation

$$\mathrm{d}U_t = s^u(U_t)\mathrm{d}W_t, \quad \textbf{[I.1]}$$

i.e. U_t is driftless, and where $s^u(\cdot)$ is to be determined.[42] Applying the Itô Rule to $v(\cdot)$, we have

$$\mathrm{d}U_t = \mathrm{d}v(P_t) = v'(P_t)\left[\mu\mathrm{d}t + \sigma\sqrt{P_t}\mathrm{d}W_t\right] + \frac{1}{2}v''(P_t)\sigma^2 P_t\mathrm{d}t. \quad \textbf{[I.2]}$$

Since we want U_t to be driftless, we now impose $\mathbb{E}_t[\mathrm{d}U_t] = 0$. Then, since $\mathbb{E}_t[\mathrm{d}W_t] = 0$, Eq. **[I.2]** implies that $v'(\cdot)$ must satisfy the following ODE for all $P \in (0, b)$:

$$v''(P) = -\tilde{\mu}\frac{v'(P)}{P}, \qquad \text{where we define} \quad \tilde{\mu} \equiv \frac{2\mu}{\sigma^2}.$$

A monotone increasing solution to this ODE is

$$v(P) = c_1 \begin{cases} \ln(P) & \text{if } \tilde{\mu} = 1, \\ \frac{P^{1-\tilde{\mu}}}{1-\tilde{\mu}} & \text{if } \tilde{\mu} \neq 1. \end{cases} + c_0 \quad \textbf{[I.3]}$$

For simplicity, we choose constants $c_1 = 1$ and $c_0 = 0$ (note that we have to choose c_1 positive since we wanted v to be monotone increasing). We already see that the space transform $v(\cdot)$ has very different properties in the two cases distinguished by the proposition: If $\mu \geq \sigma^2/2$ ($\tilde{\mu} \geq 1$), then the transformed process U_t has no lower bound, i.e. $\lim_{P\to 0} v(P) = -\infty$. If, on the other hand, $\mu < \sigma^2/2$ ($\tilde{\mu} < 1$), then the process U_t is lower-bounded and $\lim_{P \to 0} v(P) = 0$.[43] In the following, we will also need the inverse of the space transformation, which is given by

$$P = v^{-1}(u) = \begin{cases} \exp(u) & \text{if } \tilde{\mu} = 1, \\ (1-\tilde{\mu})^{\frac{1}{1-\tilde{\mu}}} u^{\frac{1}{1-\tilde{\mu}}} & \text{if } \tilde{\mu} \neq 1. \end{cases} \quad \text{[I.4]}$$

Now, the variance of the process U_t can be obtained from [I.2] as

$$[s^u(U_t)]^2 = [s^u(v(P_t))]^2 = [v'(P_t)\sigma]^2 \, P_t$$

Kolmogorov Backward Equations: We will now study expected exit times and exit probabilities of the process P_t from intervals (a, b), where $a \in [0, b)$. In our argument, we will first choose a positive lower bound, $a \in (0, b)$, since this yields a well-behaved problem, and then take the limit $a \downarrow 0$ to learn about the process's behavior at the boundary. Note here that the exit times and probabilities of U_t from $(v(a), v(b))$ are one-for-one related with the exit times and probabilities of P_t from (a, b), since $v(\cdot)$ is monotone increasing. Let us now consider functions $F(\cdot)$ defined by integrals of the form

$$F(U_0) = \mathbb{E}_0 \int_0^{T_{(v(a),v(b))}} f(U_t)\mathrm{d}t + \mathbb{E}_0[g(U_{T(v(a),v(b))})], \quad \text{[I.5]}$$

where $T_{(v(a),v(b))}$ is the first exit time of U_t from the interval $(v(a), v(b))$ – which is a random variable – and where $f(U_t)$ and $g(U_t)$ are arbitrary functions. Note now that setting $f(u) = 1$ and $g(u) = 0$, $F(U_0)$ is the expected exit time from the interval $(v(a), v(b))$ given that the process is started at U_0. On the other hand, when setting $f(u) = 0$, $g(v(a)) = 0$, and $g(v(b)) = 1$, $F(U_0)$ gives us the probability that U_t exits the interval $(v(a), v(b))$ at the upper end (again given that the process is started at U_0).

Since U_t is a diffusion, we can now derive the Kolmogorov Backward Equation for $F(\cdot)$ from [I.5]. For any interior $U_t \in (v(a), v(b))$ and $\mathrm{d}t$ small, we have $F(U_t) = f(U_t)\mathrm{d}t + \mathbb{E}_t\big[F(U_{t+\mathrm{d}t})\big]$. By the Ito Formula, we get $\mathbb{E}_t\big[F(U_{t+\mathrm{d}t})\big] = F(U_t) + \frac{1}{2}F''(U_t)[s^u(U_t)]^2\mathrm{d}t$ (note that the term in F' is zero since U_t is driftless). Thus we obtain the following second-order ODE for F that has to hold for all $u \in (v(a), v(b)$:

$$F''(u) = -\frac{2f(u)}{[s^u(u)]^2}. \quad \text{[I.6]}$$

Exit probabilities: For the case of the exit probability, we are almost done. Since we set $f(U) = 0$ in this case, Eq. [I.6] tells us that we need to find a function $F(\cdot)$ with zero second derivative everywhere – a line – that links the boundary points $F(v(a)) = 0$ and $F(v(b)) = 1$. At least for $a > 0$, this is a well-behaved problem and the solution for the exit probability at the top is given

$$P(U_t \text{ exits } (v(a), v(b)) \text{ at } v(b) | U_0 = v(P_0)) = \frac{v(P_0) - v(a)}{v(b) - v(a)} \quad \text{[I.7]}$$

$$= P(P_t \text{ exits } (a, b) \text{ at } b | P_0) \qquad \forall P_0 \in [a, b], \forall a \in (0, b), \quad \text{[I.8]}$$

Now, taking the limit as $a \to 0$ gives us the statements about the exit probability in the proposition. Consider first the case $\tilde{\mu} \geq 1$: Since, for fixed U_0, we certainly have $P(P_t \text{ exits } (0,b) \text{ at } b) \geq P(P_t \text{ exits } (a,b) \text{ at } b)$ – any path that exits (a,b) above certainly also exits $(0,b)$ above–, and since the probability of exiting (a,b) above approaches one as $a \to 0$, we must conclude that the probability of exiting $(0,b)$ at the top must be one, from which it follows that the probability of exiting at the bottom zero. Thus $P = 0$ is not accessible if $\tilde{\mu} \geq 1$, as is claimed in Point 1 of the proposition. If $\tilde{\mu} < 1$, however, as $v(a) \to 0$ the probability of exiting at the top approaches $v(P_0)/v(b)$ and the probability of exiting at the bottom approaches $1 - v(P_0)/v(b)$, which yields the claims in Point 2.(ii) of the proposition.

Expected local time and Green Functions. We now go back to the ODE **[I.6]** for the – more interesting – case where $f(U) \neq 0$. We will, however, set the boundary conditions zero from now on: $F(a) = F(b) = 0$. Note that we are *not* setting $f(U) = 1$ for now since we still can glean some interesting insights from maintaining a general $f(\cdot)$: We will be able to infer how much time the process spends in different areas of the state space. Since $f(\cdot)$ and $s^u(\cdot)$ are known functions, the ODE **[I.6]** is a standard problem that can be solved by Green-Function techniques. Define the Green Function for the interval $[a,b]$ as

$$G_{a,b}(X_0, x) = \begin{cases} \frac{(X_0-a)(b-x)}{b-a} & \text{if } X_0 \leq x, \\ \frac{(x-a)(b-X_0)}{b-a} & \text{if } X_0 > x. \end{cases} \qquad \textbf{[I.9]}$$

Intuitively, the Green Functions are basis ”hat” functions in x that have zero second derivative everywhere except at the kink point X_0.[44] The solution to the ODE **[I.6]** is then given by

$$F(U_0) = \int_{v(a)}^{v(b)} f(u) \overbrace{G_{v(a),v(b)}(U_0, u) \underbrace{\frac{2}{[s^u(u)]^2} \mathrm{d}u}_{\text{speed measure}}}^{\equiv L(u|U_0)\text{: expected local time at } u \text{ given } U_0} \qquad \textbf{[I.10]}$$

Note how this solution combines the Green basis functions and weighs them by the size of the second derivative specified in the ODE **[I.6]** for each point. Now it pays off that we have maintained an arbitrary function $f(\cdot)$: By setting indicator functions $f(u) = \mathbb{I}(u \in C)$ for subsets $C \in (v(a), v(b))$ of the state space, $F(U_0)$ becomes the expected time that U_t will spend in the set C before exiting the interval. Thus, the function $L(u|U_0)$ is a ”time density”: It tells us the expected time U_t will spend around u (given initial condition U_0). $L(u|U_0)$ consists of two components: First, the Green Function $G_{v(a),v(b)}(U_0, u)$, which is expected local time at u in the case that U_0 is a standard Wiener Process (with $s^u = 1$). The second component is the so-called *speed measure* $2/[s^u(u)]^2 \mathrm{d}u$ that tells us how quickly the process leaves the neighborhood of u: If the process has high variance at u, it will leave the neighborhood of u quickly and thus spend less time there than standard Brownian Motion; U_t will tend to be trapped more in the neighborhood of u if the variance is low, however. We will now see that for high-enough drift, the process can be trapped for very long time close to zero, which can make the expected exit times grow unbounded.

Expected exit time. What is left to show is Point 2.(ii) in the proposition, i.e. that the expected exit time is finite and converges to zero as $P_0 \to 0$ in the case $\tilde{\mu} < 1$. We now set $f(u) = 1$ in Eq. [I.10] in order to find the expected exit time. To be specific, denote by $S_a(P_0)$ the expected time of exit of the process P_t from the interval (a, b) given that it is started at P_0. For convenience, we change the variable of integration from u back to P:

$$S_a(P_0) = \int_a^b \frac{2v'(P)}{\sigma^2 P v'(P)^2} G_{v(a),v(b)}(v(P_0), v(P))\mathrm{d}P = \int_a^b \frac{2G_{v(a),v(b)}(v(P_0), v(P))}{\sigma^2 P^{1-\tilde{\mu}}}\mathrm{d}P, \quad \text{[I.11]}$$

where we note that the change of variable required us to set $\mathrm{d}u = v'(P)\mathrm{d}P$ in the first step. Now, as we let $a \to 0$, $\tilde{\mu} < 1$ implies $v(0) = 0$ and thus $S_a(P_0)$ converges to (use Eq. [I.3] and [I.9] to show this):

$$S_0(P_0) = \frac{2(b^{1-\tilde{\mu}} - P_0^{1-\tilde{\mu}})}{b^{1-\tilde{\mu}}\sigma^2} \int_0^{P_0} 1\mathrm{d}P + \frac{2P_0^{1-\tilde{\mu}}}{b^{1-\tilde{\mu}}\sigma^2} \int_{P_0}^b b^{1-\tilde{\mu}} - P^{1-\tilde{\mu}}\mathrm{d}P. \quad \text{[I.12]}$$

The first integral evaluates to P_0 and the second integral is clearly bounded since the integrand can be bounded on the range of integration, thus $S_0(P_0)$ is finite for any $P_0 \in (0, b)$. Also, both terms in Eq. [I.12] converge to zero as $P_0 \to 0$, which concludes the proof.[45]

■

Now that the heavy lifting is done, we can use the above lemma to characterize the boundary behavior for an arbitrary drift by approximating it close to the boundaries:

Lemma 4. *The boundary $P = 0$ is called* accessible *if it can be reached in finite time with positive probability.*

1. *If $\lim_{P\to 0}[T(P) - C(P)] \geq \frac{\sigma^2}{2}$, then $P = 0$ is inaccessible, i.e. the boundary is never reached.*

2. *If $\lim_{P\to 0}[T(P) - C(P)] < \frac{\sigma^2}{2}$, then $P = 0$ is accessible and is reached immediately for small P, i.e. $P($ exit $(0, b)$ at $0|P_0) \to 0$ as $P_0 \to 0$, and $\mathbb{E}_0[T_{(0,b)}|P_0] \to 0$ as $P_0 \to 0$.*

Proof of Lemma 4

By the law of motion for P_t, Eq. [3.4], the drift of P_t converges to $\mu \equiv \lim_{P\to 0}[T(P) - C(P)]$ as $P \to 0$. Again by Eq. [3.4], the variance of P_t is $\sigma^2 P(1 - P) = \sigma^2(P - P^2)$; keeping only the highest order term this approaches

$\sigma^2 P$ as $P \to 0$. For small ϵ, on the interval $(0, \epsilon)$ the process P_t is thus well approximated by $\mathrm{d}P_t = \mu \mathrm{d}t + \sigma\sqrt{P_t}\mathrm{d}W_t$. The results in the proposition then follow from Lemma 3. ■

J Exit times and probabilities

This section shows how we compute the expected time to bankruptcy of one of the regions and the probability that Region 1 is the one that ends up broke[46], always considering the equilibrium in which this eventually occurs.

First, consider the probability that Region 1 ends up broke, i.e. that stochastic process ends up at $P = 0$. Denote this probability by $Q(P)$. If both boundaries $P = 0, 1$ are accessible, then we have $\lim_{P\to 0} Q(P) = 1$ and $\lim_{P\to 1} Q(P) = 0$. For interior $P_t \in (0, 1)$,over a small-enough time interval $\mathrm{d}t$ the probability satisfies the martingale property

$$Q(P_t) = \mathbb{E}_t\left[Q(P_{t+\mathrm{d}t})\right]. \tag{J.1}$$

We can approximate this *Backward Equation* over a discrete grid $P_{i=1}^N$ as:

$$Q(P_i) = \Delta t\left[w_u Q(P_{i+1}) + w_d Q(P_{i-1}) + (1 - w_u - w_d)Q(P_i) +\right] \quad \forall i \in \{2, \ldots, N-1\}$$
$$Q(P_1) = 1, Q(P_N) = 0 \tag{J.2}$$

where w_d and w_u are the probabilities that the process goes up (or down) one position in the grid over the interval Δt. [J.2] is sparse linear system of N equation with N unknowns that can be solved efficiently.

The right panel of Fig. 9 plots the probability that Region 1 ends up broke starting from three different points ($P = 1/2, 1/3, 1/10$) as we vary μ. As expected, this probability is higher when i) Region 1 is less liked by the center or ii) when the starting point is closer to $P = 0$.

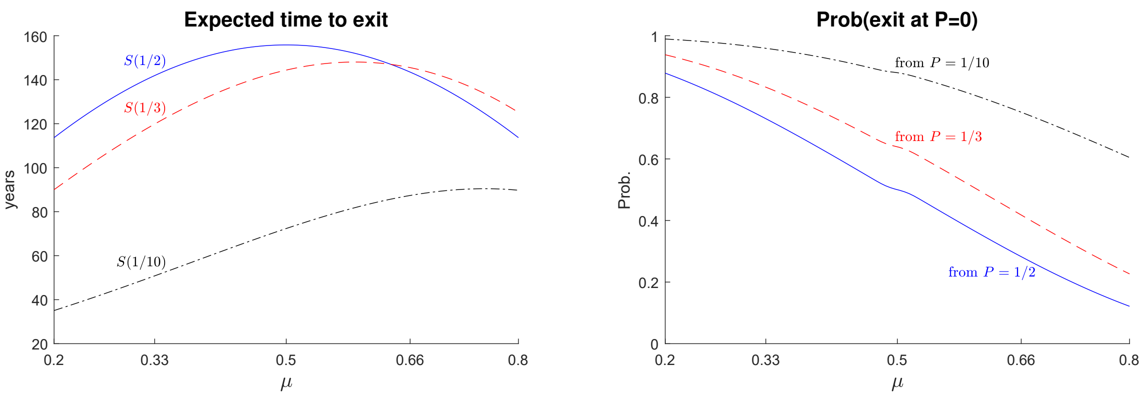

Figure 9: $S(P)$ and $Q(P)$ as a function of μ. Model parameters as in Fig. 2

[46]See Appendix I for a formal characterization of these variables.

Similarly, the expected time to bankruptcy from an interior point $P_t \in (0,1)$ has to satisfy the backward equation

$$S(P_t) = \mathrm{d}t + \mathbb{E}_t\left[S(P_{t+\mathrm{d}t})\right], \tag{J.3}$$

the boundary conditions being $\lim_{P\to 0} S(P) = \lim_{P\to 0} S(P) = 0$. The numerical approximation over a grid is analogous to [J.2]. The left panel of Figure 9 reveals that the expected time to exit from $P = 1/2$ is maximal when the two regions are equally liked.

Figure 10 shows the expected time to exit (left panel) and the probability of exiting at $P = 0$ (i.e. the probability of region 1 ending in bankruptcy) as a function of the volatility σ. On the later, we can see that, as expected, the probability does not change when the starting point is $P = 1/2$. However, as we start closer to $P = 0$ the probability of bankruptcy of region 1 is slightly lower for low values of σ, as it is easier for the center to balance wealth. Regarding the expected time, it decreases with the volatility and as we get closer to the boundary.

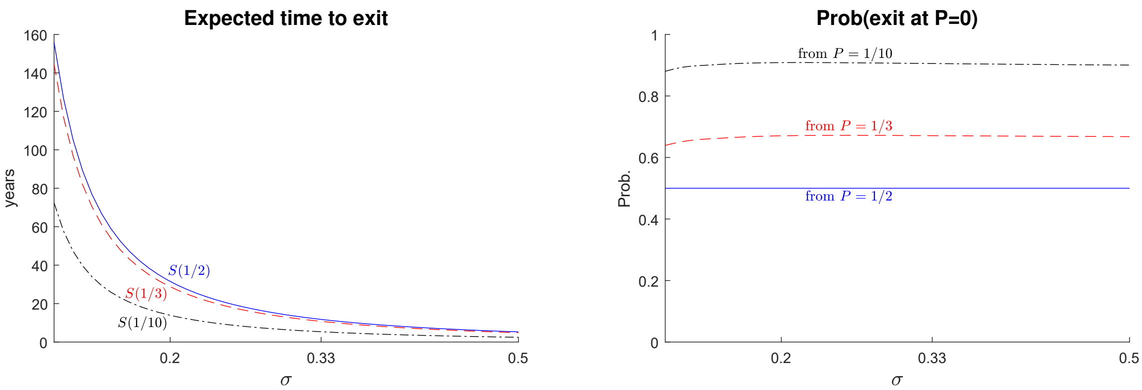

Figure 10: $S(P)$ and $Q(P)$ as a function of σ . Model parameters as in Fig. 2

K Costless redistribution

In this section we remove our assumption on costly transfers and replace it by:

Assumption 2 (Frictionless redistribution)**.** $X(\tilde{\tau}) = \tilde{\tau}$.

In this frictionless setting we find two novel types of equilibria. In the first, the center immediately expropriates the less liked region. The second type, which we call *smooth equilibria*, only exists in the knife-edge case when the two regions are equally liked by the center and features consumption policies that are increasing in the wealth share.

K.1 Expropriation equilibrium

First, notice that under Ass. 2, the center may potentially use mass transfers. In the reduced state space, a mass transfer induces a jump in the trajectory of P:

$$\lim_{\mathrm{d}t\to 0} P_{t+\mathrm{d}t} = P_t - T_m \tag{K.1}$$

For clarity, we will treat mass transfers (T_m) and flow transfers (T_f) as different mathematical objects.[47] Doing this we follow Barczyk and Kredler (2014a); we also use their extended equilibrium concept with mass transfers. We will see that, under certain conditions, the center wants to transfer all wealth to one region instantaneously, which we call *expropriations*:

Definition 4 (Expropriation)**.** *An* expropriation *policy are transfer rules that immediately remove all wealth from on Region, i.e. $T_m(P) = P$ for expropriation of Region 1, or $T'_m(P) = (1 - P)$ for expropriation of Region 2.*

In the following proposition we show that, in the case of frictionless transfers, the center immediately expropriates the less liked region and *spoon-feeds* it thereafter. That is, the center implements flow transfers that provide the less-liked region by the amount preferred by center, which is part of the center's preferred efficient allocations, see Proposition 1. However, the better-liked region overconsumes according to the wealth-pooling allocation (see Proposition 3), thus the resulting equilibrium is inefficient:

Proposition 6. *If $\mu \geq 1/2$, an equilibrium exists in which the center expropriates the less preferred Region 2 (setting $T_m = -(1 - P)$), implementing the allocation $C = \rho$ and $C' = (1 - \mu)\rho$.*

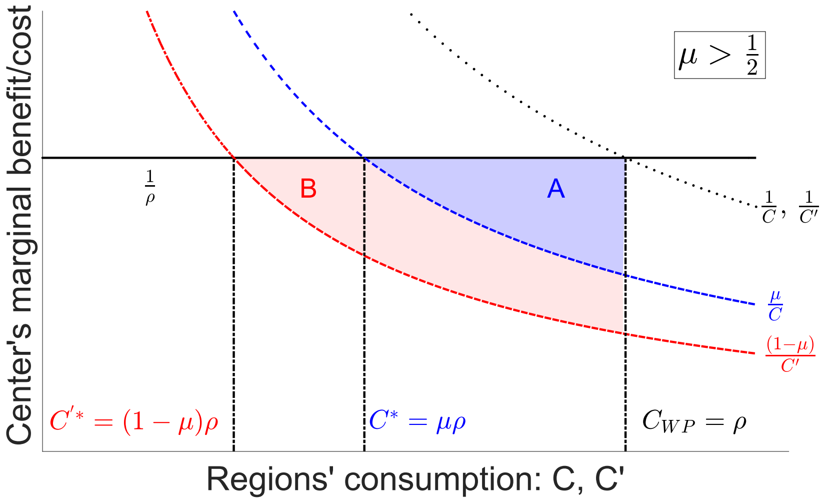

Figure 11: Expropriation equilibrium. Dashed blue line (red) represents the marginal benefit (for the center) of one extra unit of transfer to Region 1 (Region 2) if the center could control its consumption. Dotted black line represents the marginal utility of one extra unit of consumption for any region if it is not constrained.

The intuition of the proof is as follows. First, notice that in the expropriation equilibrium, all players' value functions are independent of P, since the center will intermediately make a mass transfer to bring the economy to a desired level of distribution $P - T_m$.

[47] It can be proven that, under Costly Redistribution (Assumption 1), mass transfers are ineffectual, where by ineffectual we mean that mass tax to one region will generate no revenue.

Now, if this level was in the interior, no region would be constrained and they will consume as in the *wealth pooling* scenario, resulting in a welfare loss for the center equaling the area 2A+B in Fig. 11 (with respect to the center's ideal allocation). If, however, the center expropriates Region 2 (setting $P - T_m = 1$), then Region 1 optimally consumes $C = C^{wp} = \rho$. However, Region 2 being broke can consume only the flow transfer that the center provides, which the center optimally sets to $C' = (1-\mu)\rho$, i.e. smaller than what Region 2 would prefer. Expropriating Region 2 thus yields a welfare loss equaling A in the figure. Similarly, we obtain the allocations $C = \mu\rho$ and $C' = C^{wp} = \rho$ if the center expropriates Region 1, leading to a welfare loss of A+B for the center. Thus, the center's minimal welfare loss is A and thus it expropriates the less-liked Region 2.

Formal proof of Proposition 6

For clarity, assume that $\sigma = 0$. For the case $\sigma > 0$, the proof is identical, since the extra terms that appear in the HJB's vanish. Given transfers $T_f(\cdot), T_m(\cdot)$ and the other region consumption $C'(\cdot)$, Region 1 solves:

$$V(P_t) = \max_{C \in \mathbb{C}(P,T_m)} \quad \ln(C(\cdot))\mathrm{d}t - \left[C(\cdot) + C'(\cdot) + T_f(P_t) + T_f'(P_t)\right]\frac{\mathrm{d}t}{\rho} + e^{-\rho \mathrm{d}t} V(P_{t+\mathrm{d}t})$$
$$\text{s.t} \quad P_{t+\mathrm{d}t} = P_t + \left[-(1-P)(T_f(P_t) + C) + P(T_f'(P_t) + C')\right]\mathrm{d}t - T_m$$

where $\mathbb{C}(P, T_m) = [0, -T_f(0)]$ if $P - T_m = 0$ and $[0, \infty)$ otherwise. When we substitute the linear approximation $e^{-\rho\Delta t}V(P_{t+\Delta t}) \approx (1-\rho \mathrm{d}t)V(P_t - T_m) + V_P(P_t - T_m)(\mathrm{d}P_t) + o(\Delta t)$, differently to the case of only flow transfers, $(\mathrm{d}t)^0$ terms do not cancel out. Therefore, to find the optimal choices we have to look separately at $(\mathrm{d}t)^0$ and $(\mathrm{d}t)^1$ terms. Isolating $(\mathrm{d}t)^0$ terms, we find that the policies have to satisfy the *order-0 requirement*:

$$V(P) = V(P - T_m) \qquad \text{[K.2]}$$

which implies that the value function is flat in P. Terms up to order $(\mathrm{d}t)^1$ give us the *order-1 requirement*:

$$\rho V(P) = \max_C \quad \ln(C(\cdot))\mathrm{d}t - \left[C(\cdot) + C'(\cdot) + T_f(P) + T_f'(P)\right]\frac{1}{\rho}$$
$$+ \left[-(1-P)(T_f(P) + C) + P(T_f'(P) + C')\right] V_P(P - T_m)$$

Using the fact that $T_f = -T_f'$ (balanced budget) and that $V_p(P) = 0$ (by [K.2]):

$$\rho V(P) = \max_C \quad \ln(C(\cdot))\mathrm{d}t - \left[C(\cdot) + C'(\cdot)\right]\frac{1}{\rho}$$

If the region is not broke after the mass transfer T_m, the FOC [3.8] tell us that the region it will consume the *wealth pooling* allocation $C = \rho$. If the region is broke, however, it may be limited by the flow transfer $-T_f$.

$$C(P, T_f, T_m) = \begin{cases} \rho & \text{if } P - T_m > 0 \\ \min\{\rho, -T_f\} & \text{otherwise} \end{cases} \qquad \text{[K.3]}$$

Similarly for Region 2 (just change the signs of transfers and replace P by $1-P$).
We can proceed analogously with the center's problem. The *order-0 requirement* in this case is:

$$W(P) = \max_{T_m} \quad W(P - T_m) \tag{K.4}$$

which implies $W_P(P - T_m) = 0$. Terms up to order $(\mathrm{d}t)^1$ (*order-1 requirement*) determine T_f and T_m:

$$\rho W(P - T_m) = \max_{\substack{T_f = -T'_f \\ T_m = -T'_m}} \quad \Big[\mu \ln(C(C; P, T_f, T_m)) + (1-\mu)\ln\Big(C^{'}(C'; 1-P, -T_f, -T_m)\Big)\Big]$$
$$- \Big[C(\cdot) + C^{'}(\cdot) + T_f(P) + T'_f(P)\Big]\frac{1}{\rho} + W_P(P - T_m)\dot{P}$$

where $T_m = \arg\max_{\tilde{T}_m} W(P - \tilde{T}_m)$. Notice last term vanish by [K.4].

Looking at the consumption policies of the regions [K.3] we can see that there are three scenarios depending on the mass transfer:

1. $P - T_m \in (0,1)$. No region is constrained and both consume the wealth pooling allocation. T_f, T'_f are undetermined. The payoff of the center is: $\rho W = \ln(\rho) - 2$.

2. $T_m = -(1-P)$. Region 1 (the preferred by the center) takes all the wealth and Region 2 is broke and consumes $C^{'} = -T'_f = (1-\mu)\rho$, the preferred allocation by the center. The payoff of the center is: $\rho W = \ln(\rho) + (1-\mu)\ln(1-\mu) - (2-\mu)$.

3. $T_m = P$. Region 2 takes all the wealth and Region 1 is broke and consumes $C = -T_f = \mu\rho$. The payoff of the center is: $\rho W = \ln(\rho) + \mu\ln(\mu) - (1+\mu)$.

Now, for $\mu > 1/2$ it is clear that ρW is higher in scenario 2. ■

K.2 Smooth equilibria

In this section we construct an equilibrium in which the center only uses flow-type (*smooth*) transfers. It turns out that this equilibrium only exist in the knife-edge case of equal welfare weighs ($\mu = 1/2$). In these *smooth equilibria*, regions follow a consumption policy increasing in P that coincides at the boundaries with the consumption allocations in the *expropriation equilibrium*. There is a continuum of transfer policies $T(P)$ that sustain this allocation, giving rise to multiplicity of equilibria. Different transfer schemes lead to different dynamics in P, different ergodic distributions and different value function for regions. The set of transfer policies spans between two extremes that resemble the expropriation equilibria, passing through a symmetric equilibrium in which regions avoiding bankruptcy.

K.2.1 Equilibrium candidates

We first derive three conditions for the consumption allocation in a smooth equilibrium: *center's indifference condition*, *best responding condition* and the *frontier condition(s)*.

Center's indifference condition First, for the center to use flow-type transfers (and no mass transfer at any P), the center has to be indifferent between all P, implying that $W(P) = const$ for all P.[48] Since $W_P = 0$ and $W_{PP} = 0$, the center's HJB, Eq. [4.2], then implies

$$\rho\bar{W} = \mu\ln(C(P)) + (1-\mu)\ln(C'(P)) - \left[C(P) + C'(P)\right]\frac{1}{\rho}. \qquad \text{[K.5]}$$

This, in turn, means that regions' consumption policies have to trace out an indifference curve in the planner's utility, see Fig. 12.

Best responding condition Furthermore, the planner's value is a weighted sum of the regions' values,

$$W(P) = \mu V(P) + (1-\mu)V'(P).$$

Differentiating with respect to P, using Regions's FOCs [3.8] and the fact that $W_P(P) = 0$ gives us the restriction:

$$\underbrace{\frac{1}{C'(P)} - \frac{1}{\rho}}_{\equiv M'(P)} = \frac{P}{1-P}\frac{\mu}{1-\mu}\underbrace{\left(\frac{1}{C(P)} - \frac{1}{\rho}\right)}_{\equiv M(P)} \qquad \text{for all } P \in [0,1]. \qquad \text{[K.6]}$$

Threfore, marginal utilities are related one-to-one (when fixing a value of P): If Region 1 perceives a high value of owning more of the cake, then also Region 2 must perceive a high value of owning more, and both will choose lower consumption (and thus higher marginal utility).

Conditions at the frontier Now, consider the center's problem at $P = 0$. From Region 2's FOC, we know that $C'(0) = \rho = C^{wp}$. Since Region 1 is broke, its consumption will be the minimum of the solution of its FOC [3.8] C_{unc} and the flow transfer $-T_f(0)$ provided by the center. Noticing that $W_P(0) = 0$, the problem of the center is:

$$\max_{-T_f(0)\geq 0}\left\{\mu\ln(\min\{C_{unc}(P), -T_f(0)\}) + (1-\mu)\ln(C^{wp}) - \left[\min\{C_{unc}(P), -T_f(0)\} + C^{wp}\right]\frac{1}{\rho}\right\},$$

This criterion is concave in $-T_f(0)$ on the part where $-T_f(0) \leq C_{unc}$ and it is invariant in $-T_f(0)$ on the range $-T_f(0) \geq C_{unc}$. If $C_{unc} \geq \mu\rho$, the center can set Region 1's consumption to the desired level and set $-T_f(0) = \mu\rho$. If Region 1's plan is to consume less than $\mu\rho$, then the center should set $-T_f(0) = C_{unc}0$ since the criterion is decreasing

[48] Formally, this result arises from the *order-0 requirement* of the problem of the center when mass transfers are allowed; see Eq. [K.4] in the *expropriation equilibria* proof.

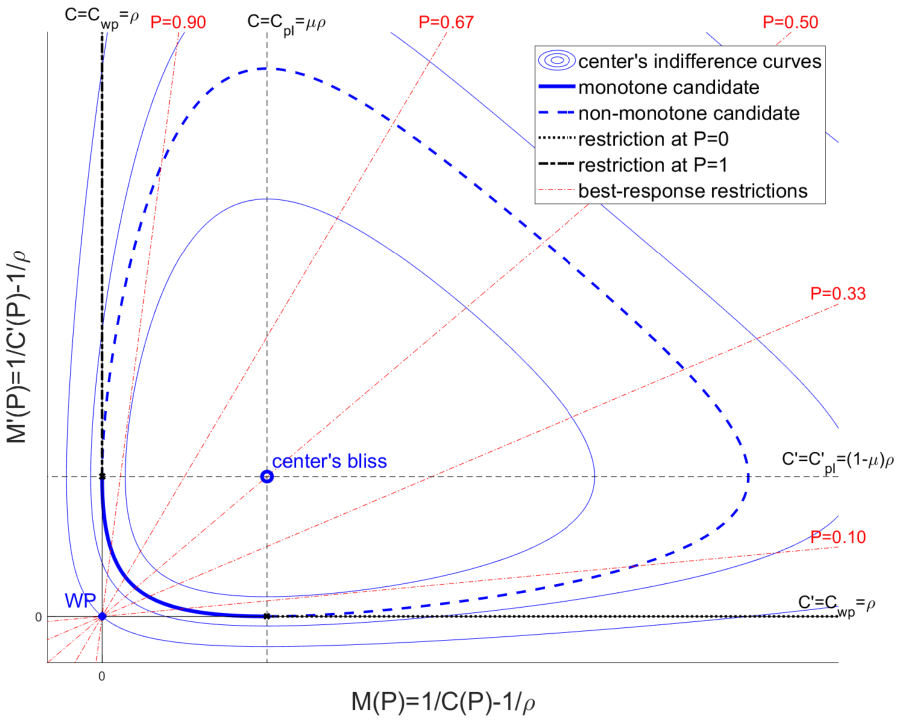

Figure 12: Smooth equilibria

Smooth equilibrium candidates for $\kappa = 0$, $\mu = 0.5$, $\sigma = 0.02$, $\rho = 0.04$. The origin represents the *wealth pooling* allocation. Dashed/pointed black lines are the *efficient allocations* that cross at the center's bliss point. Solid blue lines trace out the indifference curves of the center (*center's indifference condition*), bold dashed vertical lines the *frontier conditions* and red dashed diagonal lines the *best response conditions* for different values of P.

in $-T$. Using the analogous reasoning for Region 2 in the case $P = 1$, we conclude that in equilibrium it must be that the following *frontier conditions* are satisfied:

$$C(0) \leq \mu\rho, \qquad \Longleftrightarrow \qquad M(0) \geq \frac{\mu - 1}{\rho\mu} \qquad \text{[K.7]}$$

$$C'(1) \leq (1-\mu)\rho \qquad \Longleftrightarrow \qquad M'(1) \geq \frac{\mu}{\rho(1-\mu)} \qquad \text{[K.8]}$$

We note that all conditions that we have found so far are independent of the strength of shocks, σ.

These conditions are graphically represented on the marginal utility (M, M') space in Figure 12 for the case of identical regions ($\mu = 0.5$). Equilibrium candidates in the figure trace out the center's indifference curves (*center's indifference condition*), connect the vertical black dash-dotted line and the horizontal black pointed line that pass through the WP allocation (*frontier conditions*) and pass through regions where red (dash-dot) lines exist (region's are best responding).

We see that there is exactly one equilibrium candidate to the lower left of the center's bliss point. This equilibrium candidate is such that regions' consumption functions are

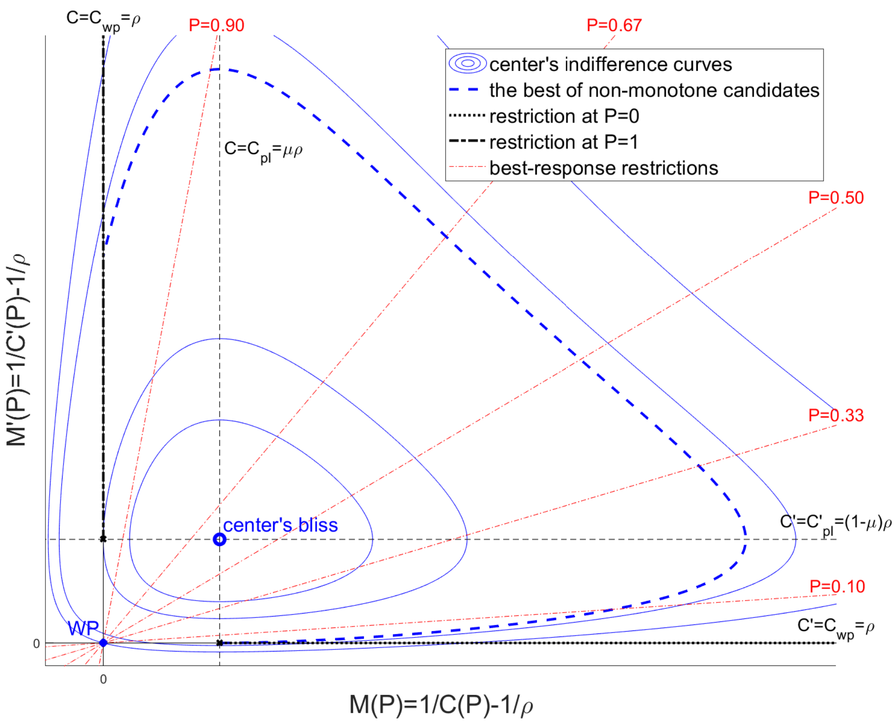

Figure 13: Smooth equilibrium candidates for $\mu = 0.6$. See legend of Figure 12

increasing in their wealth share (M increases in P and thus C decreases in P, whereas M' decreases in P). We call this equilibrium candidate the *monotone* one. Notice that for this candidate,

$$C(0) = C'(1) = \frac{\rho}{2} \quad ; \quad C(P), C'(P) \in \left[\frac{\rho}{2}, \rho\right] \forall P \in (0,1) \qquad \textbf{[K.9]}$$

However, there is also a continuum of equilibrium candidates that pass above and to the right of the center's bliss point, out of which one is marked with a dashed line. We call these candidates the *non-monotone* ones, since consumption functions are not monotone in wealth shares for these candidates. For low values of P, M is increasing in P (the area to the lower right of the center's bliss point), and thus C becomes decreasing in P. However, for higher values of P, $C(P)$ is increasing. Similarly for Region 2. Finally, note that no equilibrium candidates can be constructed that go through the north-west and south-east quadrant since none of the red dashed diagonal lines (*best response conditions*) passes through these.

For the *asymmetric* case, in which the center favors one region over the other, the monotone candidate equilibrium disappears, since it is impossible to connect the black lines that represent the center's best-response restrictions at $P \in \{0, 1\}$ by an indifference curve without leaving the north-east quadrant, meaning that the monotone equilibrium is a knife-edge result: it can only exist for $\mu = 1/2$. The continuum of non-monotone candidates to the top right of the center's bliss point persist, however, See Figure 13

K.2.2 Existence of Smooth equilibria

Notice that, so far, we have not derived any restrictions on the transfer function $T(P)$. However, the slope of the transfer function has to give regions the right incentives to save: regions' Euler Equations (EE) imply restrictions in the form of an ordinary differential equation for $T(\cdot)$. Region 1's EE is obtained differentiating its HJB [3.7] with respect to P and using the FOC [3.8]:

$$0 = \underbrace{-\dot{P}V_{PP} - \sigma^2 P(1-P)V_{PPP}}_{=-\mathcal{A}V_P} - \sigma^2(1-2P)V_{PP} + \frac{1}{\rho}C'_P + V_P\left[\rho - C - C' + T_P - PC'_P\right] \quad \text{[K.10]}$$

where we have suppressed the function arguments and their derivatives for readability. It turns out that we do not need the EE of Region 2, since it can be shown to be implied by [K.10].

Using the law of motion $\dot{P}$, [K.10] then implies that the transfer function $T(\cdot)$ obeys

$$T_P(P) = -\gamma(P)T(P) + F(P). \quad \text{[K.11]}$$

This is a first-order ODE for T with decay rate $\gamma(P) \equiv V_{PP}/V_P$ and forcing variable $F(P)$.[49] The solution of this ODE can be written as (with $T(0.5)$ being a constant):

$$T(P) = \underbrace{\exp\left(-\int_{0.5}^{P}\gamma(y)\mathrm{d}y\right)}_{\equiv\Gamma(P)} T(0.5) + \underbrace{\int_{0.5}^{P}\exp\left(-\int_{x}^{P}\gamma(y)\mathrm{d}y\right)F(x)\mathrm{d}x}_{\equiv T_{symm}(P)}, \quad \text{[K.12]}$$

i.e. T is a sum over the forcing variable F, which is compounded by γ. The following proposition summarizes the results on the *monotone equilibria*.

Proposition 7. *Suppose that $\mu = 1/2$ and that transfers are frictionless. If $T_{symm}(1) \geq \rho/2$, then there exist a continuum of smooth monotone equilibria. Allocations $\{C(P), C'(P)\}$ are given by* [K.5]*,* [K.6] *and* [K.9]*. Transfer schemes $\{T(P), T'(P)\}$ satisfy* [K.12] *with boundary conditions $T(0.5) \in [T_{min}(0.5), T_{max}(0.5)]$, where:*

$$T_{min}(0.5) = \frac{T_{symm}(1) - \frac{\rho}{2}}{\Gamma(1)} \quad ; \quad T_{max}(0.5) = -T_{min}(0.5). \quad \text{[K.13]}$$

K.2.3 Multiplicity of equilibria: Transfer schemes

Fig. [14] depicts the set of smooth equilibria under a reasonable parameterization. The bottom left panel shows that the consumption allocation is unique and increasing in the share of wealth, as described before. At the boundaries, consumption approaches the expropriation allocation. The depletion of total resources (bottom right panel) in the interior is slightly faster than in the expropriation equilibria. This is expected, since the center cannot control the regions when they are not constrained.

[49] Where $F = \left[PC' - (1-P)C - \sigma^2(2P-1)\right]\frac{V_{PP}}{V_P} + C + C' - \rho + C'_P[P - \frac{1}{\rho V_P}] + \sigma^2 P(1-P)\frac{V_{PPP}}{V_P}$.

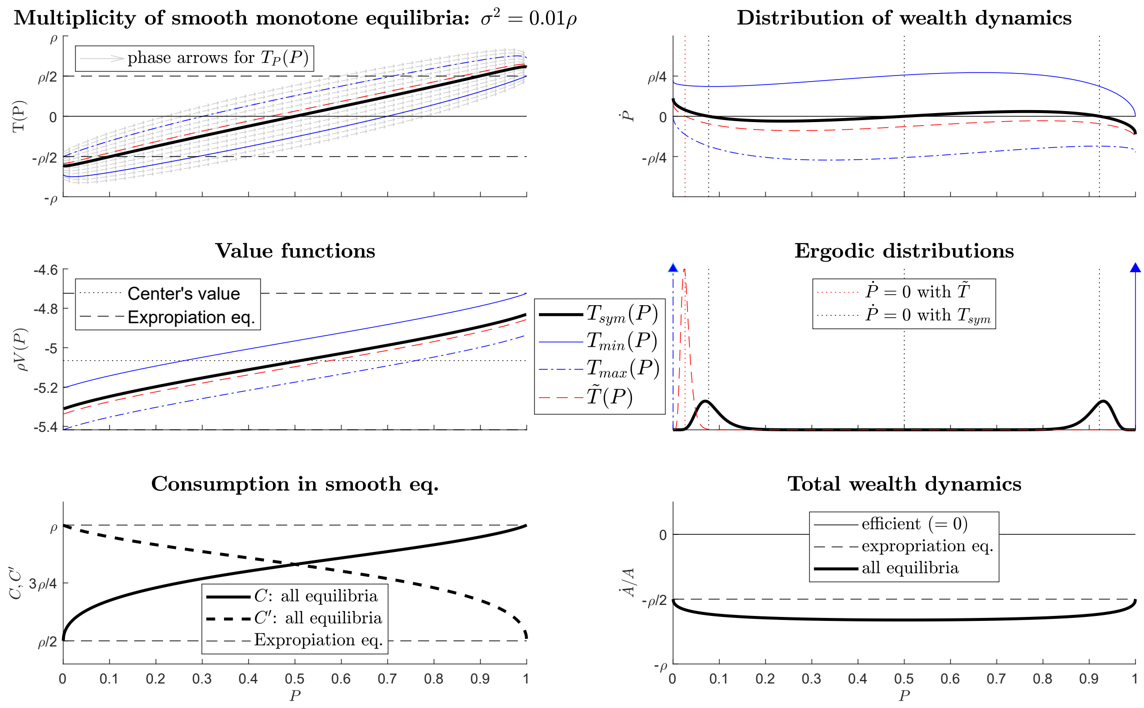

Figure 14: Smooth monotone equilibria for $\kappa = 0$, $\mu = 0.5$, $\sigma = 0.02$, $\rho = 0.04$.

As for transfer schemes $T(P)$ (see upper left panel), they appear in a band around $T_{symm}(P)$ (black line) and delimited by T_{min} and T_{max} (blue lines), which denote the least and most favorable tax scheme for Region 1 that constitute equilibria. By construction, $T_{symm}(0)$ is zero at $P = 1/2$ and, crucially, is more generous to the broke region at the boundaries than the *expropriation equilibrium*. The law of motion for P (upper right) induces one unstable steady state at $P = 1/2$ and two stable ones close to the boundaries. [50]

The ergodic distribution generated by this law of motion (middle right panel) is centered these two stable steady states. When started at $P = 1/2$, luck determines which region turns poorer, the center not taxing aggressively enough to balance the wealth distribution. Once one region turns poor enough, savings incentives for the poor region are strong enough to avoid bankruptcy, which gives the poor region a higher value than under expropriation (middle left). Interestingly, at $P = 1/2$ both regions are indifferent between the symmetric monotone smooth equilibrium and a fair lottery of the two expropriation equilibria.

When the tax schedule is tilted against Region 1 ($\tilde{T}$, dashed red lines) the more favorable steady state for Region 1 vanishes. Region 1 eventually spends most of the time close to bankruptcy, enjoying a lower value. The maximum tax on Region 1 that can be imposed (dashed blue line) yields the expropriation allocation once Region 1 turns broke, which occurs probability 1 in the long run (i.e. the ergodic distribution is a mass point at $P = 0$). In this sense, the smooth equilibria span a continuum between the two

[50] As σ grows, the stable steady states move closer to $P = 1/2$.

expropriation equilibria (which both exist for $\mu = 1/2$).

Non-monotone candidates Up to now, we have focused on *monotone candidates*. Now we turn our attention to candidates that imply consumption policies that are non-monotonous in P – an empirically unappealing property. Still, we check numerically if such candidates can constitute equilibria. To do so, we pick a consumption allocation for a non-monotone candidate; start with the minimum transfer possible at the boundary, $T_{min}(1) = -C'(1)$; and solve the ODE [K.11] backwards. If $T_{min}(0) \geq -C(0)$, we have found an equilibrium, otherwise the guess is ruled out. We do not find any $T(P)$ that sustain any monotone equilibrium in our simulations, considering a wide range of guesses.

References

ÁBRAHÁM, ÁRPÁD AND SAROLTA LACZÓ (2018): "Efficient risk sharing with limited commitment and storage," *The Review of Economic Studies*, 85 (3), 1389–1424.

AGUIAR, MARK, MANUEL AMADOR, EMMANUEL FARHI, AND GITA GOPINATH (2015): "Coordination and crisis in monetary unions," *The Quarterly Journal of Economics*, 130 (4), 1727–1779.

ATTANASIO, ORAZIO P. AND NICOLA PAVONI (2011): "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption," *Econometrica*, 79 (4), 1027–1068.

BARCZYK, DANIEL AND MATTHIAS KREDLER (2014a): "A Dynamic Model of Altruistically-Motivated Transfers," *Review of Economic Dynamics*, 17 (2), 303–328.

——— (2014b): "Altruistically motivated transfers under uncertainty," *Quantitative Economics*, 5 (3), 705–749.

BATTAGLINI, MARCO AND STEPHEN COATE (2007): "Inefficiency in legislative policy-making: a dynamic analysis," *American Economic Review*, 97 (1), 118–149.

——— (2008): "A dynamic theory of public spending, taxation, and debt," *American Economic Review*, 98 (1), 201–36.

BOZ, EMINE (2011): "Sovereign default, private sector creditors, and the IFIs," *Journal of International Economics*, 83 (1), 70–82.

BREIMAN, LEO (1992): *Probability*, SIAM.

CHARI, VARADARAJAN V AND PATRICK J KEHOE (2007): "On the need for fiscal constraints in a monetary union," *Journal of Monetary Economics*, 54 (8), 2399–2408.

——— (2008): "Time Inconsistency and Free-Riding in a Monetary Union," *Journal of Money, Credit and Banking*, 40 (7), 1329–1356.

FARHI, EMMANUEL AND IVÁN WERNING (2017): "Fiscal unions," *American Economic Review*, 107 (12), 3788–3834.

FINK, FABIAN AND ALMUTH SCHOLL (2016): "A quantitative model of sovereign debt, bailouts and conditionality," *Journal of International Economics*, 98.

HELLAND, INGE (1996): "One-dimensional diffusion processes and their boundaries," *Preprint series. Statistical Research Report http://urn. nb. no/URN: NBN: no-23420*.

KOCHERLAKOTA, NARAYANA R. (1996): "Implications of Efficient Risk Sharing without Commitment," *The Review of Economic Studies*, 63 (4), 595–609.

PERSSON, TORSTEN AND GUIDO TABELLINI (2002): "Chapter 24 - Political Economics and Public Finance," Elsevier, vol. 3 of *Handbook of Public Economics*, 1549–1659.

ROCH, FRANCISCO AND HARALD UHLIG (2018): "The dynamics of sovereign debt crises and bailouts," *Journal of International Economics*, 114.

SANNIKOV, YULIY (2007): "Games with imperfectly observable actions in continuous time," *Econometrica*, 75 (5), 1285–1329.

SIMON, LEO K AND MAXWELL B STINCHCOMBE (1989): "Extensive form games in continuous time: Pure strategies," *Econometrica: Journal of the Econometric Society*, 1171–1214.

THOMAS, JONATHAN AND TIM WORRALL (1988): "Self-Enforcing Wage Contracts," *The Review of Economic Studies*, 55 (4), 541–553.