Fiscal Rules and Transfers in a Union*

Chiara Felli†

- Job Market Paper -

Please click here for the latest version

April 27, 2021

Abstract

I study dynamic optimal fiscal rules in a supranational setting in which national governments with quasi-hyperbolic preferences are subject to privately observed idiosyncratic shocks. In this context, fiscal rules aim at striking a balance between flexibility to react to shocks, and commitment to avoid excessive government spending. I compare optimal rules in two different environments: one in which the supranational authority is allowed to transfer resources across countries (i.e. a fiscal union) and one in which transfers are forbidden. I find that optimal fiscal rules can be implemented as deficit limits and are complemented with a combination of grants and loans in a fiscal union. All instruments are debt-contingent: higher public debt contemporaneously tightens deficit limits and reduces the entity of both transfers and credits. Welfare gains from setting up a transfer system are positive, but vanish in the limit case in which governments only care about their own consumption. I present a sample calibration of the model using EU data. Optimal deficit limits are not far from Maastricht 3%; member countries under extreme distress receive help in the form of grants and loans; grants account for 30% of the overall financial help and are at most 4.5% of GDP.

Advanced economies' debt has been on a continuously increasing trend since the early 1980s, and a growing body of empirical evidence shows that larger deficits are associated with countries having short-lived governments, more ideological polarization or political fragmentation and with a proportional (rather than majoritarian) electoral system.[1] This evidence hints at the political nature of the bias behind advanced economies' sky-rocketing liabilities. Governments and international institutions have tried to limit sovereign

*I am deeply grateful to Facundo Piguillem, Pierpaolo Benigno, Pietro Reichlin and Liyan Shi for their invaluable guidance and support. I also gratly benefited from discussions with Luca Gallorini, Valerio Nispilandi, Fabiana Sabatini, Paolo Santucci De Magistris, Francesco Sobbrio, Pierre Yared and all seminar participants at EIEF and LUISS.

†LUISS University, Viale Romania 32, 00198 Rome, Italy. Email: cfelli@luiss.it.

[1]See Mbaye et al. (2018), Roubini and Sachs (1989), Woo (2003), Crivelli et al. (2015), Persson and Tabellini (2004). For a comprehensive discussion see Yared (2019).

debt growth by imposing a multiplicity of fiscal rules: at the beginning of the 90's less than 10 countries had a fiscal rule in place, for a grand-total of a dozen rules; the corresponding numbers for 2015 are over 90 nations and more than 250 rules.[2] Yet, design of fiscal rules "is generally based on ad hoc criteria rather than theoretical considerations" (Eyraud et al., 2018), and debt, deficit or expenditure limits are typically chosen on a country by country basis as a function of historical records of default and similarly relevant economic variables.[3] Given the impressive proliferation of rules imposed by national or international fiscal institutions, it seems pivotal to provide a theory-based approach to rule selection that can have practical and empirical validity.

This work falls within the commitment versus flexibility literature, pioneered by Amador et al. (2006), which formalizes the common rationale for having fiscal rules.[4] Rules are necessary to offset a spending bias in fiscal policy: they provide a *commitment* device trough which governments can avoid overspending. Yet, rules also come at the cost of reducing fiscal policy's *flexibility* in reacting to adverse economic conditions. My research extends this strand of the literature in several ways, but its main contribution is to characterize optimal rules in a fiscal union, namely, under the assumption that the central authority can provide cross-country subsidies. The model, which is calibrated on European data, supplies a useful framework to discuss whether financial assistance should be provided to members of a union and, if so, whether it should be in the form of grants or credits – a point of disagreement among EU member countries in the recent design of the Recovery Fund.

Transferring resources across countries subject to idiosyncratic shocks provides advantages in terms of insurance, yet, when union members have private information on the shock and have a tendency to overspend, the additional funds might result in increased wasteful spending, which is detrimental to welfare. In response to this trade-off, a wide variety of arrangements offering different bundles of rules and grants has been set-up across the globe. While the U.S. combines sizable federal transfers with a requirement that state budgets balance in the medium or short run; the withholding of regional subsidies can be used as punishment for fiscal rules' breaching in Argentina.[5] Europe, on the other hand, mostly relies on deficit and debt rules and is reluctant to set-up risk-sharing arrangements: the European Union (EU) is grounded in a legal framework that explicitly forbids cross-country bailouts or joint debt liability, it enforces a balanced budget rule, and its emergency credit institution (the ESM) can normally only extend credit lines under strict conditionality.[6]

Set up & Main Results. This paper focuses on the design of optimal fiscal rules at a supranational level and, in particular, considers two distinct environments: one in which transfers across union members

[2] IMF Fiscal Rules Dataset, Schaechter et al. (2012).

[3] For more details, see Baum et al. (2018).

[4] See, for instance, Vitor Gaspar's keynote address delivered during the recent workshop on "Fiscal rules in Europe" organized by the Directorate General of Economic and Financial Affairs (DG ECFIN).

[5] For a comprehensive survey of fiscal rules around the globe see Lledó et al. (2017).

[6] See the *Treaty on the Functioning of the EU*, Art.125 and Art.310 and the *European Stability Mechanism Treaty*.

are not allowed and one in which they are.[7] I set up a mechanism design problem, in which a planner, who is unrestricted in her instrument selection, chooses allocations ensuring that *(i)* union members reveal their information truthfully and *(ii)* either each member's budget constraint (in the no-transfer set-up), or a union-wide resource constraint (in the alternative set-up) are satisfied. I consider a continuous time, infinite horizon model, with a continuum of identical governments having stochastic duration. I make two key assumptions in the model. First, governments have time-inconsistent, politically-biased preferences. Second, preferences over the value of public spending are governments' private information. Present biased preferences have been shown to arise naturally from the interaction of rational agents driven by political self-interest as a consequence of heterogeneous discounting, pecuniary externalities or political turnover; while private information over the value of spending is meant to capture the familiar argument that it is difficult to foresee or verify all possible contingencies.[8]

I find that optimal fiscal rules are of the threshold kind in both environments, namely, governments having current needs below a certain threshold should be given complete flexibility, while the rest should be restrained. However, rules are weakly more stringent when transfers are allowed. Further, in a fiscal union, debt-dependent transfers complement the set of rules. Resources are redistributed towards the countries most in need, but the more indebted they are, the fewer resources they receive. Since transfers simultaneously provide insurance and strengthen government's incentives to overspend, welfare gains derived from setting up a transfer system vanish in the limit case in which governments exclusively care about their own consumption.

Fiscal rules can be implemented as deficit limits and complemented with a combination of grants and loans when cross-country subsidies are allowed. All instruments are debt-contingent: higher public debt contemporaneously tightens deficit limits and reduces financial assistance. Further, the proposed transfer implementation has a very simple form and can easily be added to preexisting fiscal rules. When a new government is formed, it is tasked with the preparation of a budgetary document detailing its spending needs for the next subsequent years. Based on this document, the union grants an initial transfer to the country and opens a credit line from which the government can draw at any time. The only condition on this credit is that loans will automatically decrease the entity of the next transfer. In other terms, the cost of the credit line is paid in the form of decreased insurance opportunities in the future.

Maastricht, Recovery Fund & Policy Implications. One of the main policy implications of this paper is that uniform, constant thresholds across countries, like the Maastricht 3% deficit limit, are sub-optimal. Fiscal constraints contingent on preexisting debt-levels, like some of the ones detailed in the more recent

[7] Although the terminology of the paper refers to a supranational setting, the model can equivalently be applied to a supraregional environment where nations are substituted by regions, national governments by local ones and the international planner by the central government.

[8] For micro-funded models featuring a present bias see, among others Tabellini (1991), Cukierman and Meltzer (1989) (heterogeneous discounting); Weingast et al. (1981), Beetsma and Uhlig (1999), Halac and Yared (2018) (pecuniary externalities); Alesina and Tabellini (1990), Battaglini and Coate (2008) (political turnover).

fiscal compact, are much closer to the derived optimal rule. Further, the model details the optimal transfer system that should be set-up in a fiscal union. It shows under which conditions the addition of transfers should be complemented with a tightening of the fiscal rules. It can be used to frame the discussion on the entity of the overall financial help member countries should have access to, and on how this assistance should be divided between grants and credits. Lastly, while the general consensus is that current rules, and in particular European ones, are too complicated, this paper provides simple, easily enforceable rules having a single operational target.

I present a calibration of the model using EU data which shows how optimal deficit limits (as a percentage of GDP) in Europe – although debt-contingent – are not far from Maastricht 3% . Under extreme distress, member countries are entitled to transfers ranging between 3% and 4.5% of GDP depending on the level of previously accumulated debt. For example, a country having a 90% debt-to-GDP ratio and a 35% revenue-to-GDP ratio would receive a grant amounting to 3.9% of GDP when hit with the worst possible shock realization. Further, transfers should represent about 30% of the overall financial help (including the credit-line) provided by the union. For a quick comparison, consider that, under the European pandemic relief program Next Generation EU (NGEU), grants amount to 52% of the total available resources (750 billion Euros), with considerable cross-country variations. In Italy, which is one of the worst hit nations in the union, the percentage of grants is around 39% (209 billion Euros ca., of which 81.4 in grants and 127.4 in loans). Moreover, notice that the Recovery and Resilience Facility (part of NGEU) amounts to 672.5 billion, 70% of which will be distributed in the next two years. Back of the envelope calculations reveal that the planned yearly disbursement is around 1.7% of European GDP.[9]

Related Literature This paper is closest to the work in Amador et al. (2006) and Halac and Yared (2014, 2018), which falls within the mechanism design literature in self-control settings. When governments have private information on the state of the economy and a tendency to systematically exceed the socially optimum level of consumption, a trade-off arises between allowing authorities the flexibility in spending required to react to macroeconomic shocks and the commitment society would like to impose on them to dampen biased expenditures. Fiscal rules, in this setting, reduce the ability to smooth consumption at the national level, but also impose predetermined fiscal constraints that narrow the gap between socially optimal and actual policy. In Amador et al. (2006), the same trade-off between commitment and flexibility arises and the authors show that optimal fiscal rules are of the threshold kind. Depending on the specifics of the model, this threshold has been shown to vary with, among other things, the extent of the political friction (Amador et al., 2006), the persistence of shocks (Halac and Yared, 2014) and the framework in which rules

[9]NGEU resources are in 2018 prices. The 2018 EU-27 countries GDP is slightly lower than 1.3518 billions, while the Italian one is around 1.771 billion Euros (Eurostat).

are imposed (Halac and Yared, 2018).[10]

The modeling approach of this paper is akin to Amador et al. (2006), in that I assume a reduced form political bias and focus on normative prescriptions of the set-up. However, I analyze a setting in which transfers are allowed, recast the framework in continuous time and introduce random government duration, which is a tractable way to capture political turnover.[11] Although stylized, the framework detailed in here gains enough flexibility as to incorporate a range of additional features, including the possibility of sovereign default, which I explore in a companion paper. Further, it allows me to construct a viable implementation of fiscal rules, in addition to their characterization.

This work also contributes to the extensive literature on fiscal unions, including Sibert (1992), Dixit and Lambertini (2001), Cooper and Kempf (2004) and Aguiar et al. (2015), who focus on the conflicts between fiscal and monetary authorities; Von Hagen and Eichengreen (1996) who explore the possible determinants of fiscal rules; Evers (2012) and Azzimonti et al. (2016) who evaluate specific fiscal constraints; Abrahám et al. (2018) and Ferrari et al. (2020) who study insurance provision within a fiscal or monetary union.[12] Perhaps most related to this paper are Chari and Kehoe (2007) and Dovis and Kirpalani (2020), in which the need to impose fiscal constraints in a union arises from a time-inconsistency problem.[13] Differently from them, however, fiscal rules are not here meant to solve a lack of commitment on the part of the central authority. Rather, they are designed to mitigate member governments' political bias – the institutions' common rationale for having fiscal rules – as in, among others, Aizenman (1998) and Beetsma and Uhlig (1999). While most of the above-mentioned literature in this strand assumes *a priori* restrictions on the set of instruments available to the supranational fiscal authority, I solve a more general, mechanism design problem.[14] Farhi and Werning (2017) present a comprehensive analysis of policy instruments available in the context of a fiscal union, finding that state-contingent transfers provide larger benefits the more asymmetric the shocks affecting the members of a union, the more persistent these shocks, and the less open the member economies. The authors set-up a New Keynesian environment with the aim of isolating the effects of aggregate demand externalities on optimal risk sharing, explicitly setting aside concerns arising from incentive provision. This paper complements their analysis by abstracting from nominal considerations and focusing, instead, on the design of incentives when members have private information on the state of the economy.

[10]More specifically, Halac and Yared (2018) show that when interest rates are an equilibrium object, the supranational planner can account for the pecuniary externality generated by governments' accumulation strategies. Amador et al. (2006) frame the discussion around a general principal-agent problem.

[11]The mentioned papers feature a two-period model in which the incumbent values spending in the first period (while he is in charge) more than in the second period, the assumption being that some other government will be in charge in the future with probability one.

[12]For policy, rather than rule coordination at a supranational level, see, among others, Chari and Kehoe (1990), Persson and Tabellini (1995), Cooley and Quadrini (2003), Alesina and Barro (2002).

[13]In Chari and Kehoe (2007) a union-wide central bank is tempted to increase inflation when member countries have sizable debts, while in Dovis and Kirpalani (2020) it is the fiscal authority who is assumed not to have commitment.

[14]For instance, Beetsma and Uhlig (1999) limit their analysis to the European Stability and Growth Pact, while Dovis and Kirpalani (2020) set-up a Ramsey problem.

Finally, the paper relates to the vast literature on the political economy of fiscal policy, including Alesina and Tabellini (1990), Krusell and Rios-Rull (1999), Persson and Svensson (1989), Battaglini and Coate (2008) and Azzimonti (2011). As in Acemoglu et al. (2008) and Yared (2010), I study the provision of dynamic incentives to self-interested politicians, but I concentrate on an international context, rather than on the conflict between citizens and their own national government.[15] Yet, contrary to the general-equilibrium set-ups in Song et al. (2012) and Halac and Yared (2018), where supranational coordination results in an endogenously determined interest rate, unions of countries are here solely characterized by their joint fiscal constraints. I abstract form debt pricing considerations (i.e. interest rates are exogenous) to focus on optimal mechanisms and the debate on whether transfers should be part of the fiscal instruments in a union.[16]

Broadly speaking, the paper also relates to the literature on hyperbolic discounting and commitment devices à la Phelps and Pollak (1968).[17] In particular, the model presented here converges to the quasi-hyperbolic preferences set-up in Harris and Laibson (2012) for extreme values of government turnover.

The paper is organized as follows: sections 1 and 2 provide, respectively, the model set-up and the characterization of the resulting optimal allocations. Section 3 provides an implementation of optimal rules; while Section 4 presents a calibrated version of the model using European data. Concluding remarks are in Section 5.

1 Environment

There is a unit mass of countries, ruled over time by a series of governments having a stochastic duration and indexed with $n \in N = \{0, 1, 2...\}$. Time is continuous and infinite, but each government n has a finite life: is formed at time $t = \tau_n$ and dissolved at time $t = \tau_{n+1}^-$, where the time of dissolution is ex-post observable and assumed to be stochastic. I denote with $F(\cdot\ ; \lambda)$ the cdf of such exponentially distributed random variable. The arrival rate λ captures in a tractable way the frequency with which governments undergo radical transformations, thus providing a proxy for political (in)stability.[18]

The arrival of a new incumbent determines a preference change. Depending on the value they attribute to public spending, governments can be of different types θ. Types with high θ place more weight on

[15] Acemoglu et al. (2008) show, for instance, that when elected officials are as patient as their citizens, no additional distortions arise, other than those implied by their incentive compatibility constraints.

[16] More generally, this work also contributes to the literature on international or inter-regional risk-sharing, including Atkeson and Bayoumi (1993) and Bucovetsky (1998). Persson and Tabellini (1996a) explore the effectiveness of different fiscal agreements in a theoretical model comprising moral-hazard, while Persson and Tabellini (1996b) investigate how different fiscal constitutions shape insurance provision. This paper is closest to Lockwood (1999), who also sets-up a mechanism design problem in an environment in which regional authorities have private information on their idiosyncratic shocks. However, I do not model externalities in the public good provision and provide, instead, an extension focusing on political bias. Finally, for some empirical work on cross-country or cross-regional risk-sharing see, among others, Asdrubali et al. (1996), Canova and Ravn (1996) Mélitz and Zumer (2002), Afonso and Furceri (2008).

[17] See also Laibson (1997), Barro (1999), Krusell and Smith (2003), Krusell et al. (2010), Bisin et al. (2015), Lizzeri and Yariv (2017).

[18] In parliamentary systems the uncertain duration of the governments is built into the system. In presidential systems, it can be interpreted as changes in the ruling majority after midterm elections. Alternatively, stochastic duration can be though of as the risk of anticipated elections due to a political crisis.

spending than low types, who have low marginal utility of current consumption. Government preferences can be interpreted as arising from the underlying constituency's opinions on the social value of spending, which can change over time and determine an alteration of the country's stance on fiscal policy.[19] Another interpretation is that demographic changes in the constituency's composition or power struggles between different parties induce the preference shock.[20]

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space described as follows: the sample space Ω is such that $\{\omega : \mathbb{R}_+ \to \Theta | \omega$ is right-continuous with a finite number of jumps in any interval $[0,t]\}$; $\mathcal{F}$ is a σ-*field* on Ω; $\{\mathcal{F}_t\}_{t\geq 0}$ is the filtration denoting information up to time t; and P is the probability measure of a Poisson process such that jumps arrive with intensity λ and – conditional on a jump occurring at time t – the value of the process at t, $\omega(t)$, is drawn from from a continuous distribution function $H(\theta)$, within a bounded set $\Theta \equiv [\underline{\theta}, \bar{\theta}] \in \mathbb{R}_+$. Preference shocks are distributed independently over time and across governments and, without loss of generality, are normalized so as to have mean one. A key assumption is that the realization of θ is *privately observed* by the current government.

Every instant $t \geq 0$ governments receive a fixed portion κ of their country's endowment y, so that κy can be thought of as tax revenues. Utility is logarithmic and all governments suffer from a *political bias*, which I model with quasi-hyperbolic preferences (see Laibson (1997) or Harris and Laibson (2012)). Namely, although all governments discount the future exponentially at rate γ, they value spending less when they are not in office or, in other terms, they discount utility by the extra-term β whenever they are not in power. This political friction is such that the bias is stronger for lower values of the parameter β, where $0 < \beta \leq 1$. Quasi-hyperbolic preferences of this kind can be micro-funded by appealing to the interaction between turnover and political polarization as in the seminal work by Alesina and Tabellini (1990), or invoking "pork barrel" spending, as in Battaglini and Coate (2007). According to this interpretation, introducing the discount term β is a reduced form way of capturing disagreement within a country over the composition of public spending, rather than over its level. A second interpretation, is that the preference structure arises naturally from the aggregation of time consistent preferences with heterogeneous discount rates (see Jackson and Yariv (2014, 2015)).[21]

I consider the problem of a benevolent planner who can be thought of as a supranational authority and allocates governments' consumption under incomplete information about their types. Formally, I set up a direct mechanism problem in which, after observing its type, a newly formed government n provides a

[19] One possibility, for example, is that preference shocks capture responses to the actual economic conditions of the country: people may think that a lean-against-the-wind type of policy is more effective during crises. Notice that in this set-up, if utility is exponential, taste shocks are equivalent to income shocks.

[20] Think, for example, about the proportion of young and old citizens in a country, or see the entrepreneurs/workers conflict in Azzimonti et al. (2014).

[21] As it is well known, both Alesina and Tabellini (1990) and Battaglini and Coate (2007) are isomorphic to the standard quasi-hyperbolic discounting set-up in Laibson (1997). Indeed, I show that when $\lambda \Rightarrow \infty$ our model maps to a continuous time equivalent of the quasi-hyperbolic discounting framework in Harris and Laibson (2012). The possibility of achieving this mapping implies that the assumed political friction can arise from the aggregation of time consistent preferences with heterogeneous discount rates.

report $\widehat{\theta}_n$. The path of government reports is given by $\widehat{\omega} : \mathbb{R}_+ \to \Theta$ defined as $\widehat{\omega}(t) \equiv \omega(t) - \theta_n + \widehat{\theta}_n$, for $\tau_n \leq t < \tau_{n+1}$. I let $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ be the filtration generated by $\widehat{\omega}(\cdot)$. Note that $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ contains the public information up to time t, including past government reports and times of formation. Let $\sigma_n : \mathcal{F}_{\tau_n} \times \widehat{\mathcal{F}}_{\tau_n^-} \to \Theta$ be the reporting strategy of government n. I denote with σ_n^* the truthful-reporting strategy, i.e. the strategy such that $\widehat{\theta}_n = \theta_n$, for all histories. Since the planner does not observe governments' types, consumption can only depend on reports, that is, $g_t : \widehat{\mathcal{F}}_t \to \mathbb{R}_+$. Let g denote the entire sequence of consumption g_t.

Take any time $\tau_n \leq t < \tau_{n+1}$ such that government n is in power. Given consumption sequence g and the reporting strategies of all governments different from n, denoted with σ_{-n}, utility of the incumbent government is given by

$$U_t(\sigma_n|g,\sigma_{-n}) \equiv \mathbb{E}_t\left[\int_t^{\tau_{n+1}} e^{-\gamma(s-t)}\theta_n \log(g_s)ds + \beta\sum_{j=n+1}^{\infty} e^{-\gamma\tau_j}\theta_j\left(\int_{\tau_j}^{\tau_{j+1}} e^{-\gamma(s-\tau_j)}\log(g_s)ds\right)\right], \tag{1}$$

where I omitted explicit dependence on the history $\mathcal{F}_{\tau_n} \times \widehat{\mathcal{F}}_{\tau_n^-}$ to simplify notation. Preferences of the benevolent planner, instead, are described by the expected present value of governments' consumption

$$V(g,\sigma) \equiv \mathbb{E}_-\left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n}\theta_n\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)}\log(g_t)dt\right]. \tag{2}$$

Notice that there are two key differences between individual governments' and planner's preferences: the latter *(i)* has no information on governments' true types θ_n and *(ii)* places equal weights on countries' consumption irrespective of which specific government is in place (i.e. there is no β). Finally, truthful reporting is incentive compatible, given the sequence g, if, for all histories and governments,

$$U_{\tau_n}(\sigma_n^*|g,\sigma_{-n}^*) \geq U_{\tau_n}(\sigma_n|g,\sigma_{-n}^*), \text{ for all } \sigma_n. \tag{IC}$$

The incentive-compatibility condition (IC) restricts the set of allocations available to the planner by requiring that, after any history, governments must be better off truthfully reporting their type rather than lying.

I consider two separate environments, depending on whether the planner can transfer resources across countries. When such transfers are allowed, the planner must satisfy an aggregate resource constraint requiring that the present value of total allocated consumption does not to exceed the present value of collectively available resources. Formally, allocation g must satisfy the resource constraint

$$\mathbb{E}_-\left[\int_0^{\infty} e^{-rt}g_t dt\right] \leq \mathbb{E}_-\left[\int_0^{\infty} e^{-rt}\kappa y dt\right]. \tag{RC}$$

On the contrary, when transfers across countries are forbidden, the planner must ensure that each country avoids consuming more than its own resources, so constraint (RC) must hold individually for every history and for every country. In particular, given a sequence of government consumption g, let a country's wealth at any time t be defined as $a_t \equiv \int_0^t e^{r(t-s)}(\kappa y - g_s)\,ds + \kappa y/r$, and initial wealth a_0 be given by the present value of future tax income, $a_0 = \kappa y/r$. In the no-trasfer set-up, the planner must satisfy the following budget constraint for each country:

$$\dot{a}_t = ra_t - g_t, \quad a_t \geq 0, \forall t. \tag{BC}$$

Summarizing, the planner's problem when transfers are allowed is to maximize welfare subject to incentive compatibility and the resource constraint

$$v^{tr} \equiv \max_g V(g, \sigma^*), \text{ s.t. } (IC), (RC). \tag{$\mathcal{P}_{TR}$}$$

while, in the alternative environment (without transfers), the planner maximizes welfare subject to incentive compatibility and each country's budget constraint

$$v^{nt} \equiv \max_g V(g, \sigma^*), \text{ s.t. } (IC), (BC). \tag{$\mathcal{P}_{NT}$}$$

Notice that Problem ($\mathcal{P}_{TR}$) is considerably more relaxed than Problem ($\mathcal{P}_{NT}$). The planner has to provide incentives according to (IC) in both setting. However, the only other limitation she has when transfers are allowed is the aggregate resource constraint (RC), which requires expected consumption across histories to be equal to the total available resources in the union. Without transfers, instead, the planner has to insure that the collection of budget constraints summarized by (BC) holds in every possible history and for all countries. It is then intuitive that the planner could do better in the relaxed problem, namely when transfers are allowed. However, this begs the question as to whether the presence of transfers interacts with incentive provision and, if so, how it alters optimal allocations.

2 Optimal Allocations

Two frictions prevent the attainment of the first best allocation in this model: (i) the fact that preference shocks θ are private information and (ii) the presence of a political bias for $\beta < 1$. Absent private information, both in the no-transfer and in the transfer set-up, the the mechanism design problem is a relaxed one, in which the incentive constraint (IC) can be dropped. If government types were observable, no incentives would have to be provided to governments for truthful revelation since the information would be public.

When governments do not have biased preferences (i.e. $\beta = 1$) their objective function coincides with the planner's one. As a result, when transfers are forbidden, planner's and governments' preferred allocations coincide, and the incentive constraint is trivially satisfied. In the transfer setting, instead, governments may still be tempted to use their private information to exploit the insurance system. The model, then, collapses to an incomplete-information insurance problem à la Atkeson and Lucas (1992).

I will refer to the *full information* allocations under the transfer and no-transfer assumption as the solutions to, respectively, problems ($\mathcal{P}_{TR}$) and ($\mathcal{P}_{NT}$) when the incentive compatibility constraint (IC) is slack. It is easy to show that, under full information, the planner only manages to deliver perfect insurance when transfers are allowed.[22] Intuitively, this is because whenever cross-subsidies between different countries are forbidden, intertemporal allocation of resources remains the only available tool to provide insurance, so the planner is unable to equate marginal utility across government types.

2.1 No Transfers

In this section, I recast the sequential problem ($\mathcal{P}_{NT}$) in its recursive formulation and characterize the solution to the planner's problem under the no-transfer assumption. Generally speaking, the planners' problem can be written as a mechanism that is recursive in promised utilities whenever preferences are standard and shocks i.i.d. – even if planner and agents have differing degrees of patience.[23] In this setting, however, although the political friction does, in some sense, make agents relatively more impatient, governments' preferences are quasi-hyperbolic.[24]

I exploit the fact that government formation is observable to solve the problem in two separate steps. In the first step, the planner determines consumption allocations for any given promised utility. In the second step, the planner chooses the overall level of expected utility and continuation utility for, respectively, the period in which the incumbent remains in charge and the time after its dissolution. The first step, then, is devoted to choosing consumption *within* the current government's tenure. The second step, instead, is the one concerned with providing the right incentives by selecting expected utility levels *across* governments.

Recursive problem. Given initial wealth $\bar{a}$, let $\mathcal{V}_0(\bar{a})$ be the set of planner's payoffs such that for all $v_0 \in \mathcal{V}_0(\bar{a})$ there exists a sequence of spending g and an associated wealth process $\{a_t\}$, which *(i)* satisfy the governments' budget constraint (BC) with initial assets $\bar{a}$, *(ii)* are such that truthful reporting is incentive compatible (i.e. constraint (IC) is satisfied) and *(iii)* deliver utility $v_0 = V(g, \sigma^*)$. Define v_n, the utility promised to government n at the time of formation τ_n, as

$$v_n = \mathbb{E}_{\tau_n^-}\left[\sum_{s=n}^{\infty} e^{-\gamma(\tau_s - \tau_n)}\,\theta_s\,\mathbb{E}_{\tau_s}\left[\int_{\tau_s}^{\tau_{s+1}} e^{-\gamma(t-\tau_s)}\log(g_t)dt\right]\right],$$

and the set $\mathcal{V}_n(\bar{a})$ analogously to $\mathcal{V}_0(\bar{a})$. Standard properties of logarithmic preferences imply that, for any $\widetilde{v}_n \in \mathcal{V}_n(\bar{a})$, we can find some $v_n \in \mathcal{V}_n(1)$ such that $\widetilde{v}_n = v_n + \log(\bar{a})/\gamma$. It is thus sufficient to only characterize the set $\mathcal{V}_n(1)$. Moreover, the combination of exponential discounting and the assumption that new governments are formed according to a standard Poisson process imply that the set $\mathcal{V}_n(1)$ is independent of time. We can thus simplify notation by dropping the subscript n.

The value of the sequential problem ($\mathcal{P}_{NT}$), v^{tr}, equals $\overline{v} + \log(\kappa y/r)/\gamma$, where $\overline{v}$ is the highest payoff in $\mathcal{V}(1)$. In the appendix, I show that $\overline{v}$ is the solution to a simple recursive problem. Formally, the planner chooses policies $\widehat{g} : \Theta \times [0, \infty) \to \mathbb{R}_+$, $\widehat{a} : \Theta \times [0, \infty) \to \mathbb{R}_+$, $u, w : \Theta \to \mathbb{R}$ so as to solve the following problem

[22] Unless $\lambda = 0$, in which case there is no uncertainty in the first place: governments are in charge and keep their type forever.
[23] See Green (1987), Sleet and Yeltekin (2006), Farhi and Werning (2007).
[24] Preferences are quasi-hyperbolic in Amador et al. (2006) too, but the authors solve a static problem.

$$\overline{v} = \max_{u,\widehat{g},\widehat{a},w\in\mathcal{V}(1)} \mathbb{E}_{-}\left[\theta u(\theta) + e^{-\gamma\tau}w(\theta) + \frac{1}{\gamma}e^{-\gamma\tau}\log(\widehat{a}_\tau)\right], \tag{$\mathcal{P}_{NT:Rec}$}$$

$$\text{s.t. } \theta \in \arg\max_{\tilde{\theta}\in\Theta}\left\{\frac{\theta}{\beta}u(\tilde{\theta}) + \frac{\lambda}{\gamma+\lambda}w(\tilde{\theta}) + \frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma\tau}\log(\widehat{a}_\tau)|\tilde{\theta}\right]\right\}, \tag{3}$$

$$\int_0^\tau e^{-rt}\,\widehat{g}_t dt + e^{-r\tau}\,\widehat{a}_\tau = 1, \tag{4}$$

$$u(\theta) = \mathbb{E}\left[\int_0^\tau e^{-\gamma t}\log(\widehat{g}_t)dt\right], \tag{5}$$

where, in the recursive formulation, the planner faces one generic government at a time.[25] Notice that the planner now chooses expected utility u and processes $(\widehat{g}, \widehat{a})$ of spending and wealth for the incumbent while in charge, together with its continuation value w after dissolution.

In general, to find $\overline{v}$ and the the optimal allocation that supports it, we would first need to characterize the entire set $\mathcal{V}(1)$, from which continuation values are chosen. It turns out, however, that $\overline{v}$ satisfies a simple self-generating property, namely that the continuation value following $\overline{v}$ is also equal to $\overline{v}$, independently of the government's reports. In other terms, when utility is logarithmic, shocks are i.i.d and transfers are forbidden, the mechanism is static: incentives are automatically provided through the budget constraint since spending more today directly translates in having fewer disposable resources tomorrow. To see this, first notice that standard arguments imply that the incentive constraint (3) is equivalent to

$$\frac{\theta}{\beta}u(\theta) + \frac{\lambda}{\gamma+\lambda}w(\theta) + \frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma\tau}\log(\widehat{a}_\tau)|\theta\right] \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta}u(z)dz + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w}, \tag{6}$$

where $(\underline{\theta}u(\underline{\theta})/\beta + \underline{w})$ is the lifetime utility of the government with the lowest type; plus a monotonicity constraint on $u(\cdot)$, which must be non-decreasing. We can thus rewrite problem $(\mathcal{P}_{NT:Rec})$ by replacing constraint (3) with (6).

Consider now the choice of continuation values. By choosing a higher continuation, the planner can increase the objective function and, at the same time, relax the incentive constraint. It is then immediate that the optimal continuation values must be such that $w(\theta) = \overline{v}$ for all $\theta \in \Theta$. Therefore, problem $(\mathcal{P}_{NT:Rec})$ becomes

$$\overline{v} = \max_{u,\widehat{g},\widehat{a},\underline{w}} \frac{1}{1-\delta}\mathbb{E}_{-}\left[\theta u(\theta) + \frac{1}{\gamma}e^{-\gamma\tau}\log(\widehat{a}_\tau)\right] \tag{$\mathcal{P}'_{NT:Rec}$}$$
$$\text{s.t. } (4), (5), (6).$$

Step 1. (Within Government) The problem can be further simplified by noticing that the incentive constraint does not directly depend on $\widehat{g}$, and that a higher end-of-life wealth $\widehat{a}_\tau$ increases the objective function while contemporaneously relaxing the incentive constraint. As a result, ceteris paribus, the planner

[25] It is possible to drop time indexes, so this generic government is formed at time $t = 0$ and dissolved at a random date τ^-. Notice that I use "hats" to denote spending and wealth *within* the incumbents' tenure.

will want to choose the highest possible $\widehat{a}_\tau$. We can then characterize the optimal instantaneous consumption for any given (expected) utility level $\overline{u}$ by solving the first step of the planner's problem, namely by choosing $(\widehat{g}, \widehat{a})$ such that

$$\max_{\widehat{g},\widehat{a}} \mathbb{E}_- \left[\frac{1}{\gamma} e^{-\gamma\tau} \log(\widehat{a}_\tau)\right], \qquad (\mathcal{P}_{NT:S1})$$
$$\text{s.t. } \overline{u} = \mathbb{E}\left[\int_0^\tau e^{-\gamma t} \log(\widehat{g}_t) dt\right] \text{ and (4).}$$

The following lemma contains the solution to sub-problem $(\mathcal{P}_{NT:S1})$ and characterizes the optimal consumption and wealth allocated by the planner to the government currently in charge, for a given utility level $\overline{u}$.

Lemma 1 *Let $k(\overline{u})$ be the solution to $\overline{u}\,(\gamma+\lambda)^2 = \log(k(\overline{u}))(\gamma+\lambda) + (r - k(\overline{u}))$. Then, for all θ, $\overline{u}$, t, the solution to $(\mathcal{P}_{NT:S1})$ is given by*

$$\widehat{g}_t = k(\overline{u}) e^{(r-k(\overline{u}))t}$$

with associated wealth process $\widehat{a}_t = e^{(r-k(\overline{u}))t}$.

Proof. In the appendix. ■

Notice that, once the type of the current government is revealed, consumption follows a deterministic path. Yet, the incumbent's time of dissolution is random, and so is the wealth left at the end of its tenure.

Step 2. (Across Governments) We now turn to the second step of the planner's problem, namely the one of choosing utility u and promised utility w subject to the governments' incentive constraint. If we substitute the allocations described in Lemma 1 into the maximization problem $(\mathcal{P}'_{NT:Rec})$ and use the incentive constraint (6) to rewrite the objective function, we obtain that the optimal choice of $(\underline{w}, u)$ must be the solution to

$$\max_{\underline{w},u\in\Phi} \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} (1 - M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w}, \qquad (\mathcal{P}_{NT:S2})$$
$$\text{s.t. } \frac{\theta}{\beta}u(\theta) + W(u(\theta)) \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta} u(z)dz + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w},$$

where

$$M(\theta) \equiv H(\theta) + \theta(1-\beta)h(\theta),$$
$$W(\overline{u}) \equiv (r - k(\overline{u}))\mathbb{E}[e^{-\gamma\tau}\tau]/\gamma + \mathbb{E}[e^{-\gamma\tau}]\overline{w},$$
$$\Phi = \{\underline{w}, u \mid \underline{w} \in W(\mathbb{R}), u : \Theta \to \mathbb{R}, u \text{ non-decreasing}\}.$$

Notice that bunching types in the upper tail of the shock distribution is always feasible, and in particular is incentive compatible under the monotonicity of u. The following lemma shows that such bunching is, in fact, optimal.

Lemma 2 *The optimal allocation* $(\underline{w}^{nt}, u^{nt})$ *satisfies* $u^{nt}(\theta) = u^{nt}(\theta^\star)$ *for all types* $\theta > \theta^\star$*, where* $\theta^\star$ *is the smallest value such that*

$$\int_{\widetilde{\theta}}^{\overline{\theta}} (1 - M(x))dx \leq 0,$$

for all $\widetilde{\theta} \geq \theta^\star$*.*

Proof. In the appendix. ■

The proof follows the lines of Amador et al. (2006) and has a very intuitive interpretation. Since shocks are multiplicative, the planner is generally unable to distinguish between a non-biased government subject to shock $\theta = \tilde{\theta}/\beta$ and a biased one with type $\tilde{\theta}$. However, government types having a sufficiently high marginal utility (i.e. $\theta > \beta\overline{\theta}$) cannot disguise themselves because there exists no value of the shock such that their preferences are equivalent to those of an unbiased government with a higher type. Separating high types would require allocating a consumption that is increasing in θ. High types, however, are already over-consuming and increasing their consumption can only reduce welfare, so the planner is better off bunching them. This result is in line with the previous literature: optimal fiscal rules feature bunching at the top of the shock distribution.

Following Amador et al. (2006), I restrict attention to shock distributions that satisfy

Assumption 1 $M(\theta)$ *is nondecreasing.*

Assumption 1 is a relatively weak requirement on the shock process which is satisfied by all log-concave distributions.[26] For differentiable densities, this is equivalent to a lower bound on the distribution's elasticity

$$\frac{\theta h'(\theta)}{h(\theta)} \geq -\frac{2-\beta}{1-\beta}.$$

When Assumption 1 is satisfied, the threshold $\theta^\star$ is implicitly defined by the following equation

$$\beta\mathbb{E}\left[\theta|\theta \geq \theta^\star\right] = \theta^\star. \tag{7}$$

Notice that the threshold above which utility becomes constant depends on the degree of political bias β. In particular, when $\beta \leq \underline{\theta}$ national governments' political bias is extremely severe and all types are allocated the same utility since $\theta^\star = \underline{\theta}$. At the same time, no type is bunched when there is no political friction at all: for $\beta = 1$, $\theta^\star = \overline{\theta}$ and each θ is offered a type-specific utility level.

I now turn to the full characterization of problem ($\mathcal{P}_{NT:S2}$) solution. The following proposition presents the main result of this section and characterizes optimal allocations in the economy without transfers.

Proposition 1 *Under Assumption 1, optimal spending and associated wealth process for* $\tau_n \leq t < \tau_{n+1}$ *are given by*

$$g_t^{nt} = a_t \cdot \begin{cases} k^{nt}(\theta_n) & for \theta < \theta^\star \\ k^{nt}(\theta^\star) & for \theta \geq \theta^\star \end{cases} \quad and \quad a_t = a_{\tau_n} \cdot \begin{cases} \exp\left((r - k^{nt}(\theta_n)(t-\tau_n)\right) & for \theta < \theta^\star \\ \exp\left((r - k^{nt}(\theta^\star)(t-\tau_n)\right) & for \theta \geq \theta^\star \end{cases}$$

[26]See Halac and Yared (2018), Amador et al. (2006).

where

$$k^{nt}(\theta) \equiv k(u(\theta)) = \frac{\gamma\theta(\gamma+\lambda)}{\gamma\theta+\lambda\beta} \quad \textit{and} \quad a_{\tau_n} = \frac{\kappa y}{r} \exp\left(\sum_{i=0}^{n-1}(r - k^{nt}(\theta_i))\tau_{i+1}\right).$$

Proof. In the appendix. ■

Proposition 1 characterizes governments' current consumption. It shows that the planner allocates a type-dependent portion of wealth to consumption. Further, this proportion is increasing in θ, meaning that optimal spending is higher when its social worth increases. Dependence on types disappears when uncertainty about the value of future spending vanishes. This uncertainty is captured by the political turnover parameter λ. Thus, when $\lambda = 0$, there is no uncertainty about subsequent spending needs and the planner allocates a constant proportion γ of wealth to current spending (i.e. $k^{nt}(\theta) = \gamma$, for all θ).

Assumption 1 guarantees that, for all types $\theta \leq \theta^\star$, optimal spending coincides with what the government would choose in a consumption-saving problem.[27] This feature yields the following simple interpretation of the planner's solution. Governments are granted full flexibility over spending decisions as long as their spending is below a certain level. Due to hyperbolic preferences, they always allocate a greater fraction of their wealth to current spending than what would be optimal for the planner. To counteract governments' desire for excessive spending, the planner limits their flexibility by introducing a bound on spending. More specifically, the optimal mechanism consists in allowing types below the threshold $\theta^\star$ to make their unconstrained choice – i.e. to enjoy full flexibility when choosing spending – and in bunching all the others. What is more, the threshold becomes tighter (i.e. more government types are constrained) when the political bias is stronger (lower β). In fact, when there is no bias at all ($\beta = 1$) full flexibility ($\theta^\star = \bar{\theta}$) is optimal, while when the bias is very high ($\beta \leq \underline{\theta}$) all types should be constrained ($\theta^\star = \underline{\theta}$). If $\beta \in (\underline{\theta}, 1)$ the threshold is monotonically decreasing in β under Assumption 1.

Remark. (Full Information) A useful benchmark is the *full information* case, which is defined as the solution to the relaxed problem obtained by dropping the incentive constraint from problem ($\mathcal{P}_{NT}$). In this benchmark, the planner obtains the preferred level of spending, which equals $g_t^f = k^f(\theta)a_t$, where $k^f(\theta) \equiv \gamma\theta(\gamma+\lambda)(\gamma\theta+\lambda)^{-1}$, and invests the remaining wealth. Comparing g^f with g^{nt}, we see that asymmetric information has a bite, that is, it distorts allocations relative to full information only when hyperbolic discounting is present (i.e. $\beta < 1$).

2.2 Transfers

I now characterize the solution to the planner's problem when transfers among countries are allowed. First, I rewrite the sequential problem of Section 1 recursively. Similarly to the no-transfer case in Section 2.1, the problem of choosing optimal allocations can be solved in two steps. In the first step, the planner takes as

[27] I show this in the implementation in Section 3.

given the current utility and the continuation value that must be delivered to the government in charge and chooses the optimal sequence of instantaneous spending. In the second step, the planner chooses current utility and continuation value optimally.

Recursive Problem. I here focus on the recursive version of the planner's dual problem, that is, the problem of minimizing the expected resources of delivering a certain lifetime utility to a given country. Standard arguments, which I present in the appendix, imply that the history of a country until the formation of a new government can be summarized by the country's continuation utility.

Let $K(v)$ be the expected amount of resources that are necessary to deliver lifetime utility v to a country, when the n-th government is formed. It satisfies the recursion

$$K(v) = \min_{\widehat{g},u,w} \mathbb{E}_{-} \left[\int_0^{\tau} e^{-rt} \widehat{g}_t dt + \frac{\lambda}{r+\lambda} K(w(\theta)) \right], \qquad (\mathcal{P}_{TR:Rec})$$

$$\text{s.t. } \theta \in \arg\max_{\tilde{\theta}\in\Theta} \left\{ \frac{\theta}{\beta} u(\tilde{\theta}) + w(\tilde{\theta}) \right\}, \qquad (8)$$

$$(\gamma+\lambda)v = \mathbb{E}_{-} \left[\theta u(\theta) + \lambda w(\theta) \right], \qquad (9)$$

$$u(\theta) = \mathbb{E} \left[\int_0^{\tau} e^{-\gamma t} \log(\widehat{g}_t) dt \right] \qquad (10)$$

where $u, w : \Theta \to \mathbb{R}$ and $\widehat{g} : \Theta \times [0,\infty) \to \mathbb{R}_+$. Notice that the planner minimizes resources subject to the constraint of delivering lifetime utility v to the country and subject to truthful reporting by the current government.

Step 1. (Within Government) Maximization problem ($\mathcal{P}_{TR:Rec}$) can be simplified by noticing that government spending g only features in the objective function and in the constraint on the incumbent expected utility (10). We can therefore characterize the optimal instantaneous consumption by minimizing the cost of delivering a given (expected) utility level $\overline{u}$.

Let $G(\overline{u})$ be the expected resources of delivering current utility $\overline{u}$ to a generic government formed at time 0 and remaining in charge until the random time τ. Then G is given by

$$G(\overline{u}) \equiv \min_{\widehat{g}} \mathbb{E}_{-} \left[\int_0^{\tau} e^{-rt} \widehat{g}_t dt \right], \qquad (\mathcal{P}_{TR:S1})$$

$$\text{s.t. } \overline{u} = (\gamma+\lambda) \mathbb{E} \left[\int_0^{\tau} e^{-\gamma t} \log(\widehat{g}_t) dt \right].$$

For simplicity, I will assume in what follows, that the interest rate is equal to the rate of time preference, $r = \gamma$. It is immediate to see that, under this assumption, the solution to the problem above entails a constant consumption over the government's lifetime.

Lemma 3 *For all θ, $\overline{u}$, t, the solution to ($\mathcal{P}_{TR:S1}$) is given by*

$$\widehat{g}_t = e^{\overline{u}}.$$

As a result, $G(\overline{u}) = \exp(\overline{u})(\gamma+\lambda)^{-1}$.

Notice that government consumption would grow (fall) deterministically over time if, instead, $r > \gamma$ ($r < \gamma$).

Step 2. (Across Governments) I now characterize the planner's choice of incentive compatible current utilities and continuation values across governments. Incorporating the results from Step 1 in the recursive problem we have that

$$K(v) = \min_{u,w} \mathbb{E}_{-}\left[G(u(\theta)) + \frac{\lambda}{\gamma+\lambda} K(w(\theta))\right], \qquad (\mathcal{P}_{TR:S2})$$
$$\text{s.t. } (8), (9)$$

First, notice that, using the homotheticity properties of logarithmic preferences, it is immediate to verify that $K(v) = K(0)\exp(\gamma v)$. Similarly, if we denote with (u_v^{tr}, w_v^{tr}) the solution to problem ($\mathcal{P}_{TR:S2}$) for some lifetime utility v, then $u_v^{tr}(\theta) = u_0^{tr}(\theta) + \gamma v$ and $w_v^{tr}(\theta) = w_0^{tr}(\theta) + v$. Therefore, it is sufficient to characterize the solution for $v = 0$.

Second, to facilitate the comparison with the no-transfer case, I recast the problem in its primal form and, proceeding as in Section 2.1, replace the incentive constraint (8) with a a monotonicity condition on u and a more convenient constraint, featuring only the continuation utility of the lowest type

$$0 = \max_{u,w,\underline{w}} \mathbb{E}_{-}\left[\theta u(\theta) + \lambda w(\theta)\right],$$
$$\text{s.t. } \mathbb{E}_{-}\left[G(u(\theta)) + \frac{\lambda}{\gamma+\lambda} K(0)\exp(\gamma w(\theta))\right] \leq K(0)$$
$$\frac{\theta}{\beta} u(\theta) + \lambda w(\theta) \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta} u(z)dz + \frac{\underline{\theta}}{\beta} u(\underline{\theta}) + \lambda \underline{w}$$

with u non-decreasing.

Third, I perform a simple change of variable which emphasizes the fact that shifting resources across types is possible in this setting. In particular, notice that $K(0)$ is the expected resources needed to deliver the desired overall utility level (i.e. $v = 0$) to the incumbent before its type is revealed. Once the shock is realized, the planner uses resources $G(u(\theta))$ and $K(0)\exp(\gamma w(\theta))$ for, respectively, current and continuation utility. I then let transfers $T(\cdot)$ capture the difference between the initial expectation and actual realization of employed resources, as a function of the incumbent's type. Formally, let $T(\theta) \equiv G(u(\theta)) - K(0) + \lambda(\gamma+\lambda)^{-1}K(0)\exp(\gamma w(\theta))$, and recast the planner's choice in terms of transfers T instead of continuation utility

w:

$$0 = \max_{(\underline{w},u,T)\in\Phi'} \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} (1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda\underline{w}, \qquad (\mathcal{P'}_{TR:S2})$$
$$\text{s.t. } \mathbb{E}[T(\theta)] \leq 0,$$
$$\frac{\theta}{\beta}u(\theta) + \lambda W\Big(K(0) - G(u(\theta)) + T(\theta)\Big) \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta} u(z)dz + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda\underline{w},$$

where

$$W(x) \equiv \frac{1}{\gamma}\log\left(\frac{\gamma+\lambda}{\lambda K(0)}x\right),$$
$$\Phi' = \Big\{\underline{w}, u, T \mid \underline{w} \in W(\mathbb{R}), u : \Theta \to \mathbb{R}, u \text{ non-decreasing}, T : \Theta \to (-K(0), \infty)\Big\}.$$

and $M(\cdot)$ is defined as in Section 2.1.

Problem ($\mathcal{P'}_{TR:S2}$) bears a close resemblance with the no-transfer problem ($\mathcal{P}_{NT:S2}$). The crucial difference is that here the planner is allowed to transfer resources across countries. Such transfers must be zero in the aggregate, as specified by the first constraint. To solve problem ($\mathcal{P'}_{TR:S2}$), I use the common approach of ignoring the monotonicity constraint on u, and then showing that the solution to this relaxed problem verifies monotonicity.

I begin the solution characterization by showing that the problem with transfers also features bunching at the top.

Lemma 4 *Under Assumption 1, the policy functions u_v^{tr}, T_v^{tr} satisfy $u_v^{tr}(\theta) = u_v^{tr}(\theta^\star)$ and $T_v^{tr}(\theta) = T_v^{tr}(\theta^\star)$ for all types $\theta > \theta^\star$ and for all continuation values v.*

Proof. In the appendix ■

This lemma is the analogue of Lemma 2 and shows that the solution to problem ($\mathcal{P'}_{TR:S2}$) is constant for types above the same threshold $\theta^\star$. Notice, however, that the lemma does not exclude the possibility that solutions become constant for types below $\theta^\star$, as it will be the case for some type distributions.

Further, to obtain a sharp solution characterization, I restrict attention to shock distributions satisfying the following assumption.

Assumption 2 *The shock distribution $h(\theta)$ is such that*

$$\textit{if} \quad \tilde{\theta}\,\frac{\varphi'(\tilde{\theta})}{\varphi(\tilde{\theta})} < -\frac{\lambda\beta}{\gamma\tilde{\theta}+\lambda\beta} \quad \textit{for some } \tilde{\theta} < \theta^\star, \quad \textit{then} \quad \theta\,\frac{\varphi'(\theta)}{\varphi(\theta)} < -\frac{\lambda\beta}{\gamma\theta+\lambda\beta} \quad \textit{for all } \theta \in [\widetilde{\theta}, \theta^\star],$$

where $\varphi(\cdot)$ is a non-negative function defined as

$$\varphi(\theta) \equiv \frac{1-M(\theta^\star)}{h(\theta)} + \frac{\gamma\theta+\lambda\beta}{\gamma\beta}\cdot\frac{m(\theta)}{h(\theta)}$$

and $m(\cdot)$ is the derivative of $M(\cdot)$.

This is a weak requirement on the distribution of types, whose role will be clear once I provide the characterization of optimal allocations.[28] Intuitively speaking, this assumption guarantees that the relaxed problem delivers a policy u that is a single-peaked function of types θ. This, in turn, implies that the solution to problem ($\mathcal{P}'_{TR:S2}$) will still be characterized by a single threshold rule on spending.

The following proposition characterizes optimal spending in a fiscal union, namely when the planner is allowed to transfer resources across countries.

Proposition 2 *Suppose assumptions 1 and 2 are satisfied. Then, there exists a threshold $\theta^{\star\star} \leq \theta^{\star}$, such that optimal government consumption and transfers are given by*

$$g_t^{tr} = K(v_n) \cdot \begin{cases} k^{tr}(\theta_n) & \text{for } \theta < \theta^{\star\star} \\ k^{tr}(\theta^{\star\star}) & \text{for } \theta \geq \theta^{\star\star} \end{cases} \qquad \text{and} \qquad T_{v_n}^{tr} = K(v_n) \cdot \begin{cases} (\alpha(\theta_n) - 1) & \text{for } \theta < \theta^{\star\star} \\ (\alpha(\theta^{\star\star}) - 1) & \text{for } \theta \geq \theta^{\star\star} \end{cases}$$

for $\tau_n \leq t < \tau_{n+1}$, where $\{v_n\}$ is the sequence of government's continuation values constructed recursively using the policy function w_v^{tr} obtained in problem ($\mathcal{P}'_{TR:S2}$); $k^{tr}(\theta) \equiv \alpha(\theta)\, k^{nt}(\theta)$; and

$$\alpha(\theta) \equiv \frac{\varphi(\theta)}{\int_{\underline{\theta}}^{\theta^{\star\star}} \varphi(\theta) h(\theta) d\theta + \varphi(\theta^{\star\star}) \int_{\theta^{\star\star}}^{\overline{\theta}} h(\theta) d\theta}.$$

Proof. In the appendix. ■

Proposition 2 shows that optimal spending takes a particularly simple form. In particular, at any point in time, the planner guarantees that the resources given to each country are enough to achieve its promised continuation utility (i.e. $K(v_n)$). Each government will then consume a type-dependent fraction k^{tr} of such resources. What is more, this fraction coincides with its counterpart in the no-transfer case (i.e. k^{nt}) rescaled using weights α – where $\mathbb{E}[\alpha] = 1$ and $\alpha > 0$. Below, in the decentralization of the optimal allocation, I show that such resources can be interpreted as a country's wealth at the moment the previous government is dissolved. I will thus use this interpretation to compare optimal spending with its counterpart in the no-transfer environment. However in the full information case (i.e. $\beta = 1$) the comparison is immediate.

Remark. (Full Information) The type distribution only matters insofar as $\beta < 1$. When $\beta = 1$ weights are $\alpha(\theta) = (\gamma\theta + \lambda)(\gamma + \lambda)^{-1}$, so $\alpha(\theta)k(\theta) = \gamma\theta$ and marginal utility is constant, meaning that the planner provides full insurance. Whenever marginal utility is not constant, this is due to the bite of incomplete information.

In contrast, full insurance was not achievable without transfers, even under complete information (i.e. $\beta = 1$). In fact, marginal utility turned out to be increasing in θ in Section 2.1.

[28] Assumption 2 is always satisfied when the shock distribution is uniform. It is also satisfied for the calibration in Section 4. In fact, the requirement was never violated in all the simulations using a normal or log-normal distribution. If Assumption 2 is not satisfied, the solution is obtained by the "*bunching and ironing*" method described in Bolton et al. (2005), Section 2.3.3.3, p.88.

When transfers are available, they do not necessarily alter the optimal bunching threshold. More specifically,

Proposition 3 *Suppose assumptions 1 and 2 are satisfied. The bunching threshold* $\theta^{\star\star}$ *is such that*

$$\begin{aligned} \theta^{\star\star} &= \theta^{\star} \qquad \text{if} \qquad \theta\,\frac{\varphi'(\theta)}{\varphi(\theta)} \geq -\frac{\lambda\beta}{\gamma\theta+\lambda\beta} \quad \text{for all } \theta \leq \theta^{\star}, \\ \theta^{\star\star} &< \theta^{\star} \qquad \text{otherwise.} \end{aligned}$$

Proof. In the appendix. ■

To understand the proposition, remember that, by Proposition 2, government spending in the transfer set-up is a weighted average of its no-transfer counterpart. In addition, weights α are proportional to φ. The first condition of the proposition, then, expresses a lower bound on the rate of change of weights as a function of types θ. When it is satisfied, the relative importance of high types is sufficient to guarantee that the monotonicity constraint does not bind. On the contrary, when weights fall sufficiently fast, the planner would like to transfer resources away from very high types in a way that violates incentive compatibility. As a result, she lowers the bunching threshold to ensure consumption monotonicity and, thus, truthful revelation.

Closed-form Solution. (Uniform Distribution) When shocks have a uniform distribution, $\varphi(\theta)$ turns out to be an increasing function of types θ, so the first condition in Proposition 3 is always satisfied, meaning that the transfer and no-transfer case have the same threshold $\theta^{\star\star} = \theta^{\star}$. Further, the expression in (7) can be used to obtain the closed-form solution for the threshold $\theta^{\star} = \overline{\theta}\beta/(2-\beta)$.

Consumption weights $\alpha(\theta)$ take a particularly simple form, since the term $m(\theta)/h(\theta)$ simplifies to $(2-\beta)$, and are a linear, increasing function of types θ. As a result, it is immediate to show that transfers as a proportion of resources satisfy

$$\frac{T_v^{tr}(\theta)}{K(v)} = \overline{T}_0(\lambda,\beta)\left(\theta - 1 + \overline{T}_1(\beta)\right),$$

for some non-negative functions $\overline{T}_0(\cdot,\cdot)$, $\overline{T}_1(\cdot)$ given in the appendix. Furthermore, $\partial\overline{T}_0(\lambda,\beta)/\partial\lambda < 0$.

Transfers induce low types (i.e. for $\theta < 1 - \overline{T}_1(\cdot)$) to spend less and, consequently, high types (i.e. for $\theta > 1-\overline{T}_1(\cdot)$) to spend more relative to the economy in Section 2.1. The reason is intuitive. The planner faces a trade-off between granting greater flexibility to governments – which have superior information about their preferences – and suffering from their excessive spending desire – which stems from hyperbolic discounting. When transfers are not allowed, the planner solves this trade-off by imposing a cap on spending. The resulting allocation is imperfect: low types spend excessively while high types spend too little. In a fiscal union, instead, the planner has an extra tool to insure governments against fluctuations in their spending needs and, at the same time, limit their excessive spending. More specifically, by transferring resources from low types to high types, the planner makes consumption more sensitive to the shock realization and, by doing so, reduces the volatility of governments' marginal utility.

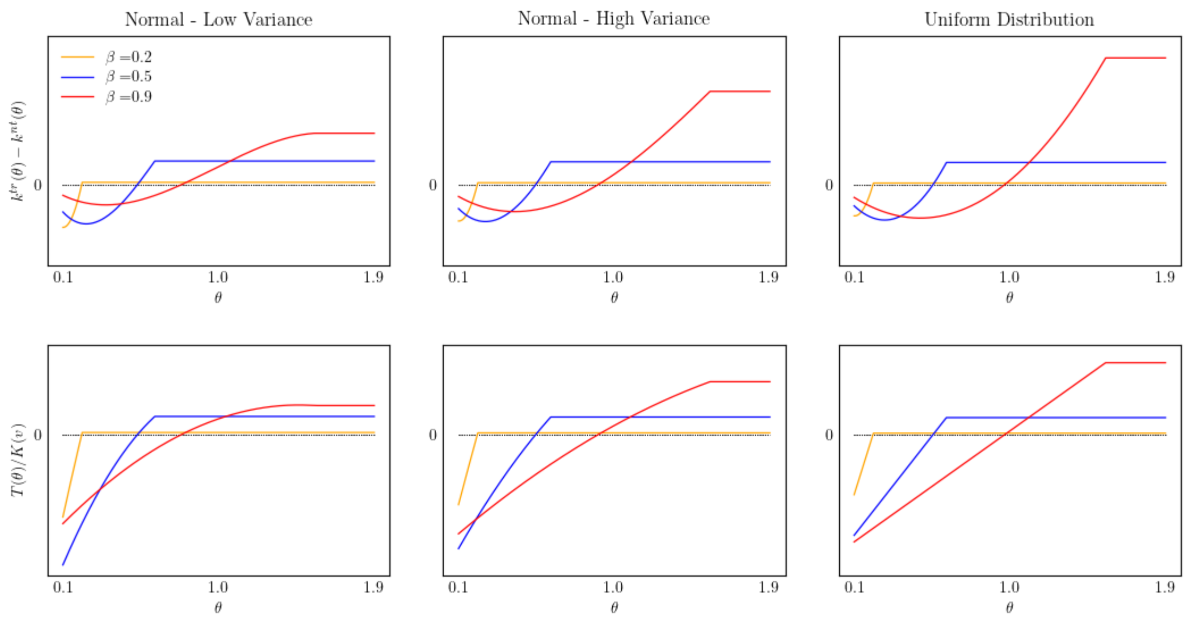

Figure I Consumption Comparison & Transfers.
Upper Panels: Difference in consumption in the transfer vs. no-transfer case as a function of θ for two governments having equal initial resources. The figures plot $(k^{tr} - k^{nt})$ for different values of β and different shock distributions. *Lower Panels*: Transfers as a percentage of resources for different values of β and different shock distributions.

At the same time, however, increasing a biased government's wealth with further resources exacerbates its incentives to lie, so the planner has to carefully balance insurance and incentives provision in its choice of transfers. When the political friction worsens (β decreases), the planner tightens fiscal rules by constraining more government types ($\partial\theta^\star/\partial\beta > 0$) while contemporaneously diminishing the entity of their subsidies ($\partial T_v^{tr}(\theta^\star)/\partial\beta > 0$). Intuitively, longer government duration relaxes this trade-off between insurance and incentives by extending the period of time in which governments discount exponentially. Formally, since $\overline{T}_0$ increases as λ diminishes (i.e. for infrequent turnover), transfers become larger in absolute value. As governments last for longer periods of time, the political friction becomes less and less important, until it completely vanishes in the limit case in which one government lasts forever (i.e. for $\lambda \to 0$). Accordingly, incentive provision also becomes less relevant, thereby allowing for more substantial cross-country subsidies.

Finally, we can compute welfare gains from transfers as the amount of additional resources the planner would require in order to give up the possibility of making transfers. The present value of initial resources in the no-transfer case is $\kappa y/r$ and delivers utility v^{nt}. To deliver the same utility in the transfer case, the planner would need resources $K(v^{nt})$. Therefore, welfare gains as a proportion of the country's endowment y are

$$\Psi = \left(\frac{\kappa}{r} - \frac{K(v^{nt})}{y}\right) = (1 - \psi(\beta))\,\frac{\kappa}{r} \tag{11}$$

where $\psi(\beta)$ is defined in the Appendix B.4. Two things are worth noticing. First, in the limit case in which current governments do not care about the future at all, for $\beta \to 0$, welfare gains completely vanish since $\lim_{\beta\to 0}\psi(\beta) = 1$. On the contrary, welfare gains are strictly positive when the present bias is not strong: $\lim_{\beta\to 1}\psi(\beta) > 1$ (proof in the appendix). Although this is not the typical moral hazard set-up, then, there still is a sense in which grants generate a trade-off between welfare improving risk-sharing and an increase of the temptation to overspend for the 'undisciplined' national governments.[29] Second, welfare gains are proportional to tax revenues.

Numerical Illustration. Figure I plots the difference in consumption between two governments who happen to have the same wealth in the transfer and no-transfer scenario, for different values of β and different distributions – where, anticipating the implementation results, I let wealth be $a = K(v^t)$ in this set-up. The numerical simulations suggest that the conclusions of the closed-form case extend to a different shock distribution and, in particular, to normally distributed types. In addition, transfers seem to be larger (in absolute value) when the distribution of types is more spread out, especially for high values of β.[30] Again, the reason has to do with the value of insurance: when governments' types are likely to be very different, insurance becomes more valuable, thus, the planner relies on transfers more heavily.

3 Implementation

The direct mechanisms described in this paper already has a natural interpretation in terms of rules and redistribution: a central institution is tasked with collecting tax-revenues on behalf of its members and allocates resources optimally. Yet, while the existence of such an arrangement seems palatable, and is indeed frequently enforced in inter-regional contexts, it might be difficult to see the practical applications of this model in a supranational environment since sovereign nations are generally reticent to surrender fiscal authority.

This section aims at describing a set of rules which implement the optimal allocation under the assumption that governments choose spending autonomously. To do so, I consider a Markov game among current and future governments within a same country, and solve for individual governments' optimal choice under the fiscal rules described in the previous section. I then focus on the possible rule implementations such that equilibrium allocations are recursive in governments' accumulated wealth. This is because, as in Albanesi and Sleet (2006), wealth turns out to be a sufficient statistic to summarize the national history: it plays the same role promised utility performs in the recursive formulation of the direct mechanism.[31]

Since government preferences are quasi-hyperbolic, their decisions are, in general, not time consistent.

[29] The set up does not feature moral hazard because member government actions do not alter the shock distribution.

[30] Notice that the uniform distribution is obtained as the limit of a sequence of Normal distributions with variance growing to infinity.

[31] Albanesi and Sleet (2006) show that, when types are i.i.d. and utility functions are separable between consumption and labor, it is possible to decentralize constrained-efficient allocations through a tax system in which taxes depend only on current period's labor income and on individual wealth.

In addition, forward-looking governments understand, and take into account, that their actions will affect the choices of future governments.[32] I thus follow the literature on hyperbolic discounting and characterize the equilibrium of a game in which each government takes future governments' best responses as given and selects spending subject to the threshold rule described in (7), and a budget constraint (either including or excluding transfers).

Formally, I consider a Markov equilibrium such that, at any instant, the state of the economy is summarized by the country's wealth a and government type θ. Governments make a simple saving-spending choice: every instant they receive tax revenues κy and can either consume or invest in a risk-free market at interest rate r. I denote with a the present value of future wealth $a \equiv x + \kappa y/r$, where x are net assets of the country.[33] Assets follow a diffusion process $dx = (rx + y - g)dt$. Further, I denote with $J(a, \theta)$ the government's equilibrium payoff when the state is (a, θ). The government's payoff and optimal spending are the solution to the following system of Hamilton-Jacobi-Bellman equations (the proof is in Appendix C)

$$\gamma\, J(a,\theta) = \max_{g} \Big\{ \theta \log(g) + J_a(a,\theta)\,(\, rx + \kappa y - g\,) \Big\} + \lambda \Big[\, \beta \mathbb{E}[\, \Upsilon(a,\theta)\,] - J(a,\theta) \Big], \tag{12}$$

$$\gamma\, \Upsilon(a,\theta) = \theta \log(g^{\star}(a,\theta)) + \Upsilon_a(a,\theta)\,(\, rx + \kappa y - g^{\star}(a,\theta)\,) + \lambda \Big[\, \mathbb{E}[\, \Upsilon(a,\theta)\,] - \Upsilon(a,\theta) \Big], \tag{13}$$

$$\text{s.t. } g \leq k^{nt}(\theta^{\star})(x + \kappa y/r). \tag{14}$$

Equation (14) features the optimal threshold rule derived in the previous section, presented here as a cap on spending. The first equation characterizes the current government's payoff J as a function of the state (a, θ) and of the value function Υ, which captures the continuation payoff of the incumbent after its replacement by a new government (J_a and Υ_a denote the derivatives with respect to wealth). Notice that the government currently in charge discounts the future at rate γ, both during the periods in which it is in power and during the periods in which it is not. When the incumbent is replaced, there is a once-and-for-all change in discounting, represented by the additional term β in equation (12). Further, since the incumbent can only make spending decisions when in power, the maximization operator shows up exclusively in equation (12) and not in equation (13). Finally, notice that the current government takes future governments' behavior as given, hence, the term $g^{\star}$ in equation (13).[34]

Equations (12) and (13) have a very intuitive interpretation. The term $\gamma\, J$ in equation (12) is the expected value of instantaneous changes in J arising from the exponential discounting. The first term of the right hand side, $\theta \ln(g)$, is the flow utility derived from government spending, while $J_a \dot{a}$ is the expected value of instantaneous changes in J arising from the (deterministic) returns process. Finally, the term $\lambda\,[\,\beta\mathbb{E}[\Upsilon] - J\,]$ represents the expected value of the instantaneous change in J due to the possible dissolution of the current government. Equations (12) and (13) are almost identical, with one important exception. From the point of

[32] See, for example, Harris and Laibson (2012).

[33] I use the convention that positive values of x are a credit, while negative ones are a debt.

[34] Formally, the first order condition gives $g^{\star}(a, \theta) = \theta / J_a(a, \theta)$ for unconstrained types.

view of the incumbent, future streams of consumption obtained after its replacement are discounted at rate β. Yet, once control has passed to a new government, any subsequent transition is not further discounted. In fact, the two equations are identical when $\beta = 1$.

Notice that, in the absence of political instability, the problem collapses to a standard saving/spending model. In fact, for $\lambda = 0$ the current value function J is equal to the continuation value function Υ: the incumbent remains in charge forever and the type never changes. For intermediate levels of instability, $\lambda \in (0, \infty)$, this model generalizes the micro-founded set-ups in Alesina and Tabellini (1990) and Battaglini and Coate (2007) by letting the social value of spending be uncertain. Finally, with extreme political instability, $\lambda \to \infty$, a new government is formed every instant, and the set-up converges to a quasi-hyperbolic discounting model in continuous time that can be considered a generalization of the two period model in Halac and Yared (2018), thus, the model with $0 < \lambda < \infty$ provides a generalization of their policy prescription.

As it turns out, the solution to this problem is recursive in the present value of wealth a and generates a government policy function closely replicating allocations in Proposition 1. The implication, then, is that countries should, within a predetermined range, be allowed complete flexibility in their spending choices. Formally, the solution to this problem is that countries with low enough spending needs – types $\theta < \theta^\star$ – are free to select $g^\star(a, \theta) = k^{nt}(\theta)a$, even if they end up overspending with respect to their unbiased (i.e. if $\beta = 1$) choice; while incumbents with larger spending needs – types $\theta \geq \theta^\star$, are limited by a binding cap $g^\star(a, \theta) = k^{nt}(\theta^\star)a$, even if this implies a loss in terms of insurance. We can then interpret the results in Section 2.1 in the following way: the threshold separating unconstrained from constrained government types was chosen as to balance the costs of limiting insurance at the top of the distribution with those deriving from overspending at the bottom.

3.1 Fiscal Rules

I here show that the threshold rule can be implemented as either an upper bound on the growth rate of debt or a debt-contingent deficit limit. For simplicity, I focus on the case without transfers since the alternative case is analogous. Remember that at the optimum countries spend a proportion $k^{nt}(\theta)$ of their wealth a, which is bounded by $k^{nt}(\theta^\star)$. Thus, countries with types above $\theta^\star$ do not have the discretion to spend as much as they desire. Instead, they can only spend up to proportion $k^{nt}(\theta^\star)$ of their assets. Although the fiscal rule is framed in terms of a threshold on the government's type, it is immediate to see that it can be implemented as a limit on the rate at which governments borrow. In particular, recall that only wealth a and spending g are observable; and that financial assets are defined as $x = a - \kappa y/r$, where negative values of x denote a debt. Debt evolution can be inferred from wealth evolution since $\dot{x} = \dot{a}$. It follows that the threshold rule can be implemented as an upper bound on the percentage growth rate of debt given by:

$$\frac{\dot{x}}{x} \leq \left(r - k^{nt}(\theta^\star) \right) \left(1 + \frac{\kappa y}{r x} \right).$$

This upper bound is debt dependent, namely it becomes tighter as debt grows.[35] In the extreme case in which $x = -\kappa y/r$ (i.e. at the "natural debt limit"), the right-hand side of the equation above becomes zero, so that no further debt is allowed.

Equivalently, the same threshold rule can be implemented using deficit limits rather than caps on the debt growth rate. Since (primary) deficit is defined as $d \equiv g - \kappa y$, the optimal mechanism would require an upper bound on the deficit/GDP ratio:

$$\frac{d}{y} \leq k^{nt}(\theta^\star)\left(\frac{x}{y} + \frac{\kappa}{r}\right) - \kappa. \tag{15}$$

Under this alternative implementation, the rule becomes a deficit limit that is contingent on the debt/GDP ratio. Notice that, insofar as $k^{nt}(\theta^\star)$ depends on political instability, the implementation of the threshold rule varies with λ. More specifically, since $\partial k(\theta^\star)/\partial\lambda \propto (\theta^\star - \beta)$ and $\theta^\star \geq \beta$ for any value of β, higher political instability (i.e. shorter average government duration) calls for looser fiscal rules.[36]

Policy Implications. (Fiscal Rules) The optimal commitment device can be implemented using a single operational target, may it be spending, debt growth rates or deficits.[37] One key feature of the optimal fiscal rule is that it tightens with a country's indebtedness: past obligations are important in that heavily indebted governments are more constrained in their spending choices.

Clearly, this implies that a uniform threshold across countries (i.e. the Maastricht requirement) is a sub-optimal instrument. However, rules that try to link fiscal constraints with preexisting debt-levels, like some of the most recent ones, are strikingly similar to the optimal mechanism implementation derived in this section. For example, under the so called "fiscal compact", European Union member states' borrowing is constrained to be lower than either 1% or 0.5% of GDP depending on whether the country's debt-to-GDP ratio is below or above 60%.[38]

A second interesting conclusion, is that limits should only be imposed on the *speed* of debt accumulation, not on debt levels. In other terms, there is no debt anchor towards which member countries should strive in this model. The result is in contrast with many existing frameworks, including the European one, which requires long-run convergence to a 60% debt-to-GDP ratio. The implication, then, is that the existence of a political bias is not enough, from a theoretical perspective, to understand why or when debt anchors might be useful. Another friction is needed to rationalize the type of fiscal rules currently in place, and,

[35] The inequality follows from the fact that $x < 0$. When $x > 0$ the inequality must be reversed.

[36] The optimal upper bound on the percentage growth rate of debt and on the deficit as a percentage of GDP are, respectively, decreasing and increasing in λ. At the optimal threshold $\theta^\star = \beta\mathbb{E}[\theta|\theta \geq \theta^\star] \geq \beta$ since $\mathbb{E}[\theta|\theta \geq \theta^\star] \geq 1$.

[37] The specific choice of operational target is unimportant as long as all target variables are perfectly observed by the central institution. In the model, I assume that everything – except governments' types – is public knowledge. A more realistic assumption, is that the quantification of some target variables might be more precise, or have fewer measurement errors than others, in which case, targets would not be equivalent.

[38] The fiscal compact is part of the Treaty on Stability, Coordination and Governance in the Economic and Monetary Union (TSCG), signed by all EU countries, except the Czech Republic and the UK in March 2012 and entered into force in January 2013. More precisely, the rule constraints countries' Medium Term Objective (MTO) not to be below a structural balance of -1% or -0.5% of GDP. For more information, see the Vade Mecum on the Stability and Growth Pact.

in particular, I show in a companion paper that the introduction of sovereign default does, under certain conditions, deliver the imposition of a strict debt limit.

Finally, it is worth noticing that the frequency of government turnover turns out to be relevant in rules selection. In particular, shorter average government duration implies that, since shocks are more frequent in the economy, there is more scope for insurance and rules should loosen up to allow it.

3.2 Transfers

The most immediate way to implement the transfer allocation is to simply let the central authority choose consumption levels. However, in an international setting, this might be politically unfeasible. An alternative, is to use a combination of transfers and loans that are conditional on countries' wealth and on government's consumption, in the spirit of Albanesi and Sleet (2006).

More specifically, let a_- be the present value of wealth of the country at the time of a generic government's dissolution, before a new incumbent is elected. Further, let the continuation value promised to the previous government in the mechanism design formulation v be summarized by wealth a_- in this implementation. That is, v is such that $a_- = K(0)\exp(\gamma v)$. In this arrangement, the planner will infer the new government type from its spending choice at the moment of formation. One possible interpretation of this process is that the government is required to prepare a document detailing future expenditure plans from which the planner can guess the government type.[39] Formally, if the government consumes $\widetilde{g}$ in the first instant of its life, the planner infers that its type is the $\widetilde{\theta}$ satisfying $\widetilde{g} = (\gamma + \lambda)G(u_v(\widetilde{\theta}))$. By incentive compatibility, the type inferred by the planner will coincide with the government's true type, that is, $\widetilde{\theta} = \theta$.

At the time of its formation, the incumbent receives a subsidy equal to

$$\chi(\theta, a_-) = (\alpha(\theta) - 1)a_-. \tag{16}$$

Let $a \equiv a_- + \chi(\theta, a_-)$ be the present value of wealth, including transfers χ, at the time of formation, then $a = \alpha(\theta)a_-$. Similarly, let a_t denote wealth at time $0 < t < \tau$. In addition to the above transfer, the central authority provides a savings account with a credit line which works as follows. At any point in time, as long as it is in charge, the government can draw (deposit) an amount b_t from the credit line (savings account). When a new government is formed, an amount $P = \lambda^{-1} b_-$ is repaid to (received from) the central authority in the form of lower (higher) transfers, where b_- is the amount used at the time of dissolution.[40]

Under this arrangement, the government's decision problem is described by the analogue of HJB equations (12) and (13) where governments choose both consumption g and the entity of the loans (savings) b; and

[39] For example, EU members that adhere to the Stability and Growth Pact have to submit a Stability and Convergence Programme detailing the country's public finance plans.

[40] Notice that, although the final payment is proportional only to the last amount b'_-, since the time of dissolution is random, governments understand that they can be called to make such a payment at any time, thus, the optimal value of b_t is finite for all t.

assets follow a jump-diffusion process

$$dx = (rx + \kappa y + b - g)dt + \left(-\lambda^{-1} b_- + \chi(\theta', a'_-)\right) dN,$$

where N_t is the jump process and a'_- is the country's wealth at the time of dissolution net of the payment $-\lambda^{-1} b_-$.

In the appendix, I prove that governments choose $b_t^\star = \bar{b}(\theta) a_t$, with $\bar{b}(\theta) = \gamma\lambda(\theta - \beta)(\gamma\theta + \lambda\beta)^{-1}$. Notice that types with higher marginal value of consumption ($\theta > \beta$) choose a positive b_t but the next government will have to pay in the form of lower transfers, while the opposite is true for low types. In addition, the amount $b_t^\star$ is such that, at the optimum, country's wealth is constant, thus, wealth at the time of formation is equal to wealth at the time of dissolution. Since spending is a constant fraction of wealth, the latter also implies that government consumption is constant throughout its life and, in particular, is equal to $k^{nt}(\theta)a$.

Finally, the reduction (increase) in the future transfer P is such that the average balance on the account is zero:

$$\mathbb{E}\left[e^{-\gamma\tau} P - \int_0^\tau e^{-\gamma t} b_t^\star dt\right] = \left[\frac{\bar{b}(\theta)}{\gamma + \lambda} - \frac{\bar{b}(\theta)}{\gamma + \lambda}\right] a = 0.$$

The savings account is necessary to implement the optimal allocation because governments are subject to the risk of sudden termination, which requires a financial instrument whose payoff is contingent on such an event. In other terms, it provides an insurance against the observable shocks of the model. Although I followed the literature in linking the formation of a new government with the draw of a new shock (the two events always happen simultaneously), a more realistic assumption would be that the two are not perfectly correlated. This alternative is explored in a companion paper focusing on the role of political uncertainty in the selection of optimal fiscal rules.

Policy Implications. (Transfers) Let me conclude the section with a few comments on the interpretation of this implementation. First, notice that, although in this implementation transfers can be negative, it is easy to rescale the problem in such a way that transfers are always positive. What is needed, is to levy a tax on the union members such that the sum of tax and transfers is equal to zero on average. This tax would be akin to, for instance, the EU budget contributions, except for the fact that –instead of being proportional to the countries' revenues– it would be proportional to members' accumulated wealth.[41]

Secondly, the proposed implementation has a very simple form and can easily be introduced in addition to preexisting fiscal rules. The idea can be summarized as follows. When a new government is formed, it is tasked with the preparation of a budgetary document detailing its spending needs for the next subsequent years. Based on this document, the union grants an initial transfer to the country and opens a credit line from which the government can draw at any time. The only condition on this credit is that loans will

[41]EU funding comes mainly from customs duties, sugar levies and a portion of value added tax (VAT) collected on behalf of the EU by the Member States. Additional contributions are made in proportion to the members' gross national income.

automatically decrease the entity of the next transfer. In other terms, the cost of the credit line is paid in the form of decreased insurance opportunities in the future. When a new government, with different spending needs is formed, the process starts anew, with the caveat that the bargaining process for grants will take into consideration the resources previously drawn from the credit line.

The described credit line would not be very different from some existing programs. The European unemployment insurance scheme SURE, for example, extends loans to member states with the aim of mitigating sudden increases in public expenditure related to employment protection.[42] The main difference would be that, rather than being conditional on pre-specified consumption items or reforms, loans would only be conditioned on wealth: the more indebted the country is, the less loans it could receive. Further, loan repayment would decrease the entity of the next transfer rather than being repaid at a set date.

4 Application: Fiscal Rules in the EU

One of the main concerns that prevents practitioners from adopting theory-inspired models in the actual selection of fiscal rules, is that theory often produces excessively stylized set-ups.[43] Although the one presented here also is an extremely simple model, this section aims at showing that it can effectively be taken to the data. Moreover, as shown in a companion paper, the set-up is flexible enough as to accommodate other practical concerns (e.g. the existence of default) and represents, as such, a first step toward the integration of policy and theory oriented literature strands.

The question we are going to ask in this section is: what should have been the fiscal rule adopted by Maastricht signatories in 1993 according to this model? The idea behind this exercise is to *(i)* use data on average government duration to estimate the political uncertainty parameter λ, *(ii)* identify the polarization parameter β and the shock distribution $H(\theta)$ through the model, *(iii)* compute the optimal thresholds $\theta^{\star}, \theta^{\star\star}$ and their corresponding fiscal rules, *(iv)* quantify optimal transfers across union members.

Data Sources. Data on debt and GDP have been obtained from the recently compiled Global Debt Database (GDD), while the series of government revenue and government expenditures as a percentage of GDP are from the Macro-economic database of the European Commission's Directorate General for Economic and Financial Affairs AMECO.[44] Historical information on government duration has been obtained from The Party Government Data Set (PGDS). All the original signatories of the Maastricht treaty have been included in the analysis, except for Denmark and the United Kingdom.[45] As previously mentioned, optimal threshold rules can be implemented in several ways. However, to allow for a more immediate comparison

[42] SURE stands for "Support to mitigate Unemployment Risks in an Emergency".

[43] The complaint is expressed, for example, in Eyraud et al. (2018).

[44] Data in GDD are in nominal terms. To compute gross debt real growth I subtract inflation from nominal growth using the World Development Indicators database of the World Bank (WDI).

[45] The two countries have been excluded because the opt-outs they managed to obtain render comparison with the other EU countries problematic. Nonetheless, results are robust to their inclusion.

Table 1: Parameters values and estimates.

Parameter	*Value*	*Source or Target*
Discount rate	$\gamma = 0.05$	Set 5% yearly interest rate.
Turnover frequency	$\lambda = 0.68$	Estimate from average government duration.
Political bias	$\beta = 0.36$	Estimate from average debt growth.
Standard deviation of shocks	$\sigma = 0.96$	Estimate from debt growth.

with Maastricht 3% deficit rule, I will focus on debt-contingent deficit limits.[46]

Political Uncertainty, Polarization & Shock Distribution. In this model, the political uncertainty parameter λ simply is the arrival rate of a new government, so it can be easily estimated as the inverse of average government's duration. An interesting issue, however, concerns the definition of "government" itself. Is it enough to change a few key figures or the ruling party coalition to establish that a new government has been formed? Or does a country also need to select a new prime minister? The answer is likely to change on a country by country basis, depending on the institutional peculiarities of the various nations.[47] Yet, to ensure consistent estimates across different national states, government duration should be defined in the same manner. I adopt the loosest possible definition of government, according to which any change is a government change. More specifically, in accordance with the PGDS definition, any official government resignation and any change in prime minister or party composition of the cabinet is considered a government switch.[48] Let $Dur_{i,n}$ be the number of years government n in country i lasts. The political uncertainty parameter in country i is the inverse of government duration's sample average $\lambda_i = N(\sum_n \mathrm{Dur}_{i,n})^{-1}$. The introduction of fiscal rules should not, according to the model, have any effect on government duration. Nonetheless, to avoid endogeneity problems, the sample only includes governments formed before 1993 (the year Maastricht entered into force).

The polarization parameter β is meant to capture the present government discount of future governments' spending. Since there is no clear consensus on how polarization in democratic systems should be measured, one advantage of this estimation strategy is that β can be identified through the model.[49] Assume that the preference rate γ is equal to the interest rate r and use (i) the unconstrained asset growth equation

[46]Maastricht Treaty actually requires countries to keep debt to GDP levels below 60%, deficits lower than 3% of GDP, or to reduce debt in excess of the 60% threshold by $1/20^{\text{th}}$ of the distance every year. As mentioned before, however, there is no reason, in this model, to constraint debt levels, just debt growth, so only the deficit rule will be considered in this application. Further, I here only consider the corrective arm of EU rules for clarity purposes, but an interesting extension would be to compare optimal rules to the new, post 2008 financial crisis, prescriptions of the EU framework's preventive arm. Medium term objectives (MTO), for example, are generally stricter than Maastricht deficit limit.

[47]To give you an example, a head-of-the-state change is extremely relevant in presidential systems like France, less so in parliamentary ones such as Italy.

[48]Including cases in which, after resigning, a government with the same prime minister and party composition of the previous one is formed. The number and time-span of observations is variable in the PGDS, as every country had its own beginning of democratic life and number of governments.

[49]An alternative, could be to use a measure of disagreement within the government (e.g. the Partisan Conflict Index of Azzimonti (2018)) to proxy for β.

$\dot{a}/a = [(r-\gamma)\gamma\theta + (r\beta - \gamma\theta)\lambda]\,(\gamma\theta + \lambda\beta)^{-1}$, and (*ii*) the fact that preference shocks have mean one. Taking expectations and rearranging we have that, for $\gamma = r$

$$\beta = \frac{\gamma}{\lambda}\mathbb{E}\left[\frac{\gamma - \dot{x}/x}{\lambda + \dot{x}/x}\right].$$

Let $Debt_{i,t}$ be the real debt of country i in year t, I compute the sample growth rate of debt as the first difference in the log series of debt: $g^d_{i,t} = log(Debt_{i,t+1}) - log(Debt_{i,t})$. We then have that the polarization parameter of country i is $\beta_i = \gamma\,(\lambda_i T)^{-1}\sum_t(\gamma - g^d_{i,t})(\lambda_i + g^d_{i,t})^{-1}$ where I set $\gamma = 5\%$. The assumption on interest rates being equal to the time preference rate substantially simplifies exposition, but is in no way necessary.

Notice that, in the model, x represents net assets, rather than gross debt. However, measurement issues in the quantification of national assets are such that debt data are to be preferred; especially considering that, when government assets are relatively stable over time, net asset and debt growth are one and the same.[50] One possible alternative would be to construct a series for government net debt by subtracting financial assets from gross debt. Growth rates estimated using general government net financial assets instead of debt are quite similar, but rules, being debt-dependent, are substantially looser when computed with net assets. The results of the calibration then, can be seen as an upper bound on how strict rules should be.

An objective interpretation of the distribution of shocks implies that, given a utility function, the distribution $H(\theta)$ can be identified from unrestricted behavior. In fact, if countries have full flexibility, the observed growth rate of debt identifies the distribution of preference shocks, given the utility functions and the polarization and political uncertainty parameters β and λ.[51] Namely, we can compute preference shocks θ from

$$\theta = \frac{\beta\lambda}{\gamma}\left(\frac{\gamma - \dot{x}/x}{\lambda + \dot{x}/x}\right).$$

Let $\theta_{i,t}$ be the preference shock experienced by country i at time t, then we have that $\theta_{i,t} = \beta_i\lambda_i(\gamma - g^d_{i,t})(\gamma\lambda_i + \gamma g^d_{i,t})^{-1}$. The country specific shocks are then pulled together, and a truncated normal distribution is fitted to the data. Finally, to have a union-wide political uncertainty and polarization parameter estimates, I compute a weighted average of the individual countries' parameters, where the weights w_i for country i are given by its relative contribution to the 1995 GDP of the union, namely $w_i = GDP_{i,1995}/\sum_i GDP_{i,1995}$, and the union-wide parameters $\widehat{\lambda} = I^{-1}\sum_i w_i\lambda_i$ and $\widehat{\beta} = I^{-1}\sum_i w_i\beta_i$.

Thresholds & Optimal Fiscal Rules. We now have all the information needed to compute both the optimal threshold $\theta^\star$ and the optimal rule each individual country should have imposed according to the model. I use equation (7) to compute the optimal threshold for the union $\theta^\star$, and the deficit rule in Section

[50] It is unclear, for example, if illiquid assets should be included in the assessment of debt sustainability, and if yes how to value them.

[51] Remember that $\dot{a}/a = \dot{x}/x$ in the model.

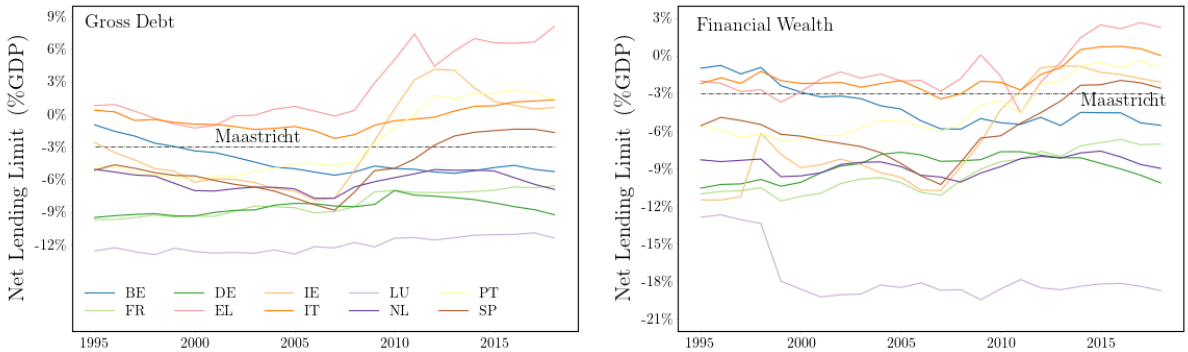

Figure II Optimal Net Lending Limit
*B*oth graphs have net lending on the vertical assets: positive numbers are a surplus while negative numbers are a deficit. Each line is a country specific net-lending limit. To comply with the rule, countries should choose net lending above their line. *Left Panel*: Net lending limit computed with general government debt data (AMECO). *Right Panel*: Net lending limit computed with financial wealth data (World Bank).

3.1 to calculate the minimum allowable net lending of the European nations for the period 1995-2018.[52] Remember that fiscal rules can be implemented as debt-contingent deficit limits, so the net-lending limits from 1995 to the present are a function of the countries' realized debt to GDP ratio in that particular year. Let $Debt\_Ratio_{i,t}$ and $Rev\_Ratio_{i,t}$ be, respectively, the debt to GDP and revenue to GDP ratio for country i at time t. Then the country's net lending limit (as a percentage of GDP) $\bar{L}_{i,t}$ is computed as

$$\bar{L}_{i,t} = \left[1 - \frac{\gamma\widehat{\theta}^{\star}(\gamma + \widehat{\lambda})}{\gamma\widehat{\theta}^{\star} + \widehat{\lambda}\widehat{\beta}} \left(\frac{Debt\_Ratio_{i,t}}{Rev\_Ratio_{i,t}} + \frac{1}{r} \right) - \frac{Debt\_Ratio_{i,t}}{Rev\_Ratio_{i,t}} r \right] Rev\_Ratio_{i,t}.$$

I compute the net lending limit $\bar{L}_{i,t}$ for each country i and time $t \in [1995, 2018]$ using the realized revenue and debt to GDP ratio for that particular year. Figures II and III summarize the results.

It is important to notice that, since debt is an endogenous variable of the model, and we are using the realized debt and revenue series, the net lending limit $\bar{L}_{i,t}$ does not provide an estimate of how countries' debt would have evolved had they implemented the model's rules. Rather, it shows what the optimal net-lending limit would have been, had they decided to introduce a fiscal rule in that specific year. The left panel in Figure II, for example, shows that had Europe, at the beginning of the century, decided to institute fiscal rules, the optimal choice would have been to allow France to run deficits up to 9% of GDP (given its 2000 debt level); while Italy and Greece should have been required to run a balanced budget. Overall, the 3% deficit limit imposed by the Maastricht Treaty was, according to this model, too loose for about half of the union members and too strict for the other half, suggesting that the political bargaining process in fiscal constraint selection managed to find a middle ground on which all member could agree.[53] However, as previously pointed out, the limit should be estimated using net assets. The right panel in Figure II presents

[52]In Section 3.1 we characterized the primary deficit limit. The net lending limit can be easily inferred by adding interest repayment and changing the sign.

[53]The extremely loose fiscal constraint in Luxembourg is due to its historically low debt to GDP ratio.

a robustness check in which net financial assets, rather than gross debt are used in the computation of lending limits $\bar{L}_{i,t}$. It shows rules that are considerably looser with respect to the ones in the right panel. If one where to include illiquid public capital (i.e. public buildings), optimal rules would become even looser, indicating that a 3% deficit limit is too tight for the union.

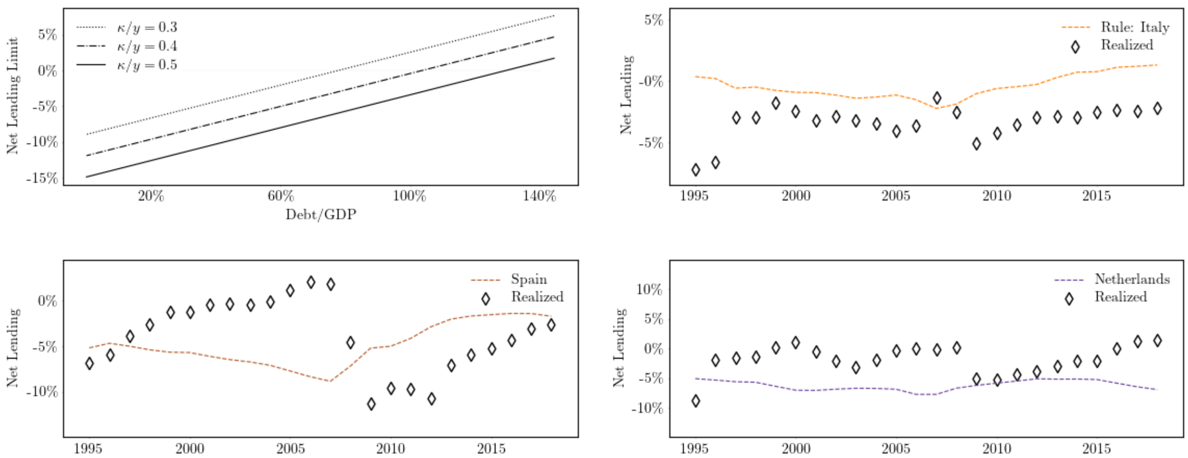

Figure III Optimal Lending Limit vs. Realized Lending
Upper Left Panel: Optimal fiscal rule as a function of Debt/GDP ratio for different Revenue/GDP levels ($\kappa/y \in [0.3, 0.4, 0.5]$). *Other Panels:* Comparison between the prescribed minimum net lending (colored line) and realized primary deficit/surplus as a percentage of GDP (diamonds) for Italy, Spain and the Netherlands (AMECO) .

Figure III depicts, in the upper-left panel, the net lending limit as a function of debt for the estimated union threshold $\widehat{\theta}^\star$ and for different revenue levels. The higher the revenue to GDP ratio, the higher the permissible deficit for a given debt to GDP ratio. For example, a 75% debt to GDP ratio corresponds to a maximum deficit of around 5% if revenues account for 40% of GDP, but only allows for a balanced budget with a revenue to GDP ratio of 30%. The other three panels, instead, compare realized net lending in Italy, Spain and the Netherlands against the country-specific net lending limit $\bar{L}_{i,t}$ for the years between 1995 and 2018. Countries are not compliant with the rule whenever the realized surplus/deficit is below the country's minimum net-lending $\bar{L}_{i,t}$ (colored line). While the Netherlands has almost always managed to sustain large enough surpluses (or low enough deficits) to comply with its optimal deficit limit, Spain has over-borrowed after the financial crisis.[54] Italy, on the other hand, would have been required to keep a balanced budget (if not a small surplus) under this rule, but has almost always run a deficit.

Grants, Credit & Welfare Gains. To quantify optimal transfers, the first step is to find consumption shares $\widehat{k}^{nt}(\theta)$ and weighs $\widehat{\alpha}(\theta)$ using Proposition 2 and the previously estimated parameters, including the shock distribution. As it turns out, $\widehat{\theta}^\star$ is such that the first condition in Proposition 3 is satisfied. In

[54] Yet, it could be argued that the shock distribution has changed after the financial crisis and that its variance should be re-estimated for the post-2008 period.

other terms, since the estimated share of consumption for the transfer case $\widehat{k}^{tr}(\theta)$ is non-decreasing in θ, the no-transfer threshold coincides with the transfer one $\widehat{\theta}^{\star\star} = \widehat{\theta}^{\star}$. If the European Union implemented the optimal fiscal rules, then, it could add transfers to its instrument set without altering deficit limits. Let $Spend\_Ratio_{i,t}$ be the planned government expenditure as a proportion of GDP for country i in year t. The planner can infer $\theta_{i,t}$ from optimal spending in Proposition 2 as

$$\widehat{k}(\theta_{i,t})\widehat{\alpha}(\theta_{i,t}) = \frac{r\, Spend\_Ratio_{i,t}}{r\, Debt\_Ratio_{i,t} + Rev\_Ratio_{i,t}},$$

and compute the due (annualized value of) transfers, $T\_Ratio_{i,t}$, as a percentage of GDP from, equivalently, Proposition 2:

$$T\_Ratio_{i,t} = (\gamma + \widehat{\lambda})(\widehat{\alpha}(\theta_{i,t}) - 1)\left(Debt\_Ratio_{i,t} + \frac{Rev\_Ratio_{i,t}}{r}\right).$$

As an example, to get a sense of the transfers' entity, we can consider the case of Spain in 2012. Had the European Union decided to start implementing fiscal rules from that year on, Spain would have been allowed to run deficits up to 3% of GDP, a rule that coincides with the actual Maastricht requirement. Using data for spending, debt and revenue ratios, it turns out that Spain should have received a positive transfer of 3.02% of GDP.[55]

Further, we can compute what the maximum grant would be under an extreme shock realization as a function of the previously accumulated debt, for a given revenue-to-GDP ratio. Assuming that revenues are 35% of GDP, under the worst case scenario (i.e for $\theta = \overline{\theta}$), member countries would be entitled to transfers between 3% and 4.5% of GDP depending on the level of previously accumulate debt. Having lower debt or higher revenues would increase the transfer's entity. For example, a country having a 90% debt-to-GDP ratio and a 35% revenue-to-GDP ratio would receive a grant amounting to 3.9% of GDP when hit with the worst possible shock realization, while a union member with debts for 150% of GDP would only be entitled to a 3.5% transfer. Using the credit-line equation in the implementation Section 3.2:

$$Credit\_Ratio_{i,t} = \frac{\gamma\widehat{\lambda}(\theta_{i,t} - \widehat{\beta})}{\gamma\theta_{i,t} + \widehat{\lambda}\widehat{\beta}}\widehat{\alpha}(\theta_{i,t})\left(Debt\_Ratio_{i,t} + \frac{Rev\_Ratio_{i,t}}{r}\right),$$

we find out that transfers should represent about 30% of the overall financial help (including the credit-line) under extreme financial distress.

For a quick comparison, consider that, under the European pandemic relief program Next Generation EU (NGEU), grants amount to 52% of the total available resources (750 billion Euros), with considerable cross-country variations. In Italy, which is one of the worst hit nations in the union, the percentage of grants is around 39% (209 billion Euros ca., of which 81.4 in grants and 127.4 in loans). Further, notice that the Recovery and Resilience Facility (part of NGEU) amounts to 672.5 billion, 70% of which will be distributed in the next two years. Back of the envelope calculations reveal that the planned yearly disbursement is around 1.7% of European GDP.

[55] Spending in Spain in 2012 was 41.5% of GDP (AMECO).

Finally, we can quantify welfare gains accruing to the EU from setting up a transfer system. More specifically, evaluating equation (11) with the previously estimated parameters and shock distribution, and data on the average revenue ratio for the union, I obtain that welfare gains are between 10% and 11% of European GDP.[56] This should be read as an upper bound on welfare gains, considering that the model only features a simplified asset market (non-contingent risk-free-bonds) with little availability of insurance against macroeconomic shocks.

5 Conclusion

The common rationale for having fiscal rules is that they are necessary to offset biases in fiscal policy. Yet, most rules comprise escape clauses, reflecting the fact that spending flexibility might be needed in adverse economic conditions to conduct counter-cyclical fiscal policies. The theory articulated in this paper formalizes this two key elements in a model with present biased governments having private information on the idiosyncratic state of the economy.

I have focused on the design of optimal fiscal rules at a supranational level in two distinct environments: one in which transfers across union members are not allowed and one in which they are. Optimal rules were found to be of the threshold kind in both environments, but weakly more stringent when transfers are allowed. In a fiscal union, debt-dependent transfers complement the set of rules. All instruments are debt-contingent: higher public debt contemporaneously tightens deficit limits and reduces financial assistance. One of the main policy implications of this paper, then, is that uniform, constant thresholds across countries, like the Maastricht 3% deficit limit, are sub-optimal. Fiscal constraints contingent on preexisting debt-levels, like some of the ones detailed in the more recent fiscal compact, are much closer to the derived optimal rule.

The described optimal rules can be implemented as simple deficit limits and complemented with a combination of grants and loans when cross-country subsidies are allowed. This paper details the optimal transfer system in a fiscal union, and shows under which conditions the introduction of transfers should be matched with a tightening of the fiscal rules. The recent creation of a European Recovery Fund in response to the Covid pandemic, has spurred new disagreements between EU members over the entity of financial assistance provided and over the relative proportion of grants and credits it should entail. This work supplies a useful benchmark to frame this now salient discussion.

[56] Variability is due to changes in the average revenue ratio for the union over the years, so welfare gains depend on when, exactly, transfers are assumed to be introduced. However, since the average revenue ratio is quite stable, welfare gains do not heavily depend on the year in which the transfer system is set-up.

References

Abrahám, A., E. Cárceles Poveda, Y. Liu, and R. Marimon (2018). On the optimal design of a financial stability fund.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2008). Political economy of mechanisms. *Econometrica 76*(3), 619–641.

Afonso, A. and D. Furceri (2008). Emu enlargement, stabilization costs and insurance mechanisms. *Journal of International Money and Finance 27*(2), 169–187.

Aguiar, M., M. Amador, E. Farhi, and G. Gopinath (2015). Coordination and crisis in monetary unions. *The Quarterly Journal of Economics 130*(4), 1727–1779.

Aizenman, J. (1998). Fiscal discipline in a union. *The political economy of reform*, 185–208.

Albanesi, S. and C. Sleet (2006). Dynamic optimal taxation with private information. *The Review of Economic Studies 73*(1), 1–30.

Alesina, A. and R. J. Barro (2002). Currency unions. *The Quarterly Journal of Economics 117*(2), 409–436.

Alesina, A. and G. Tabellini (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies 57*(3), 403–414.

Amador, M., I. Werning, and G.-M. Angeletos (2006). Commitment vs. flexibility. *Econometrica 74*(2), 365–396.

Asdrubali, P., B. E. Sørensen, and O. Yosha (1996). Channels of interstate risk sharing: United states 1963–1990. *The Quarterly Journal of Economics 111*(4), 1081–1110.

Atkeson, A. and T. Bayoumi (1993). Do private capital markets insure regional risk? evidence from the united states and europe. *Open economies review 4*(3), 303–324.

Atkeson, A. and J. Lucas, Robert E. (1992). On efficient distribution with private information. *The Review of Economic Studies 59*(3), 427–453.

Azzimonti, M. (2011). Barriers to investment in polarized societies. *American Economic Review 101*(5), 2182–2204.

Azzimonti, M. (2018). Partisan conflict and private investment. *Journal of Monetary Economics 93*, 114–131.

Azzimonti, M., M. Battaglini, and S. Coate (2016). The costs and benefits of balanced budget rules: Lessons from a political economy model of fiscal policy. *Journal of Public Economics 136*, 45 – 61.

Azzimonti, M., E. De Francisco, and V. Quadrini (2014). Financial globalization, inequality, and the rising public debt. *American Economic Review 104*(8), 2267–2302.

Barro, R. J. (1999). Ramsey meets laibson in the neoclassical growth model. *The Quarterly Journal of Economics 114*(4), 1125–1152.

Battaglini, M. and S. Coate (2007). Inefficiency in legislative policymaking: a dynamic analysis. *American Economic Review 97*(1), 118–149.

Battaglini, M. and S. Coate (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review 98*(1), 201–36.

Baum, A., L. Eyraud, A. Hodge, M. Jarmuzek, Y. Kim, S. Mbaye, and E. Ture (2018). How to calibrate fiscal rules: a primer.

Beetsma, R. and H. Uhlig (1999). An analysis of the stability and growth pact. *The Economic Journal 109*(458), 546–571.

Bisin, A., A. Lizzeri, and L. Yariv (2015). Government policy with time inconsistent voters. *American Economic Review 105*(6), 1711–37.

Bolton, P., M. Dewatripont, et al. (2005). *Contract theory.* MIT press.

Bucovetsky, S. (1998). Federalism, equalization and risk aversion. *Journal of Public Economics 67*(3), 301–328.

Canova, F. and M. O. Ravn (1996). International consumption risk sharing. *International Economic Review*, 573–601.

Chari, V. V. and P. J. Kehoe (1990). International coordination of fiscal policy in limiting economies. *Journal of Political Economy 98*(3), 617–636.

Chari, V. V. and P. J. Kehoe (2007). On the need for fiscal constraints in a monetary union. *Journal of Monetary Economics 54*(8), 2399–2408.

Cooley, T. F. and V. Quadrini (2003). Common currencies vs. monetary independence. *The review of economic studies 70*(4), 785–806.

Cooper, R. and H. Kempf (2004). Overturning mundell: Fiscal policy in a monetary union. *The Review of Economic Studies 71*(2), 371–396.

Crivelli, E., V. Gaspar, S. Gupta, and C. Mulas-Granados (2015). Fragmented politics and public debt. In *Fiscal Politics.*

Cukierman, A. and A. H. Meltzer (1989). A political theory of government debt and deficits in a neo-ricardian framework. *The American Economic Review*, 713–732.

Dixit, A. and L. Lambertini (2001). Monetary–fiscal policy interactions and commitment versus discretion in a monetary union. *European Economic Review 45*(4-6), 977–987.

Dovis, A. and R. Kirpalani (2020). Fiscal rules, bailouts, and reputation in federal governments. *American Economic Review 110*(3), 860–88.

Evers, M. P. (2012). Federal fiscal transfer rules in monetary unions. *European Economic Review 56*(3), 507–525.

Eyraud, L., V. Lledó, P. Dudine, and A. Peralta (2018). How to select fiscal rules: A primer. imf series of fiscal affairs department how-to notes, 9, 1–24.

Farhi, E. and I. Werning (2007). Inequality and social discounting. *Journal of political economy 115*(3), 365–402.

Farhi, E. and I. Werning (2017). Fiscal unions. *American Economic Review 107*(12), 3788–3834.

Ferrari, A., R. Marimon, and C. Simpson-Bell (2020). Fiscal and currency union with default and exit. *Available at SSRN 3673917*.

Green, E. J. (1987). Lending and the smoothing of uninsurable income. *Contractual arrangements for intertemporal trade 1*, 3–25.

Halac, M. and P. Yared (2014). Fiscal rules and discretion under persistent shocks. *Econometrica 82*(5), 1557–1614.

Halac, M. and P. Yared (2018). Fiscal rules and discretion in a world economy. *American Economic Review 108*(8), 2305–34.

Harris, C. and D. Laibson (2012). Instantaneous gratification. *The Quarterly Journal of Economics 128*(1), 205–248.

Jackson, M. O. and L. Yariv (2014). Present bias and collective dynamic choice in the lab. *American Economic Review 104*(12), 4184–4204.

Jackson, M. O. and L. Yariv (2015). Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics 7*(4), 150–78.

Krusell, P., B. Kuruşçu, and A. A. Smith Jr (2010). Temptation and taxation. *Econometrica 78*(6), 2063–2084.

Krusell, P. and J.-V. Rios-Rull (1999). On the size of us government: political economy in the neoclassical growth model. *American Economic Review 89*(5), 1156–1181.

Krusell, P. and A. A. Smith (2003). Consumption-savings decisions with quasi-geometric discounting. *Econometrica 71*(1), 365–375.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Lizzeri, A. and L. Yariv (2017). Collective self-control. *American Economic Journal: Microeconomics 9*(3), 213–44.

Lledó, V., S. Yoon, X. Fang, S. Mbaye, and Y. Kim (2017). Fiscal rules at a glance, international monetary fund. *Erişim 26*, 2–77.

Lockwood, B. (1999). Inter-regional insurance. *Journal of Public Economics 72*(1), 1–37.

Mbaye, S., M. M. M. Badia, and K. Chae (2018). *Global debt database: Methodology and sources*. International Monetary Fund.

Mélitz, J. and F. Zumer (2002). Regional redistribution and stabilization by the center in canada, france, the uk and the us:: A reassessment and new tests. *Journal of public Economics 86*(2), 263–286.

Persson, T. and L. E. Svensson (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *The Quarterly Journal of Economics 104*(2), 325–345.

Persson, T. and G. Tabellini (1995). Double-edged incentives: Institutions and policy coordination. *Handbook of international economics 3*, 1973–2030.

Persson, T. and G. Tabellini (1996a). Federal fiscal constitutions: Risk sharing and moral hazard. *Econometrica: Journal of the Econometric Society*, 623–646.

Persson, T. and G. Tabellini (1996b). Federal fiscal constitutions: risk sharing and redistribution. *Journal of political Economy 104*(5), 979–1009.

Persson, T. and G. Tabellini (2004). Constitutional rules and fiscal policy outcomes. *American Economic Review 94*(1), 25–45.

Phelps, E. S. and R. A. Pollak (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies 35*(2), 185–199.

Roubini, N. and J. D. Sachs (1989). Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review 33*(5), 903–933.

Schaechter, A., T. Kinda, N. T. Budina, and A. Weber (2012). Fiscal rules in response to the crisis-toward the'next-generation'rules: A new dataset.

Sibert, A. (1992). Government finance in a common currency area. *Journal of International Money and Finance 11*(6), 567–578.

Sleet, C. and Ş. Yeltekin (2006). Credibility and endogenous societal discounting. *Review of Economic Dynamics 9*(3), 410–437.

Song, Z., K. Storesletten, and F. Zilibotti (2012). Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt. *Econometrica 80*(6), 2785–2803.

Tabellini, G. (1991). The politics of intergenerational redistribution. *Journal of Political Economy 99*(2), 335–357.

Von Hagen, J. and B. Eichengreen (1996). Federalism, fiscal restraints, and european monetary union. *The American Economic Review 86*(2), 134–138.

Weingast, B. R., K. A. Shepsle, and C. Johnsen (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. *Journal of Political Economy 89*(4), 642–664.

Woo, J. (2003). Economic, political, and institutional determinants of public deficits. *Journal of Public Economics 87*(3), 387 – 426.

Yared, P. (2010). Politicians, taxes and debt. *The Review of Economic Studies 77*(2), 806–840.

Yared, P. (2019). Rising government debt: Causes and solutions for a decades-old trend. *Journal of Economic Perspectives 33*(2), 115–40.

A Proofs for Section 2.1

A.1 Recursive Formulation

I begin by rewriting the sequential problem ($\mathcal{P}_{NT}$) recursively. Remember that, given an initial level of wealth $\bar{a}$, $\mathcal{V}_0(\bar{a})$ is the set of planner's payoffs such that, for all $v_0 \in \mathcal{V}_0(\bar{a})$, there exists a sequence of spending g and an associated wealth process that *(i)* satisfy the government's budget constraint (BC) with initial assets $\bar{a}$, *(ii)* are such that truthful reporting is incentive compatible (i.e. constraint (IC) is satisfied) and *(iii)* deliver utility $v_0 = V(g, \sigma^*)$. Also, v_n is the continuation utility of government n at the time of its formation τ_n.

Now, take a sequence of spending g and associated wealth process $\{a_t\}$, with $a_0 = \kappa y/r$, satisfying incentive compatibility (1), the budget constraint (BC) and delivering utility $v_0 = V(g, \sigma^*)$. Standard properties of logarithmic preferences imply that the sequence $\widehat{g} \equiv g/a_0$ and the associated wealth process $\widehat{a}_t \equiv a_t/a_0$, with $\widehat{a}_0 = 1$, satisfy incentive compatibility (1), the budget constraint (BC) and deliver utility to the planner

$$\begin{aligned}
\widehat{v}_0 &= V(\widehat{g}, \sigma^*) \\
&= \mathbb{E}_- \left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n} \theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)} \log(\widehat{g}_t) dt \right] \\
&= v_0 - \mathbb{E}_- \left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n} \theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)} \log(a_0) dt \right] \\
&= v_0 - \frac{1}{\gamma} \log(a_0).
\end{aligned}$$

Exactly the same arguments show that, for any value $v_n \in \mathcal{V}_n(\bar{a})$, there is a corresponding value in $\widehat{v}_n \in \mathcal{V}_n(1)$ such that $\widehat{v}_n = v_n - \log(\bar{a})/\gamma$. As a result, it is sufficient to characterize the problem for $a_0 = 1$.

Consider now a sequence g with associated wealth process $\{a_t\}$, $a_{\tau_n} = 1$, delivering utility $v_n \in \mathcal{V}_n(1)$ to the planner. Let a_{τ_n} the country's wealth at the time the n-th government is formed. Using the law of iterated expectations and the above results, the expected utility of government n at time τ_n is

$$\begin{aligned}
U_{\tau_n}(\sigma_n^*|g, \sigma_{-n}^*) &= \mathbb{E}_{\tau_n} \left[\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(s-t)} \theta_n \log(g_s) ds + \beta \sum_{j=n+1}^{\infty} e^{-\gamma\tau_j} \left(\int_{\tau_j}^{\tau_{j+1}} e^{-\gamma(s-\tau_j)} \theta_j \log(g_s) ds \right) \right] \qquad (17) \\
&= \mathbb{E}_{\tau_n} \left[\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(s-t)} \theta_n \log(g_s) ds + \beta e^{-\gamma\tau_{n+1}} \mathbb{E}_{\tau_{n+1}^-} \sum_{j=n+1}^{\infty} e^{-\gamma(\tau_j - \tau_{n+1})} \left(\int_{\tau_j}^{\tau_{j+1}} e^{-\gamma(s-\tau_j)} \theta_j \log(g_s) ds \right) \right] \\
&= \mathbb{E}_{\tau_n} \left[\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(s-t)} \theta_n \log(g_s) ds + \beta e^{-\gamma\tau_{n+1}} v_{n+1} \right] \\
&= \mathbb{E}_{\tau_n} \left[\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(s-t)} \theta_n \log(g_s) ds + \beta e^{-\gamma\tau_{n+1}} \left(\widehat{v}_{n+1} + \frac{1}{\gamma} \log\left(a_{\tau_{n+1}}\right) \right) \right],
\end{aligned}$$

where $\widehat{v}_{n+1}$ in the last equality is such that $\widehat{v}_{n+1} \in \mathcal{V}_{n+1}(1)$. By the same arguments, planner's utility equals

$$v_n = \mathbb{E}_{\tau_n^-}\left[\int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(s-t)}\theta_n \log(g_s)ds + e^{-\gamma\tau_{n+1}}\widehat{v}_{n+1} + e^{-\gamma\tau_{n+1}}\frac{1}{\gamma}\log\left(a_{\tau_{n+1}}\right)\right].$$

I now consider the planner's problem at time 0, when the first government is formed (the other formation periods τ_n, $n = 1, 2...$, are analogous). By definition, the value of such a problem – which I denoted with v^{nt} – corresponds to the point in $\mathcal{V}_0(a_0)$ such that $v^{nt} \geq v_0$, for all $v_0 \in \mathcal{V}_0(a_0)$. In addition, the arguments above imply that there exists a point in $\mathcal{V}_0(1)$ – which I denoted with $\overline{v}$ – such that $\overline{v} = v^{nt} - \log(a_0)/\gamma$ and $\overline{v} \in \mathcal{V}_0(1)$. Since $\overline{v} \in \mathcal{V}_0(1)$, it must be that

$$\overline{v} = \max_{g,a,w\in\mathcal{V}(1)} \mathbb{E}_-\left[\int_0^{\tau_1} e^{-\gamma t}\theta_0 \log(g_t)dt + e^{-\gamma\tau_1}w(\theta) + e^{-\gamma\tau_1}\frac{1}{\gamma}\log\left(a_{\tau_1}\right)\right],$$

subject to the incentive compatibility constraint (IC), which, using (17), can be equivalently expressed as

$$\begin{aligned}&\mathbb{E}\left[\int_0^{\tau_1} e^{-\gamma t}\theta_0 \log(g_t)dt + \beta e^{-\gamma\tau_1}w(\theta_0) + \beta\frac{1}{\gamma}e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\right] \\ \geq&\mathbb{E}\left[\int_0^{\tau_1} e^{-\gamma t}\theta_0 \log(g_t)dt\middle|\,\tilde{\theta}\right] + \beta w(\tilde{\theta})\mathbb{E}\left[e^{-\gamma\tau_1}\right] + \beta\frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\middle|\,\tilde{\theta}\right]\end{aligned} \tag{18}$$

and the budget constraint (BC) with initial wealth equal to 1. The latter can be equivalently written as

$$\int_0^{\tau} e^{-rt}g_t dt + e^{-r\tau}a_\tau = 1,$$

which is (4). Finally, to obtain problem ($\mathcal{P}_{NT:Rec}$), I add the constraint (5), together with the extra choice variable $u : \Theta \to \mathbb{R}$, and use it to rewrite (18) as

$$\mathbb{E}\left[\theta_0 u(\theta_0) + \beta e^{-\gamma\tau_1}w(\theta_0) + \beta\frac{1}{\gamma}e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\right] \geq \theta_0 u(\tilde{\theta}) + \beta w(\tilde{\theta})\mathbb{E}\left[e^{-\gamma\tau_1}\right] + \beta\frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\middle|\,\tilde{\theta}\right]$$

or, using $\mathbb{E}[e^{-\gamma\tau_1}] = \lambda/(\gamma+\lambda)$,

$$\mathbb{E}\left[\theta_0 u(\theta_0) + \beta\frac{\lambda}{\gamma+\lambda}w(\theta_0) + \beta\frac{1}{\gamma}e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\right] \geq \theta_0 u(\tilde{\theta}) + \beta\frac{\lambda}{\gamma+\lambda}w(\tilde{\theta}) + \beta\frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma\tau_1}\log\left(a_{\tau_1}\right)\middle|\,\tilde{\theta}\right],$$

which is equivalent to (5).

A.2 Proof of Lemma 1

Consider the subproblem ($\mathcal{P}_{NT:S1}$). Integrating by parts, I can rewrite the constraint on current utility as follows:

$$\begin{aligned}\overline{u} &= \mathbb{E}\left[\int_0^{\tau} e^{-\gamma t}\log(\widehat{g}_t)dt\right] \\ &= \int_0^{\infty} \lambda e^{-\lambda t}\int_0^{t} e^{-\gamma s}\log(\widehat{g}_t)dsdt \\ &= \int_0^{\infty} e^{-(\lambda+\gamma)t}\log(\widehat{g}_t)dt.\end{aligned}$$

The Lagrangian associated to the problem is then

$$\mathcal{L}=\int_0^\infty \lambda e^{-\lambda t}\log(\widehat{g}_t)dt-\int_0^\infty\left(\int_0^t e^{-r\cdot s}\widehat{g}_s ds+e^{-rt}\widehat{a}_t-1\right)d\Phi_t+\mu\int_0^\infty e^{-(\lambda+\gamma)t}\log(\widehat{g}_t)dt,$$

where μ and Φ are Lagrange multipliers on, respectively, the current-utility constraint above and the budget constraint (4). Integrating by parts the second term,

$$\mathcal{L}=\frac{1}{\gamma}\int_0^\infty \lambda e^{-(\lambda+\gamma)t}\log(\widehat{a}_t)dt-\int_0^\infty (e^{-rt}\widehat{a}_t-1)d\Phi_t-\left(\Phi_t\int_0^t e^{-r\cdot s}\widehat{g}_s ds\Big|_0^\infty-\int_0^\infty \Phi_t e^{-r\cdot t}\widehat{g}_t dt\right)+\mu\int_0^\infty e^{-(\lambda+\gamma)t}\log(\widehat{g}_t)dt$$

or

$$\mathcal{L}=\frac{1}{\gamma}\int_0^\infty \lambda e^{-(\lambda+\gamma)t}\log(\widehat{a}_t)dt-\int_0^\infty (e^{-rt}\widehat{a}_t-1)d\Phi_t-\int_0^\infty (\Phi_\infty-\Phi_t)e^{-r\cdot t}\widehat{g}_t dt+\mu\int_0^\infty e^{-(\lambda+\gamma)t}\log(\widehat{g}_t)dt,$$

where $\Phi_\infty\equiv\lim_{t\to\infty}\Phi_t$. The first-order conditions with respect to $\widehat{a}_t$ and $\widehat{g}_t$ are, respectively,

$$\frac{1}{\gamma}\lambda e^{-(\lambda+\gamma)t}\frac{1}{\widehat{a}_t}-\dot{\Phi}_t e^{-rt}=0$$

and

$$\mu e^{-(\lambda+\gamma)t}\frac{1}{\widehat{g}_t}-(\Phi_\infty-\Phi_t)e^{-r\cdot t}=0.$$

Conjecture $\widehat{a}_t=e^{(r-\Delta)t}$ and $\widehat{g}_t=\Gamma e^{(r-\Delta)t}$, for some positive scalars Δ, Γ, with $\Delta>\lambda+\gamma$. Then, the first condition yields

$$\frac{1}{\gamma}\lambda e^{-(r+\lambda+\gamma-\Delta)t}-\dot{\Phi}_t e^{-rt}=0$$

and, thus,

$$\Phi_t=\frac{1}{\gamma}\lambda\frac{1}{\Delta-(\lambda+\gamma)}e^{-(\lambda+\gamma-\Delta)t}.$$

Also, since $\Delta>\lambda+\gamma$, $\Phi_\infty=0$, and the second condition becomes

$$\mu e^{-(\lambda+\gamma)t}\frac{1}{\Gamma}e^{-(r-\Delta)t}+\frac{1}{\gamma}\lambda\frac{1}{\Delta-(\lambda+\gamma)}e^{-(\lambda+\gamma-\Delta)t}e^{-r\cdot t}=0$$

or

$$\mu\frac{1}{\Gamma}+\frac{1}{\gamma}\lambda\frac{1}{\Delta-(\lambda+\gamma)}=0. \tag{19}$$

Also, $\{\widehat{a}_t,\widehat{g}_t\}_{t\geq 0}$ must satisfy the budget constraint (4), which given the conjectures above is

$$\Gamma\int_0^\tau e^{-\Delta t}dt+e^{-\Delta\tau}=1.$$

The latter is true for all possible realizations of τ if and only if $\Gamma=\Delta$. Combining the latter result with (19) yields

$$\Gamma=\frac{\mu(\lambda+\gamma)}{\frac{1}{\gamma}\lambda+\mu}.$$

The Lagrange multiplier μ must be chosen to satisfy the current-utility constraint:

$$\begin{aligned}\overline{u} &= \int_0^\infty e^{-(\gamma+\lambda)t}\log(\widehat{g}_t)dt \\ &= \log(\Gamma)\int_0^\infty e^{-(\gamma+\lambda)t}dt + (r-\Gamma)\int_0^\infty e^{-(\gamma+\lambda)t}tdt \\ &= \log\left(\frac{\mu(\lambda+\gamma)}{\frac{1}{\gamma}\lambda+\mu}\right)\frac{1}{\gamma+\lambda} + \left(r - \frac{\mu(\lambda+\gamma)}{\frac{1}{\gamma}\lambda+\mu}\right)\frac{1}{(\gamma+\lambda)^2}.\end{aligned}$$

Thus, the conjecture is verified.

Finally, by denoting with $k(\overline{u})$ the solution to

$$\overline{u} = \log(k(\overline{u}))\frac{1}{\gamma+\lambda} + (r-k(\overline{u}))\frac{1}{(\gamma+\lambda)^2},$$

the optimal levels of spending and wealth are, respectively, $\widehat{g}_t = k(\overline{u})e^{(r-k(\overline{u}))t}$ and $\widehat{a}_t = e^{(r-k(\overline{u}))t}$.

A.3 Proof of Lemma 2

The proof follows from the arguments in Proposition 2 in Amador et al. (2006). In particular, notice that the objective function can be written as

$$\frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}(1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w} = \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}(1-M(\theta))u(\theta)d\theta + \frac{1}{\beta}\int_{\theta^\star}^{\overline{\theta}}(1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w}.$$

Also, using $u(\theta) = \int_{\theta^\star}^{\theta} du + u(\theta^\star)$, for $\theta \geq \theta^\star$, after integrating by parts, the second term becomes

$$\frac{1}{\beta}\int_{\theta^\star}^{\overline{\theta}}\int_{\theta}^{\overline{\theta}}(1-M(x))dxdu + \frac{1}{\beta}\int_{\theta^\star}^{\overline{\theta}}(1-M(\theta))u(\theta^\star)d\theta.$$

Since u is non-decreasing it must be that $du \geq 0$. Also, since $\int_{\theta}^{\overline{\theta}}(1-M(x))dx \leq 0$ for all $\theta \geq \theta^\star$, the term above is maximized for $du = 0$ or, equivalently, for $u(\theta) = u(\theta^\star)$ for all $\theta \geq \theta^\star$. Finally, notice that bunching types in the upper tail is always incentive compatible since the incentive constraint

$$\frac{\theta}{\beta}u(\theta) + W(u(\theta)) \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta}u(z)dz + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w}$$

is satisfied for all $\theta > \theta^\star$ if it is satisfied for $\theta \leq \theta^\star$.

A.4 Proof of Proposition 1

Consider the Lagrangian associated to problem ($\mathcal{P}_{NT:S2}$):

$$\mathcal{L} = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}(1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w} + \int_{\underline{\theta}}^{\overline{\theta}}\left(\frac{1}{\beta}\theta u(\theta) + W(u(\theta)) - \int_{\underline{\theta}}^{\theta}\frac{1}{\beta}u(x)dx - \frac{\underline{\theta}}{\beta}u(\underline{\theta}) - \underline{w}\right)d\Lambda(\theta)$$

for some non-decreasing function $\Lambda(\cdot)$. Integrating by parts yields

$$\mathcal{L} = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}(\Lambda(\theta)-M(\theta))u(\theta)d\theta + \left(\frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \underline{w}\right)\Lambda(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}}\left(\frac{1}{\beta}\theta u(\theta) + W(u(\theta))\right)d\Lambda(\theta).$$

Following Amador et al. (2006), I set $\Lambda(\underline{\theta}) = 0$, $\Lambda(\theta) = M(\theta)$, for $\theta \leq \theta^\star$, and $\Lambda(\theta) = 1$, for $\theta > \theta^\star$. Notice that, by Assumption 1, $\Lambda(\cdot)$ is non-decreasing, as required. The arguments in Amador et al. (2006) then imply that the value of u that maximizes the Lagrangian with this particular choice of $\Lambda(\cdot)$ is the solution to problem ($\mathcal{P}_{NT:S2}$). In particular, the first-order condition with respect to $u(\theta)$, for $\theta \leq \theta^\star$, gives

$$\frac{1}{\beta}\theta + W'(u(\theta)) = 0.$$

Since $W(x) = (r-k(x))\mathbb{E}[e^{-\gamma\tau}\tau]/\gamma + \mathbb{E}[e^{-\gamma\tau}]\overline{w}$ and $\mathbb{E}[e^{-\gamma\tau}\tau] = \lambda/(\gamma+\lambda)^2$, then $W'(x) = -k'(x)\lambda/[\gamma(\gamma+\lambda)^2]$ and the condition above implies

$$\frac{1}{\beta}\theta = \frac{\lambda k'(u(\theta))}{\gamma(\gamma+\lambda)^2}. \tag{20}$$

Also, from Lemma 1, $k(x)$ is the solution to the equation

$$(\gamma+\lambda)x = \log(k(x)) + (r - k(x))\frac{1}{\gamma+\lambda}.$$

Differentiating both sides gives $-k'(x)/(\gamma+\lambda) = \gamma+\lambda - k'(x)/k(x)$ which, combined with (20), yields

$$-\frac{1}{\beta}\theta\frac{\gamma}{\lambda} = 1 - \frac{\frac{1}{\beta}\theta\gamma(\gamma+\lambda)}{\lambda k(u(\theta))}$$

or

$$k(u(\theta)) = \frac{\gamma\theta(\gamma+\lambda)}{\gamma\theta+\lambda\beta} \equiv k^{nt}(\theta).$$

The statement then follows directly from Lemma 1 with initial wealth $a_0 = \kappa y/r$.

B Proofs for Section 2.2

I begin by rewriting the sequential problem ($\mathcal{P}_{TR}$) recursively. It is equivalent, but more convenient to work with the dual problem of minimizing expected resources of delivering a given lifetime utility. Formally, let $K(v_0)$ be the minimum amount of resources necessary to delivery utility v_0:

$$K(v_0) \equiv \min_g \mathbb{E}_- \left[\int_0^\infty e^{-rt} g_t dt\right], \tag{21}$$

subject to (IC) and

$$v_0 = \mathbb{E}_- \left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n}\theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)}\log(g_t)dt\right].$$

The value of ($\mathcal{P}_{TR}$) (i.e. v^{tr}) is then given by $K(v^{tr}) = \kappa y/r$.

I write the recursive version of (21) at time 0, that is, the time at which the first government is formed (the other cases are analogous). Take a sequence g and let $v_1(\theta_0)$ be the associated continuation value at time τ_1 (i.e. the time at which the next government is formed), as a function of the current government's type. It is equal to

$$v_1(\theta_0) = \mathbb{E}_{\tau_1^-} \left[\sum_{n=1}^{\infty} e^{-\gamma(\tau_n-\tau_1)}\theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)}\log(g_t)dt\right]. \tag{22}$$

As a result,

$$v_0 = \mathbb{E}_- \left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n} \theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)} \log(g_t) dt \right] \tag{23}$$
$$= \mathbb{E}_- \left[\theta_0 \int_0^{\tau_1} e^{-\gamma t} \log(g_t) dt + e^{-\gamma\tau_1} v_1(\theta_0) \right]$$
$$= \mathbb{E}_- \left[\theta_0 \int_0^{\tau_1} e^{-\gamma t} \log(g_t) dt + \frac{\lambda}{\gamma+\lambda} v_1(\theta_0) \right],$$

where the last line uses $\mathbb{E}_- \left[e^{-\gamma\tau_1} v_1(\theta_0) \right] = \mathbb{E}_- \left[e^{-\gamma\tau_1} \mathbb{E}_{\tau_1^-} \left[v_1(\theta_0) \right] \right] = \lambda \mathbb{E}_-[v_1(\theta_0)]/(\gamma+\lambda)$. The same arguments imply that the incentive constraint (IC) at time 0 can be rewritten as

$$\mathbb{E} \left[\int_0^{\tau_1} e^{-\gamma t} \theta_0 \log(g_t) dt \right] + \beta \frac{\lambda}{\gamma+\lambda} v_1(\theta_0) \geq \mathbb{E} \left[\int_0^{\tau_1} e^{-\gamma t} \theta_0 \log(g_t) dt \middle| \tilde{\theta} \right] + \beta \frac{\lambda}{\gamma+\lambda} v_1(\tilde{\theta}). \tag{24}$$

Therefore, problem (21) can be equivalently stated as

$$K(v) \equiv \min_g \mathbb{E}_- \left[\int_0^{\infty} e^{-rt} g_t dt \right],$$

subject to (24), the incentive constraint (IC) from time τ_1 onward, (23), and (22). In addition, the value function $K(v)$ then satisfies the following property:

$$K(v) \equiv \min_g \mathbb{E}_- \left[\int_0^{\infty} e^{-rt} g_t dt \right]$$
$$= \min_g \mathbb{E}_- \left[\int_0^{\tau_1} e^{-rt} g_t dt + e^{-r\tau_1} \mathbb{E}_{\tau_1^-} \left[\sum_{n=1}^{\infty} e^{-r(\tau_n-\tau_1)} \int_{\tau_n}^{\tau_{n+1}} e^{-r(t-\tau_n)} g_t dt \right] \right]$$
$$= \min_{\{g_t\}_0^{\tau_1}} \mathbb{E}_- \left[\int_0^{\tau_1} e^{-rt} g_t dt + \frac{\lambda}{r+\lambda} \min_{g|_{\tau_1}} \mathbb{E}_{\tau_1^-} \left[\sum_{n=1}^{\infty} e^{-r(\tau_n-\tau_1)} \int_{\tau_n}^{\tau_{n+1}} e^{-r(t-\tau_n)} g_t dt \right] \right],$$

subject to (24), the (IC) at $t \geq \tau_1$, (23), and (22), where $g|_{\tau_1}$ is a short-hand notation for the sequence of spending starting from time τ_1. Since constraints (24) and (23) depend only on spending until time τ_1, the latter is equivalent to

$$K(v) = \min_{\{g_t\}_0^{\tau_1}} \mathbb{E}_- \left[\int_0^{\tau_1} e^{-rt} g_t dt + \frac{\lambda}{r+\lambda} K(v_1(\theta_0)) \right],$$

subject to (24) and (23).

Finally, to obtain problem ($\mathcal{P}_{TR:Rec}$), I add the constraint (10), together with the extra choice variable $u : \Theta \to \mathbb{R}$, and use it to rewrite (23) as (9) and (24) as

$$\theta_0 u(\theta_0) + \beta \frac{\lambda}{\gamma+\lambda} v_1(\theta_0) \geq \theta_0 u(\tilde{\theta}) + \beta \frac{\lambda}{\gamma+\lambda} v_1(\tilde{\theta}),$$

which is equivalent to (8).

B.1 Proof of Lemma 3

The Lagrangian associated to problem ($\mathcal{P}_{TR:S1}$) is

$$\mathcal{L} = \int_0^{\infty} \lambda e^{-(\lambda+r)t} \widehat{g}_t dt - \mu(\gamma+\lambda) \int_0^{\infty} e^{-(\lambda+\gamma)t} \log(\widehat{g}_t) dt,$$

for some Lagrange multiplier μ. The first-order condition with respect to g_t is

$$\lambda e^{-(\lambda+r)t} - \mu(\gamma+\lambda)e^{-(\lambda+\gamma)t}\frac{1}{\widehat{g}_t} = 0.$$

As a result,

$$\widehat{g}_t = \mu(\gamma+\lambda)e^{(r-\gamma)t}. \tag{25}$$

Finally, the multiplier μ is chosen so as to satisfy the current-utility constraint. Condition (25) shows that the planner minimizes the resources to deliver a given utility level $\overline{u}$ by allocating government spending which is increasing or decreasing depending on whether $r > \gamma$ or $r < \gamma$. In the special case that $r = \gamma$, condition (25) becomes

$$\widehat{g}_t = \mu(\gamma+\lambda).$$

where μ satisfies

$$\overline{u} = (\gamma+\lambda)\log(\mu(\gamma+\lambda))\mathbb{E}\left[\int_0^\tau e^{-\gamma t}dt\right].$$

Properties of the Poisson process imply $\mathbb{E}\left[\int_0^\tau e^{-\gamma t}dt\right] = 1/(\gamma+\lambda)$, thus,

$$\mu = \frac{1}{\gamma+\lambda}e^{\overline{u}}.$$

As a result, the optimal amount of resources necessary to deliver utility $\overline{u}$ is

$$\begin{aligned} G(\overline{u}) &\equiv \min_g \mathbb{E}\left[\int_0^\tau e^{-\gamma t}\widehat{g}_t dt\right] \\ &= \frac{1}{\gamma+\lambda}e^{\overline{u}}. \end{aligned}$$

Derivation of problem ($\mathcal{P}'_{TR:S2}$). As discussed in the main text, I conjecture that the value function $K(\cdot)$ satisfies $K(v) = K(0)\exp(Dv)$, for some scalar D. With this conjecture, we obtain

$$K(0)\exp(Dv) = \min_{u,w}\mathbb{E}_-\left[\frac{1}{\gamma+\lambda}\exp(u) + \frac{\lambda}{\gamma+\lambda}K(0)\exp(Dw(\theta))\right],$$

subject to (8) and

$$(\gamma+\lambda)v = \mathbb{E}_-\left[\theta u(\theta) + \lambda w(\theta)\right].$$

Consider the change of variables: $\widetilde{u}(\theta) = u(\theta) - \gamma v$, $\widetilde{w}(\theta) = w(\theta) - v$, for some scalars A, B. Notice that the incentive constraint is not affected by this change of variables. As a result,

$$K(0)\exp(Dv) = \min_{\widetilde{u},\widetilde{w}}\mathbb{E}_-\left[\frac{1}{\gamma+\lambda}\exp(\widetilde{u}(\theta)+\gamma v) + \frac{\lambda}{\gamma+\lambda}K(0)\exp(D(\widetilde{w}(\theta)+v))\right],$$

subject to (8) and $0 = \mathbb{E}_-\left[\theta\widetilde{u}(\theta) + \lambda\widetilde{w}(\theta)\right]$. The conjecture is therefore verified by letting $D = \gamma$.

Finally, the same steps as those for the case with transfers imply that the recursive problem can be equivalently rewritten as ($\mathcal{P}'_{TR:S2}$).

B.2 Proof of Lemma 4

The proof is analogous to the one of Lemma 2 and follows the arguments of Proposition 2 in Amador et al. (2006).

B.3 Proof of Proposition 2 & 3

To prove the proposition, I consider the relaxed problem obtained by dropping the monotonicity constraint on u. I then solve the resulting problem in two steps. In the first step, I find optimal utility given transfers. Let $P(\cdot)$ the value of this problem:

$$P(T) \equiv \max_{u,U(\underline{\theta})} \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} (1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda \underline{w}, \tag{26}$$

subject to

$$\frac{\theta}{\beta}u(\theta) + \lambda W\left(K(0) - G(u(\theta)) + T(\theta)\right) \geq \frac{1}{\beta}\int_{\underline{\theta}}^{\theta} u(x)dx + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda \underline{w}.$$

In the second step, I find optimal transfers:

$$\max_{T} P(T), \tag{27}$$

subject to $\mathbb{E}[T(\theta)] \leq 0$.

Consider the Lagrangian associated to problem (26):

$$\begin{aligned}\mathcal{L} =& \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} (1-M(\theta))u(\theta)d\theta + \frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda \underline{w} \\ &+ \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\beta}\theta u(\theta) + \lambda W\left(K(0) - G(u(\theta)) + T(\theta)\right) - \int_{\underline{\theta}}^{\theta} \frac{1}{\beta}u(x)dx - \frac{\underline{\theta}}{\beta}u(\underline{\theta}) - \lambda \underline{w}\right) d\Lambda(\theta)\end{aligned}$$

for some non-decreasing function $\Lambda(\cdot)$. Integrating by parts yields

$$\mathcal{L} = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} (\Lambda(\theta) - M(\theta))u(\theta)d\theta + \left(\frac{\underline{\theta}}{\beta}u(\underline{\theta}) + \lambda \underline{w}\right)\Lambda(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\beta}\theta u(\theta) + \lambda W\left(K(0) - G(u(\theta)) + T(\theta)\right)\right) d\Lambda(\theta).$$

Follows the arguments in Amador et al. (2006), we set $\Lambda(\underline{\theta}) = 0$, $\Lambda(\theta) = M(\theta)$, for for $\theta \leq \theta^\star$ and $\Lambda(\theta) = 1$, for $\theta > \theta^\star$. In particular, the first-order condition with respect to $u(\theta)$, for $\theta \leq \theta^\star$, gives

$$\frac{1}{\beta}\theta - \lambda G'(u(\theta))W'\left(K(0) - G(u(\theta)) + T(\theta)\right) = 0.$$

Since $G(u) = \exp(u)(\gamma+\lambda)^{-1}$ and $W(x) = \log\left((\gamma+\lambda)x/\lambda K(0)\right)/\gamma$, the latter gives

$$u(\theta) = \log\left(k^{nt}(\theta)(K(0) + T(\theta))\right) \equiv U(T(\theta), \theta).$$

For $\theta > \theta^\star$, current utility and transfers must be constant by Lemma 4. As a result, $u(\theta) = U(T(\theta^\star), \theta^\star)$. Clearly, it must be that $T(\theta) > -K(0)$, for all $\theta \in \Theta$.

I now turn to problem (27). Using the results just derived, the objective function becomes

$$\begin{aligned}P(T) &= \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}(1-M(\theta))u(\theta)d\theta \\ &= \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}(1-M(\theta))u(\theta)d\theta + \frac{1}{\beta}u(\theta^\star)\int_{\theta^\star}^{\overline{\theta}}(1-M(\theta))d\theta.\end{aligned}$$

By definition of $\theta^\star$ the last integral is zero, thus,

$$P(T) = \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}(1-M(\theta))u(\theta)d\theta$$

or, integrating by parts,

$$\begin{aligned}P(T) &= \frac{1}{\beta}(1-M(\theta))\int_{\underline{\theta}}^{\theta}u(x)dx\Bigg|_{\underline{\theta}}^{\theta^\star} + \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}m(\theta)\left(\int_{\underline{\theta}}^{\theta}u(x)dx\right)d\theta \\ &= \frac{1}{\beta}(1-M(\theta^\star))\int_{\underline{\theta}}^{\theta^\star}u(\theta)d\theta + \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}m(\theta)\left(\int_{\underline{\theta}}^{\theta}u(x)dx\right)d\theta.\end{aligned}$$

Finally, combining the latter with the incentive constraint,

$$P(T) = \frac{1}{\beta}(1-M(\theta^\star))\int_{\underline{\theta}}^{\theta^\star}u(\theta)d\theta + \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}m(\theta)\left[\frac{1}{\beta}\theta u(\theta) + \lambda W\left(K(0)-G(u(\theta))+T(\theta)\right) - \frac{\theta}{\beta}u(\underline{\theta}) - \lambda\underline{w}\right]d\theta,$$

where, in addition, $u(\theta) = U(T(\theta),\theta)$. The Lagrangian associated to the second-step problem (27) is then

$$\mathcal{L} = \frac{1}{\beta}(1-M(\theta^\star))\int_{\underline{\theta}}^{\theta^\star}u(\theta)d\theta + \frac{1}{\beta}\int_{\underline{\theta}}^{\theta^\star}m(\theta)\left[\frac{1}{\beta}\theta u(\theta) + \lambda W\left(K(0)-G(u(\theta))+T(\theta)\right) - \frac{\theta}{\beta}u(\underline{\theta}) - \lambda\underline{w}\right]d\theta - \mu\mathbb{E}[T(\theta)],$$

with $u(\theta) = U(T(\theta),\theta)$, where $\mu \geq 0$ is the Lagrange multiplier on the constraint $\mathbb{E}[T(\theta)] \leq 0$. The first-order condition with respect to $T(\theta)$ is

$$\frac{1}{\beta}(1-M(\theta^\star))\frac{1}{K(0)+T(\theta)} + \frac{1}{\beta}m(\theta)\frac{\theta}{\beta}\frac{1}{K(0)+T(\theta)} + \frac{1}{\beta}m(\theta)\frac{\lambda}{\gamma}\cdot\frac{1}{K(0)+T(\theta)} - \mu h(\theta) = 0.$$

Rearranging yields

$$\begin{aligned}K(0)+T(\theta) &= \frac{1}{\mu}\cdot\frac{1}{\beta}\cdot\frac{1}{h(\theta)}\left(1-M(\theta^\star)+m(\theta)\left(\frac{1}{\beta}\theta+\frac{\lambda}{\gamma}\right)\right). \\ &\equiv \frac{1}{\mu}\cdot\frac{1}{\beta}\varphi(\theta),\end{aligned}$$

where I have used the definition of $\varphi(\cdot)$ in Assumption 2. Therefore, using $u(\theta) = U(T(\theta),\theta)$ and the definition of $U(T(\theta),\theta)$,

$$\begin{aligned}u(\theta) &= \log\left(k^{nt}(\theta)(K(0)+T(\theta))\right) \\ &= \log\left(k^{nt}(\theta)\frac{1}{\mu}\cdot\frac{1}{\beta}\varphi(\theta)\right).\end{aligned}$$

Also, from the sub-problem ($\mathcal{P}_{TR:S1}$), we know that instantaneous spending is constant throughout the tenure of a government and equals $(\gamma+\lambda)G(u(\theta)) = \exp(u(\theta))$. To compute μ, we take the average of transfers and set it equal to zero:

$$K(0) + \mathbb{E}[T(\theta)] = \frac{1}{\mu} \cdot \frac{1}{\beta} \mathbb{E}\left[\varphi(\theta)\right],$$

hence, $\mu = \mathbb{E}\left[\varphi(\theta)\right] / (\beta K(0))$. The latter can be used to replace μ in the expression for $u(\theta)$. As a result, instantaneous government spending becomes

$$\begin{aligned} \exp(u(\theta)) &= k^{nt}(\theta) \frac{1}{\mu} \cdot \frac{1}{\beta} \varphi(\theta) \\ &= k^{nt}(\theta) \frac{\varphi(\theta)}{\mathbb{E}\left[\varphi(\theta)\right]} K(0). \end{aligned} \tag{28}$$

Finally, transfers are given by

$$\begin{aligned} T(\theta) &= \frac{1}{\mu} \cdot \frac{1}{\beta} \varphi(\theta) - K(0) \\ &= \left(\frac{\varphi(\theta)}{\mathbb{E}\left[\varphi(\theta)\right]} - 1 \right) K(0). \end{aligned} \tag{29}$$

I am left to verify the monotonicity constraint on u. Simple calculation gives

$$\begin{aligned} \frac{d}{d\theta} u(\theta) &= \frac{d}{d\theta} \log \left(k^{nt}(\theta) \frac{\varphi(\theta)}{\mathbb{E}\left[\varphi(\theta)\right]} K(0) \right) \\ &= \frac{1}{k^{nt}(\theta)} \frac{d}{d\theta} k^{nt}(\theta) + \frac{\varphi'(\theta)}{\varphi(\theta)} \\ &= \frac{1}{\theta} \cdot \frac{\lambda\beta}{\gamma\theta + \lambda\beta} + \frac{\varphi'(\theta)}{\varphi(\theta)}. \end{aligned}$$

By Assumption 2, the latter is either positive for all $\theta \leq \theta^\star$ or, if it becomes negative for some $\widetilde{\theta}$, then it will be negative for all $[\widetilde{\theta}, \theta^\star]$. In the former case, utility is non-decreasing, thus, the solution to the relaxed problem satisfies the monotonicity constraint. In particular, optimal government spending and transfers are, respectively, given by (28) and (29), for all $\theta \leq \theta^\star$, and are constant thereafter. In the latter case, instead, there must exist a threshold $\theta^{\star\star} < \theta^\star$ such the solution coincides with the one of the relaxed problem for all $\theta \leq \theta^{\star\star}$ and it is constant thereafter. In particular, optimal government spending and transfers are, respectively, given by (28) and (29), for all $\theta \leq \theta^{\star\star}$, and are constant thereafter. This also proves Proposition 3.

Finally, notice that government spending (28) and transfers (29) correspond to the case in which $v = 0$. By replacing $K(0)$ with $K(v)$, we obtain the their counterparts for any v.

Uniform Distribution. I now provide an explicit solution for the special case in which shocks are uniformly distributed. Let $H(\theta) = \theta/(\overline{\theta} - \underline{\theta})$, for $\theta \in [\underline{\theta}, \overline{\theta}]$. The equation $\beta\mathbb{E}\left[\theta | \theta \geq \theta^\star\right] = \theta^\star$ immediately implies the threshold:

$$\theta^\star = \frac{\overline{\theta}\beta}{2-\beta}.$$

In addition, $M(\theta) \equiv H(\theta) + \theta(1-\beta)h(\theta) = (2-\beta)\theta/(\overline{\theta} - \underline{\theta})$, hence, $m(\theta) = (2-\beta)/(\overline{\theta} - \underline{\theta})$. I can then compute the weights α in Proposition 2 explicitly. First,

$$\begin{aligned}\varphi(\theta) &\equiv \frac{1}{h(\theta)}\left(1 - M(\theta^\star) + m(\theta)\left(\frac{1}{\beta}\theta + \frac{\lambda}{\gamma}\right)\right)\\ &= \kappa_0 + (2-\beta)\frac{1}{\beta}\theta,\end{aligned}$$

where $\kappa_0 \equiv \overline{\theta}(1-\beta) - \underline{\theta} + (2-\beta)\lambda/\gamma$. It is immediate to see that $\varphi(\cdot)$ is increasing, thus, Assumption 2 is verified. By Proposition 3, $\theta^{\star\star} = \theta^\star$. Also,

$$\begin{aligned}\int_{\underline{\theta}}^{\theta^\star} \varphi(\theta)h(\theta)d\theta + \varphi(\theta^\star)\int_{\theta^\star}^{\overline{\theta}} h(\theta)d\theta &= \kappa_0 + (2-\beta)\frac{1}{\beta}\left[\int_{\underline{\theta}}^{\theta^\star} \theta h(\theta)d\theta + \theta^\star \int_{\theta^\star}^{\overline{\theta}} h(\theta)d\theta\right]\\ &= \kappa_0 + (2-\beta)\frac{1}{\beta}\kappa_1\end{aligned}$$

where, using $\int_{\underline{\theta}}^{\theta^\star} \theta h(\theta)d\theta = 1 - \int_{\theta^\star}^{\overline{\theta}} \theta h(\theta)d\theta$,

$$\begin{aligned}\kappa_1 &\equiv 1 - \frac{1}{\overline{\theta} - \underline{\theta}}\left[\frac{1}{2}\left((\overline{\theta})^2 - (\theta^\star)^2\right) - \theta^\star\left(\overline{\theta} - \theta^\star\right)\right]\\ &= 1 - \frac{1}{2}\frac{(\overline{\theta} - \theta^\star)^2}{\overline{\theta} - \underline{\theta}}\\ &= 1 - \frac{2\overline{\theta}^2}{\overline{\theta} - \underline{\theta}}\left(\frac{1-\beta}{2-\beta}\right)^2,\end{aligned}$$

where the last line uses the definition of $\theta^\star$. Therefore, for $\theta \leq \theta^\star$,

$$\begin{aligned}\frac{T_v^{tr}(\theta)}{K(v)} &= \alpha(\theta) - 1\\ &= \frac{\varphi(\theta)}{\int_{\underline{\theta}}^{\theta^\star} \varphi(\theta)h(\theta)d\theta + \varphi(\theta^{\star\star})\int_{\theta^\star}^{\overline{\theta}} h(\theta)d\theta} - 1\\ &= (2-\beta)\frac{1}{\beta}\cdot\frac{1}{\kappa_0 + (2-\beta)\frac{1}{\beta}\kappa_1}\left(\theta - 1 + \frac{2\overline{\theta}^2}{\overline{\theta} - \underline{\theta}}\left(\frac{1-\beta}{2-\beta}\right)^2\right)\end{aligned}$$

and the decomposition in the main text follows by letting

$$\overline{T}_0(\lambda, \beta) \equiv (2-\beta)\frac{1}{\beta}\cdot\frac{1}{\kappa_0 + (2-\beta)\frac{1}{\beta}\kappa_1}$$

and

$$\overline{T}_1(\beta) \equiv \frac{2\overline{\theta}^2}{\overline{\theta} - \underline{\theta}}\left(\frac{1-\beta}{2-\beta}\right)^2.$$

Finally, since κ_0 is increasing in λ while κ_1 is independent of λ, it follows immediately that $\partial\overline{T}_0(\lambda,\beta)/\partial\lambda < 0$, as claimed in the main text.

B.4 Welfare Gains

To compare welfare in the transfer and no transfer scenario we *(i)* compute welfare in the no-transfer case, for given resource $\kappa y/r$, namely $v^{nt}(\kappa y/r)$ *(ii)* compute the amount of resources necessary to deliver the same utility in the transfer case $K(v^{nt})$, define welfare gain as the difference between $K(v^{nt})$ and $\kappa y/r$ divided by GDP.

In the no transfer problem, we know that the planner's value function is

$$v^{nt}(a) = \bar{A} + \frac{1}{\gamma}\log(a),$$

where

$$\bar{A} = \frac{1}{\gamma}\mathbb{E}\left[\theta\log(k^{nt}(\theta)) - k^{nt}(\theta)\frac{\gamma\theta+\lambda}{\gamma(\gamma+\lambda)} + 1\right],$$

and we set initial amount of resources to $a = \kappa y/r$. Further, we know that in the transfer case $K(v^{nt}) = K(0)\exp(\gamma\bar{A})(\kappa y/r)$, where

$$K(0) = \exp\left(-\mathbb{E}\Big[\,\theta\log(k^{nt}(\theta)\alpha(\theta)) + \lambda\log\left((\gamma+\lambda-k^{nt}(\theta))\frac{\alpha(\theta)}{\lambda}\right)\,\Big]\right).$$

Now define welfare gains as the difference in resources needed to make the planner indifferent between making transfers or not, as a proportion of the endowment:

$$\Psi(\beta) \equiv \left(\frac{\kappa y}{r} - K(v^{nt})\right)\frac{1}{y} = \big(1 - K(0)\exp(\gamma\bar{A})\big)\,\frac{\kappa}{r}.$$

Welfare gains clearly depend on all the parameters of the model, but I make explicit the dependency on β to emphasize the following. Define tha function $\psi(\beta)$ as

$$\psi(\beta) \equiv K(0)\exp(\gamma\bar{A}) = \exp\left(-\mathbb{E}\left[\theta\log(\alpha(\theta)) + \lambda\log\left((\gamma+\lambda-k^{nt}(\theta))\frac{\alpha(\theta)}{\lambda}\right) + k^{nt}(\theta)\frac{\gamma\theta+\lambda}{\gamma(\gamma+\lambda)} - 1\right]\right),$$

then welfare gains are

$$\Psi(\beta) = (1-\psi(\beta))\,\frac{\kappa}{r}.$$

Notice that when everybody is constrained (for β very low), weights and consumption share are constant, namely $\alpha(\theta) = 1$; $k^{nt}(\theta) = \gamma\ \forall\theta$. Substituting those values in the expression for welfare gains we get that $\psi(\beta) = 1$, namely there is no gain in setting-up a transfer system when the political friction is extreme. When $\beta = 1$, instead, we have $\alpha(\theta) = (\gamma\theta+\lambda)(\gamma+\lambda)$ and $k^{nt}(\theta) = \gamma\theta(\gamma+\lambda)(\gamma\theta+\lambda)^{-1}$, meaning that $\psi(1) = \exp(-\mathbb{E}[\theta\log(\alpha(\theta))])$, which is grater than one by Jensen's inequality. In other terms, welfare gains are positive when there is no political friction and vanish when governments' exclusively care about their own consumption.

C Proofs for Section 3

I provide a heuristic derivation of the HJB system (12), (13). Let us start with the equation for the value function Υ in (13). This function represents the value of a sequence of spending, after the government has been dissolved, generated by policy function $g^{\star}(a,\theta)$ which prescribes government spending as a function of current wealth a and type θ. Formally, take any time $t_0 < \tau_0$ (the case with τ_n, $n > 0$, is analogous) and suppose country's wealth is $a_{t_0} = a$ and current government's type is $\theta_0 = \theta$. Then

$$\Upsilon(a, \theta) = \mathbb{E}_{t_0}\left[\sum_{n=0}^{\infty} e^{-\gamma\tau_n}\theta_n \int_{\tau_n}^{\tau_{n+1}} e^{-\gamma(t-\tau_n)} \log(g^\star)dt\right]$$

where wealth evolves according to the process $\dot{a}_t = ra_t - g^\star(a_t, \theta_t)$.

At time t_0, the government enjoys flow utility of $\theta \log(g^\star(a, \theta))$. Moreover, only two things can happen in the next instant of time $t_0 + dt$. First, with probability $e^{-\lambda dt}$, the type θ remains unchanged and wealth increases by a deterministic amount da. When this occurs, the expected discounted payoff becomes $e^{-\gamma dt}\Upsilon(a+da, \theta)$. Second, with probability $(1-e^{-\lambda dt})$, a new taste shock θ will be drawn from the distribution $H(\theta)$. When this occurs, the expected discounted payoff becomes $\mathbb{E}[e^{-\gamma dt}\Upsilon(a + da, \widetilde{\theta})]$. Putting the pieces together, and weighting them by their respective probabilities, I obtain

$$\Upsilon(a, \theta) = \theta \log(g^\star(a, \theta))dt + e^{-\lambda dt}e^{-\gamma dt}\Upsilon(a + da, \theta) + (1 - e^{-\lambda dt})\mathbb{E}\left[e^{-\gamma dt}\Upsilon(a + da, \widetilde{\theta})\right].$$

Using the approximations

$$e^{-\gamma dt} = 1 - \gamma dt + O(dt^2),$$

$$e^{-\lambda dt} = 1 - \lambda dt + O(dt^2)$$

and ignoring higher-order terms, the expression above can be rewritten as

$$\Upsilon(a, \theta) = \theta \log(g^\star(a, \theta))dt + (1 - \lambda dt - \gamma dt)\Upsilon(a + da, \theta) + \lambda dt\mathbb{E}\left[\Upsilon(a + da, \widetilde{\theta})\right].$$

Since changes in wealth are deterministic, the term $\Upsilon(a + da, \theta)$ is simply $\Upsilon(a, \theta) + \Upsilon_a(a, \theta)da$, where Υ_a is the derivative of the value function with respect to its first argument, and $da = (ra - g^\star(a, \theta))dt$. As a result,

$$\begin{aligned}\Upsilon(a, \theta) =& \theta \log(g^\star(a, \theta))dt + (1 - \lambda dt - \gamma dt)\left[\Upsilon(a, \theta) + \Upsilon_a(a, \theta)(ra - g^\star(a, \theta))dt\right] \\ &+ \lambda dt\mathbb{E}\left[\Upsilon(a, \widetilde{\theta}) + \Upsilon_a(a, \widetilde{\theta})(ra - g^\star(a, \widetilde{\theta}))dt\right].\end{aligned}$$

Ignoring second-order terms, subtracting $\Upsilon(a, \theta)$ from both sides, and dividing through by dt (letting $dt \longrightarrow 0$) yields

$$(\lambda + \gamma)\Upsilon(a, \theta) = \theta \log(g^\star(a, \theta)) + \Upsilon_a(a, \theta)(ra - g^\star(a, \theta)) + \lambda\mathbb{E}\left[\Upsilon(a, \widetilde{\theta})\right],$$

which is (13).

The proof for (12) follows analogous arguments. There are two differences though. First, in this case the policy function $g^\star(a, \theta)$ is not taken as given, but chosen by the government in charge, hence, the maximization operator in (12). In particular, the maximization problem is subject to the fiscal rule (14). Second, when the shock causing a type change occurs, the government is dissolved. As a result, this event is discounted by $\beta e^{-\gamma dt}$, instead of the standard discount $e^{-\gamma dt}$ and, in addition, the value function switches from J to Υ.

To find a solution to the HJB system (12), (13), I guess and verify that the value functions take the following form:

$$\Upsilon(a,\theta) = \overline{\Upsilon}(\theta,\theta^\star) + A(\theta)\log(a),$$
$$J(a,\theta) = \overline{J}(\theta,\theta^\star) + B(\theta)\log(a),$$

for some functions $\overline{\Upsilon}$, A, $\overline{J}$ and B. The first-order condition of the maximization problem in (12) is then

$$g^\star(a,\theta) = \frac{\theta}{J_a(a,\theta)} = \frac{\theta}{B(\theta)}a,$$

for $\theta < \theta^\star$ and simply $g^\star(a,\theta) = k^{nt}(\theta^\star)a$, otherwise. Substituting $g^\star(a,\theta)$, together with the conjectures above, into (12) and rearranging gives

$$(\gamma+\lambda)\left(\overline{J}(\theta,\theta^\star) + B(\theta)\log(a)\right) = \theta\log\left(\frac{\theta}{B(\theta)}a\right) + B(\theta)\left(r - \frac{\theta}{B(\theta)}\right) + \lambda\beta\mathbb{E}\left[\overline{\Upsilon}(\widetilde{\theta},\theta^\star) + A(\widetilde{\theta})\log(a)\right],$$

for $\theta < \theta^\star$ and an analogous expression for $\theta \geq \theta^\star$. Similarly, equation (12) becomes

$$(\gamma+\lambda)\left(\overline{\Upsilon}(\theta,\theta^\star) + A(\theta)\log(a)\right) = \theta\log\left(\frac{\theta}{B(\theta)}a\right) + A(\theta)\left(r - \frac{\theta}{B(\theta)}\right) + \lambda\mathbb{E}\left[\overline{\Upsilon}(\widetilde{\theta},\theta^\star) + A(\widetilde{\theta})\log(a)\right],$$

for $\theta < \theta^\star$ and an analogous expression for $\theta \geq \theta^\star$.

Consider the case with $\theta < \theta^\star$. Equalizing terms multiplying $\log(a)$, the second equation immediately gives

$$(\gamma+\lambda)A(\theta) = \theta + \lambda\mathbb{E}\left[A(\widetilde{\theta})\right],$$

hence, taking expectations of both sides,

$$\mathbb{E}\left[A(\widetilde{\theta})\right] = \frac{1}{\gamma}$$

and

$$A(\theta) = \frac{\gamma\theta+\lambda}{\gamma(\gamma+\lambda)}.$$

Using the latter in the first equation and equalizing terms multiplying $\log(a)$, I obtain

$$(\gamma+\lambda)B(\theta) = \theta + \lambda\beta\frac{1}{\gamma}$$

and, thus,

$$B(\theta) = \frac{\gamma\theta+\lambda\beta}{\gamma(\gamma+\lambda)}.$$

Spending is therefore

$$g^\star(a,\theta) = \frac{\theta}{B(\theta)}a = k^{nt}(\theta)a.$$

The latter, together with the law of motion for wealth, generates the optimal spending in Proposition 1. Finally, the functions $\overline{\Upsilon}(\theta,\theta^\star)$, $\overline{J}(\theta,\theta^\star)$ can be obtained by solving the system of equations

$$(\gamma+\lambda)\overline{J}(\theta,\theta^\star)=\theta\log\left(\frac{\theta}{B(\theta)}\right)+B(\theta)\left(r-\frac{\theta}{B(\theta)}\right)+\lambda\beta\mathbb{E}\left[\overline{\Upsilon}(\widetilde{\theta},\theta^\star)\right],$$

$$(\gamma+\lambda)\overline{\Upsilon}(\theta,\theta^\star)=\theta\log\left(\frac{\theta}{B(\theta)}\right)+A(\theta)\left(r-\frac{\theta}{B(\theta)}\right)+\lambda\mathbb{E}\left[\overline{\Upsilon}(\widetilde{\theta},\theta^\star)\right],$$

for $\theta<\theta^\star$ and analogous expressions for $\theta\geq\theta^\star$. The latter also verify our original conjecture.

Transfers Consider now the implementation with transfers. The derivation of the HJB equations is analogous to the no-transfer case, so I will not repeat it here. There are two main differences. First, at the time a new government is formed, the country's assets change discontinuously due to the payment for the credit line and the new transfer. Formally, assets now evolve according to

$$dx_t=(rx_t+\kappa y+b_t-g_t)dt+\left(-\lambda^{-1}b_{t-}+\chi(\widetilde{\theta}_t,a'_{t-})\right)dN_t,$$

where N_t is the jump process and a'_{t-} is wealth at the time of dissolution net of the payment $-\lambda^{-1}b_{t-}$. In particular, notice that, when the Poisson jump occurs, assets x_t jump by the amount $-\lambda^{-1}b_{t-}+\chi(\widetilde{\theta}_t,a'_{t-})$. Second, now the government in charge has two choice variables, spending and credit-line drawdown.

Using (16), after the government is dissolved and a new government with type $\widetilde{\theta}$ is formed, the country's wealth becomes $a_t=\alpha(\widetilde{\theta})\left(a_{t-}-\frac{1}{\lambda}b(\theta,a_{t-})\right)$. As a result, the system of HJB equations is

$$(\lambda+\gamma)J(\theta,a)=\max_{g,b}\left\{\theta\log(g(\theta,a))+J_a(\theta,a)(ra+b(\theta,a)-g(\theta,a))\right\}+\beta\lambda\mathbb{E}\left[\Upsilon\left(\widetilde{\theta},\alpha(\widetilde{\theta})\left(a-\frac{1}{\lambda}b(\theta,a)\right)\right)\right],$$

$$(\lambda+\gamma)\Upsilon(\theta,a)=\theta\log(g^\star(\theta,a))+\Upsilon_a(\theta,a)(ra+b^\star(\theta,a)-g^\star(\theta,a))+\lambda\mathbb{E}\left[\Upsilon\left(\widetilde{\theta},\alpha(\widetilde{\theta})\left(a-\frac{1}{\lambda}b^\star(\theta,a)\right)\right)\right],$$

where the maximization problem in the first equation is subject to the constraint $g\leq k^{tr}(\theta^{\star\star})a$. As for the no-transfer case, I conjecture

$$\Upsilon(a,\theta)=\overline{\Upsilon}(\theta,\theta^\star)+A(\theta)\log(a),$$

$$J(a,\theta)=\overline{J}(\theta,\theta^\star)+B(\theta)\log(a),$$

for some functions $\overline{\Upsilon}$, A, $\overline{J}$ and B. The first-order condition with respect to g is exactly the same as in the no-transfer case. Also, the same arguments for the no-transfer case yield $\mathbb{E}[A(\widetilde{\theta})]=1/\gamma$ and $B(\theta)=(\gamma\theta+\lambda\beta)\gamma^{-1}(\gamma+\lambda)^{-1}$. As a result, the first-order condition with respect to b is

$$\frac{\gamma\theta+\lambda\beta}{\gamma(\gamma+\lambda)}\cdot\frac{1}{a}-\beta\lambda\frac{1}{\gamma}\cdot\frac{\frac{1}{\lambda}}{a-\frac{1}{\lambda}b^\star(\theta,a)}=0.$$

Straightforward algebra gives

$$b^{\star}(\theta, a) = \gamma\lambda\frac{\theta - \beta}{\gamma\theta + \lambda\beta}a \equiv \bar{b}(\theta)a.$$

Finally, by replacing the expressions for $g^{\star}(\theta, a)$ and $b^{\star}(\theta, a)$ into the HJB equations, I obtain two equations that can be used to solve for $\overline{\Upsilon}(\theta, \theta^{\star})$, $\overline{J}(\theta, \theta^{\star})$, thus verifying our original conjecture.