The Optimal Degree of Flexibility in Fiscal Policy: A Quantitative Exploration

Work in progress

Guillaume Sublet

Université de Montréal and CIREQ †

June 25, 2026

The design of a fiscal rule involves a trade-off between disciplining a deficit-biased government and granting it the flexibility to respond to shocks to its fiscal needs. The optimal degree of flexibility depends on the elasticity of the right tail of the distribution of shocks, which measures the incentive cost of imposing discipline. This paper measures this need for flexibility. Because fiscal needs are not directly observable, I use a tractable model of government spending to infer past fiscal needs from government finance data, and I then apply tools from heavy-tail analysis to estimate the tail elasticity of their distribution. Applying the method to members of the European Union, I find novel evidence of a Pareto tail of the distribution of shocks to fiscal needs, which indicates a large need for flexibility and suggests avenues to reform the Stability and Growth Pact.

Keywords: Fiscal Rule, Delegation, Mechanism Design, Money Burning, Price vs. Quantity

JEL classification: D02, D82, E02, E6, F36, F45, F55, H1, H62, P16

† Email: guillaume.sublet@umontreal.ca.

Part of this paper previously circulated under the title "The Optimal Degree of Discretion in Fiscal Policy." I thank Manuel Amador and Pierre Yared for comments and suggestions that greatly improved the paper. I thank Marco Bassetto, Anmol Bhandari, Roel Beetsma, Job Boerma, Rui Castro, Louphou Coulibaly, Alessandro Dovis, Mike Golosov, Barış Kaymak, Tim Kehoe, Rishabh Kirpalani, Ellen McGrattan, Thomas Phelan, Todd Schoellman, Smriti Upadhyay, Pierre-Olivier Weill and the audiences of academic seminars and conferences for helpful suggestions and comments. I thank Joan Gieseke for expert copyediting. I am responsible for any shortcomings of this paper.

Introduction

The European Union is currently debating the design of the Excessive Deficit Procedure of its fiscal rule (the Stability and Growth Pact). The Stability and Growth Pact constrains national fiscal policy with two limits: a limit on the deficit (3% of GDP) and a limit on the debt (60% of GDP). This paper focuses on the issues raised by inefficiently high deficits stemming from a deficit bias and abstracts from the sustainability of the debt. This centers the debate on the fundamental trade-off in the design of a fiscal rule: how to optimally balance the need to impose discipline on a deficit-biased government with the need for flexibility to respond to shocks to the country's fiscal needs. Should society impose discipline on its government with a fiscal rule featuring a graduated schedule of penalties on excessive deficits, a cap on deficits, or a combination of the two? The thickness of the right tail of the distribution of shocks to fiscal needs is essential to the optimal design of a fiscal rule. Irrespective of whether the optimal rule is a cap on the deficit (Amador, Werning, and Angeletos (2006), Halac and Yared (2022)) or a graduated schedule of penalties, the thickness of the tail provides a measure of the need for discretion (Sublet (2026)).

This paper takes a first step toward quantifying the trade-off between discipline and flexibility in the design of a fiscal rule. It builds on a companion theory paper, Sublet (2026), which studies the optimal design of a fiscal rule enforced by penalties whose burden is borne asymmetrically by the government and the society, as in the "leaky bucket" model of government finances (Amador and Bagwell (2013, 2020)). The analysis there covers the full range of degrees of asymmetry in the cost of discipline. At one end of the spectrum, the asymmetry is maximal in the sense that the society is immune to the penalty on the government, and the optimal fiscal rule is a graduated schedule of penalties reminiscent of a Pigouvian tax. The other end of the spectrum nests the symmetric case, and the optimal fiscal rule is a cap on the deficit, in line with the main insight from the literature. For degrees of asymmetry between the two polar cases, while the main insight from the literature prevails under more stringent conditions, under a complementary set of conditions, a fiscal rule featuring a graduated schedule of penalties is optimal. The key lesson of the theory for the present paper is that the thickness of the right tail of the distribution of shocks to fiscal needs governs the optimal degree of flexibility.

The contribution of this paper is to measure this need for flexibility. I propose a tool to measure the need for flexibility based on government finance data. As the theory shows, the elasticity of the right tail of the distribution of shocks to fiscal needs gives a quantifiable measure of the incentive cost of imposing discipline, and hence of the need for flexibility (Sublet (2026)). I find evidence of a Pareto tail of the distribution of shocks to fiscal needs of members of the European Union, which indicates a large need for flexibility. A by-product of the analysis, which is not part of the companion theory paper, is to establish that the sufficient optimality conditions characterizing an optimal fiscal rule remain valid for distributions with an unbounded support—such as the Pareto distribution documented in the measurement—so long as the first moment is finite.

I obtain these contributions by building on a line of research modeling the trade-off between discipline and flexibility in the design of a fiscal rule (Amador, Werning, and Angeletos (2006), Halac and Yared (2014, 2022a, 2023)). Consider a small open economy whose government decides how much to spend in response to shocks to the country's fiscal needs. Suppose that while the government and the society agree on the need for flexibility to respond to the shocks, they disagree on the relative value between present and future public spending. In particular, the government is shortsighted in the sense that it overvalues current spending. The shortsightedness captures, in a reduced form, the incentive to overborrow on the part of the government due to various reasons: political turnover (Alesina and Passalacqua (2016) offer a survey), heterogeneity in discounting among members of the society (Jackson and Yariv (2014, 2015)), or, for members of an economic union, the common pool problem caused by a common monetary authority that lacks the ability to commit (Beetsma and Uhlig (1999), Chari and Kehoe (2007), and Aguiar, Amador, Farhi, and Gopinath (2015)). The combination of shocks and a present-biased objective calls for both flexibility and discipline. Lastly, a trade-off between flexibility and discipline arises because the realization of the shock to fiscal needs is private information to the government, which implies that the prescription of the fiscal rule cannot be contingent upon the shock. I follow Halac and Yared (2022) in modeling a fiscal rule as a penalty schedule, which is a function of the government's choice.

The departure from the standard framework is that the cost of meting out a penalty on the government is borne asymmetrically by the government and the society. The problem of designing a fiscal rule then maps to a mechanism design problem without transfers but with asymmetric "money burning," where money burning refers to a penalty that is meted out. The companion theory paper, Sublet (2026), solves this problem with global Lagrangian methods, blending insights from the first-order approach to obtain easy-to-check optimality conditions (the appendix records the methods needed below).

The key object is a closed-form marginal penalty—the foa-wedge—that balances the benefit of discipline, measured by the degree of present bias, with the incentive cost of imposing discipline through on-equilibrium penalties. The main insight from the theory, which is the point of departure for this paper, is that this incentive cost is governed by the elasticity of the right tail of the distribution of shocks: because an on-equilibrium penalty is borne by all governments with larger fiscal needs, its incentive cost is the inverse of the rate at which the tail thins out. The elasticity of the right tail of the distribution of shocks to fiscal needs thus gives a quantifiable measure of the incentive cost of imposing discipline and hence of the need for flexibility (Sublet (2026)).

For a broad class of environments, the optimal fiscal rule is a hybrid rule featuring no penalties below a threshold, a graduated schedule of penalties conforming with the foa-wedge above the threshold, and a cap on the deficit. The thickness of the tail governs the stringency of the cap: a thin tail calls for a binding cap, whereas a thick tail calls for more flexibility and, for a sufficiently thick tail, for no rule at all. When on-equilibrium penalties are used, it is optimal to exempt low levels of spending from penalties, which lowers the level of the penalty schedule and creates a kink—in contrast with the notch in the penalty schedule of the Stability and Growth Pact.

The theory indicates that the elasticity of the right tail of the distribution of fiscal needs is a quantifiable measure of the incentive cost of imposing discipline, and hence of the need for flexibility, and a key input to the design of an optimal fiscal rule (Sublet (2026)). The task of measuring the need for flexibility from government finance data has two challenges, however. First, fiscal needs are not directly observable. Second, the distribution of fiscal needs is an infinite dimensional object. A parametric assumption on the distribution would, to a large extent, presume the thickness of the tail instead of deducing it from the data (e.g., irrespectively of its mean and variance, a normal distribution has a thin tail).

In a first step, I use a tractable positive model to infer past fiscal needs from government finance data. The positive model is a standard model of (government) spending augmented with taste shocks that may be persistent. The taste shocks capture fiscal needs as a catchall for shocks to the need for a deficit, irrespectively of whether it originates in a shock to the need for public spending, a shock to the cost of servicing the debt, or a shock to government revenue, as in the normative theory. The identification of fiscal needs obtains from inverting the government policy function to recover the fiscal needs that rationalize the government finance data. A difficulty, however, is that the policy function depends on the distribution of fiscal needs, which is the object of interest to be inferred in the second step. Assuming log utility keeps the two-step approach tractable because I obtain a closed-form expression for the policy function that depends only on the first moment of the distribution of fiscal needs. As a result, the first step alone delivers the exact identification of past fiscal needs from government finance data.

In a second step, I infer the thickness of the tail of the distribution of measured fiscal needs. Because rare events featuring large observations—say, during recessions—are a common feature of government finance data, I use tools from heavy-tail analysis. Following Gabaix and Ibragimov (2011), a log-log plot of rank and size of measured spending needs can be used to detect evidence of a Pareto tail and to measure the thickness of the tail. I adopt a common practice in heavy-tail analysis by substantiating the analysis with a Hill plot (Resnick (2007), Chapter 4). I apply the two-step methodology to the case of the European Union and find novel evidence of a thick (Pareto) tail of the distribution of shocks to fiscal needs. The concluding section contains suggestions to reform the Stability and Growth Pact.

Related literature. More broadly, the theory relates to the literature on delegation, money burning, and the design of rules to discipline a policy-making authority to act in the best interest of the society (Holmström (1977), Melumad and Shibano (1991), Alonso and Matoushek (2008), Kováč and Mylovanov (2009), Athey, Atkeson, and Kehoe (2005), Amador, Werning, and Angeletos (2006), Ambrus and Egorov (2013, 2017), Amador and Bagwell (2013, 2020, 2022), Clayton and Schaab (2022), and Halac and Yared (2014, 2020a, 2020b, 2022a, 2022b, 2023)). The companion theory paper, Sublet (2026), contributes to this literature in two ways. First, it contains a near-complete characterization of the optimal delegation contract with asymmetric money burning in a model with a multiplicative bias. Second, the analysis there blends insights from the first-order approach, reminiscent of Diamond (1998)'s ABC formula, to obtain easy-to-check and intuitive optimality conditions from the global Lagrangian methods. The optimality conditions elicit the central role of the thickness of the tail of the distribution of shocks for the optimal balance between the needs for discipline and flexibility.

The present paper takes a first step toward a quantitative application of this rich theory on the trade-off between discipline and flexibility in the design of fiscal rules on deficits. The quantitative literature on fiscal rules has, so far, focused on inefficient debt levels—for example, as a result of the possibility of default (see, e.g., Azzimonti, Battaglini, and Coate (2016), Hatchondo, Martinez, and Roch (2020), Alfaro and Kanczuk (2019), and Aguiar, Amador, and Fourakis (2020); an exception is Bassetto and Sargent (2006)). Felli, Piguillem, and Shi (2021) show that the risk of default makes it optimal to introduce a default rule.

Because a Pigouvian tax is a price regulation and a cap is a quantity regulation, this paper also contributes to the literature on the choice of regulatory instrument. In an influential contribution, Weitzman (1974) finds that the relative curvature of the benefit and cost functions of economic activity determines whether fixing the price or fixing the quantity is preferable.[2] Perhaps surprisingly, the distribution of shocks does not affect the ranking of these two simple instruments. Allowing for a richer set of instruments, as I do in this paper, reveals the role of the distribution of shocks: a thick tail of the distribution of shocks calls for flexibility, which favors on-equilibrium penalties.

1 Model

This section introduces a standard model used to analyze the design of a fiscal rule and relaxes the assumption of symmetry in the cost of burning money (Amador, Werning, and Angeletos (2006), Halac and Yared (2014, 2022a)).

1.1 Economic environment

Consider a small open economy whose government decides how much to spend and borrow in response to shocks to the fiscal needs of its society. The government is present biased in the sense that it discounts the future at a higher rate than its society. The combination of shocks to fiscal needs and a present-biased objective creates a need for both flexibility and discipline.

[2]For instance, the consensus in environmental economics is that the marginal benefit curve of abatement is relatively flat, whereas the marginal cost curve is steep, which according to Weitzman (1974) favors price regulation over quantity regulation (see McKibbin Wilcoxen (2002)).

Formally, the preferences of the government over the allocation of public spending over time are represented by the objective function

$$\theta U(g) + \beta(W(x) - P(g)), \tag{1}$$

where $\theta U(\cdot)$ denotes the utility from government spending $g \geq 0$, and the continuation value $W(\cdot)$ is a function of future assets $x \in \mathbb{R}$.[3] Both U and W are twice continuously differentiable and strictly increasing. The index U is strictly concave and satisfies Inada conditions, and the index W is concave. Shocks to the country's *fiscal needs*, denoted θ, are private information to the government. The shocks follow a distribution F whose support is an interval Θ with lower bound $\underline{\theta} > 0$ and supremum $\bar{\theta} > \underline{\theta}$.[4] Although the supremum can be infinite, the first moment is assumed to be finite. The distribution F is twice continuously differentiable with density f. The *tail* of the distribution refers to $1 - F$.

The degree of *present bias* of the government is $1 - \beta \in (0, 1]$, and, to simplify the notation, the discount factor of the society is subsumed into the continuation value. A *fiscal rule* is a penalty schedule $P(\cdot)$ that is non-negative, $P(g) \geq 0$ for $g \geq 0$.

The government's budget constraint is $g + x = T$, where $T > 0$ denotes government revenues. For simplicity, the gross interest rate is exogenous and normalized to one. Substituting the budget constraint into the objective of the government (1) gives the government's problem

$$g(\theta) \in \arg\max_{g \geq 0} \; \theta U(g) + \beta(W(T - g) - P(g)). \tag{2}$$

In contrast, the society's objective is not present biased, and, from an ex ante perspective,

$$\int_{\Theta} \left[\theta U(g(\theta)) + W(T - g(\theta)) - \rho P(g(\theta))\right] dF(\theta), \tag{3}$$

where $1 - \rho$ denotes the *asymmetry* in the cost of penalties for the economic union. This paper covers the full spectrum of the degrees of asymmetry in the cost of the penalties, $\rho \in [0, 1]$. At one end of the spectrum (i.e., $\rho = 0$), the society is immune to the penalties on the government. At the other end of the spectrum (i.e., $\rho = 1$), penalties symmetrically affect the government

[3] Amador, Werning, and Angeletos (2006) show that the results apply to a multi-period environment with iid shocks. Halac and Yared (2014) study an infinite horizon environment with persistent shocks.

[4] Shocks to government revenues map to shocks to fiscal needs with a CARA utility index (see Section 5.4 of Amador, Werning, and Angeletos (2006)).

and the society.[5] For $\rho \in (0,1)$, an on-equilibrium penalty mutually affects the society and the government, albeit asymmetrically.

Formally, the design of a fiscal rule consists of determining a penalty schedule $P(\cdot) \geq 0$ to maximize society's welfare (3) subject to the implementability constraint (2).

Economic union. Note that the analysis in this paper applies, at the cost of additional notation, to the design of a fiscal rule for an economic union that is not fiscally integrated. Appendix A contains the additional notation needed for an economic union and the mapping to the environment of this section. In short, because the lack of fiscal integration limits the use of transfers between countries, the key trade-off remains between flexibility and discipline—separately from insurance.[6]

Some useful definitions. An *allocation* is a contingent plan for government spending $g(\cdot)$ with $g(\theta) \geq 0$ for $\theta \in \Theta$. A fiscal rule $P(\cdot)$ *implements* an allocation $g(\cdot)$ if for $\theta \in \Theta$, $g(\theta)$ solves the government's problem (2). Consider a rule $P(\cdot)$ and the allocation it implements $g(\cdot)$. A strictly positive penalty on g is *on-equilibrium* if there exists $\theta \in \Theta$ such that, if θ realizes, the penalty is meted out (i.e., $g(\theta) = g$). Otherwise, the penalty is *off-equilibrium*. Let g_f denote the *flexible allocation*, which is the allocation that solves the government problem (2) in the absence of a fiscal rule. Define the *wedge* Δ in the Euler equation of the government as follows: $(1 - \Delta(g,\theta))\theta U'(g) = \beta W'(T - g)$. The wedge allows us to read the discipline imposed by a fiscal rule off the allocation it implements. The wedge evaluated at the flexible allocation is $\Delta(g_f(\theta), \theta) = 0$. A rule that constrains the government to spend less (i.e., $g(\theta) \leq g_f(\theta)$) induces a positive wedge.

1.2 Designing a mechanism with asymmetric "money burning"

The optimal design of a fiscal rule maps to a mechanism design problem without transfers but with money burning. Using the revelation principle, the composition of the penalty schedule

[5] The economic environment with $\rho = 1$ nests the model in Amador, Werning, and Angeletos (2006) and the model in Halac and Yared (2022a) without an upper bound on punishments (i.e., with unlimited enforcement).

[6] Harstad (2007) provides a rationale for limits on transfers to curb the incentive of members to strategically delay reaching an agreement while bargaining over the design of a common rule.

and the allocation gives the *money-burning* schedule $t(\theta) = P(g(\theta))$ for $\theta \in \Theta$. Note that on-equilibrium penalties impose discipline at the cost of "burning money," whereas off-equilibrium penalties do not "burn money."

Incentive compatibility constraints guarantee the implementability of the allocation as in (2). An allocation $g(\cdot)$ is *incentive compatible* given a money-burning schedule $t(\cdot)$ if

$$\theta U(g(\theta)) + \beta\left(W(T - g(\theta)) - t(\theta)\right) \geq \theta U(g(\hat{\theta})) + \beta\big(W(T - g(\hat{\theta})) - t(\hat{\theta})\big), \quad \text{for } \theta, \hat{\theta} \in \Theta. \quad \text{(IC)}$$

A fiscal rule is *optimal* if the allocation it implements $g(\cdot)$ and the associated schedule $t(\cdot)$ solve

$$\max_{g(\cdot),\ t(\cdot)} \left\{ \int_{\Theta} \big[\theta U(g(\theta)) + W(T - g(\theta)) - \rho t(\theta)\big] dF(\theta) \;\Big|\; \text{(IC) and } t(\theta) \geq 0 \text{ for } \theta \in \Theta \right\}. \quad (4)$$

The intercept of the schedule, if left implicit, is $t(\underline{\theta}) = 0$ and $P(g) = 0$ for $g \leq g(\underline{\theta})$.

The non-negativity constraint on money burning sets program (4) apart from the design of a mechanism with transfers because it rules out cross-subsidization across types.[7] The solution method exploits powerful Lagrangian techniques to allow for the non-negativity constraint on money burning (Section B of the Appendix contains a description of the solution method).

2 On the balance between discipline and flexibility

In this section, I use the first-order approach to analyze the economics of on-equilibrium penalties. I first study a relaxed problem and then explore the conditions under which the relaxed constraints may be binding. The first-order approach provides an analytical formula for the (locally) optimal balance between the competing needs for discipline and flexibility.

The following standard result exploits the incentive compatibility constraints to characterize, for a given intercept $t(\underline{\theta})$, the money-burning schedule associated with a non-decreasing allocation (Myerson (1981)). The proof is in the companion theory paper, Sublet (2026).

Lemma 1 (Incentive compatible allocations)**.** *An allocation $g(\cdot)$ is incentive compatible given a money-burning schedule $t(\cdot)$ if and only if $g(\cdot)$ is non-decreasing and*

$$t(\theta) = t(\underline{\theta}) + \frac{\theta}{\beta} U(g(\theta)) + W(T - g(\theta)) - \frac{\theta}{\beta} U(g(\underline{\theta})) - W(T - g(\underline{\theta})) - \frac{1}{\beta} \int_{\underline{\theta}}^{\theta} U(g(\tilde{\theta}))\, d\tilde{\theta}. \quad (5)$$

[7] Atkeson and Lucas (1992) study the case with transfers but without present bias, and Galperti (2015) studies the case with transfers and present bias.

Lemma 1 is useful in substituting the money-burning schedule with a function of the allocation in the objective of program (4). The resulting objective functional reads as follows:

$$\int_{\Theta}\Big[(1-\beta)W(T-g(\theta))+(1-\tfrac{\rho}{\beta})\left(\theta U(g(\theta))+\beta W(T-g(\theta))\right)+\tfrac{\rho}{\beta}\tfrac{1-F(\theta)}{f(\theta)}U(g(\theta))\Big]\,dF(\theta) \qquad (6)$$
$$-\tfrac{\rho}{\beta}\big(t(\underline{\theta})-\underline{\theta}\;U(g(\underline{\theta}))-\beta W(T-g(\underline{\theta}))\big).$$

The first-order approach in this context consists of maximizing the objective (6) point-wise for $\theta > \underline{\theta}$ while ignoring three constraints (the usual monotonicity condition, the non-negativity constraint on money burning, and the non-negativity constraint on spending). Rearranging the first-order condition to express it as a wedge suggests the following definition.

Definition (foa-wedge)**.** *Let $\rho \in [0,1)$. Define the foa-wedge as follows:*

$$\Delta_n(\theta)=\frac{1}{1-\rho}\left((1-\beta)-\rho\frac{1-F(\theta)}{\theta f(\theta)}\right). \qquad (7)$$

The foa-wedge is the product of a scaling factor and a term capturing the key trade-off for on-equilibrium penalties. The scaling factor captures society's aversion to on-equilibrium penalties. It is the inverse of the degree of asymmetry in the cost of on-equilibrium penalties and ranges between 1 and infinity. It takes its smallest value for the largest degree of asymmetry. Indeed, for $\rho = 0$, society does not bear any cost associated with on-equilibrium penalties. In contrast, for symmetric penalties (i.e., $\rho = 1$), society is infinitely averse to on-equilibrium penalties, and only the sign of the term in parentheses matters.[8]

The term in parentheses in (7) balances the benefit and cost of imposing discipline with on-equilibrium penalties. The benefit of discipline is simply captured by $1-\beta$. The term $\rho\frac{1-F(\theta)}{f(\theta)}$ captures the *incentive cost* of a marginal (on-equilibrium) penalty. To be compatible with incentives, on-equilibrium penalties are cumulative in the sense that a marginal penalty on spending $g_n(\theta)$ is borne with weight ρ and probability $1-F(\theta)$ that the fiscal needs are above θ. The incentive cost is therefore governed by the elasticity of the right tail of the distribution of shocks—the rate at which the tail thins out, $\frac{\theta f(\theta)}{1-F(\theta)}$. Because the incentive cost is the inverse of this elasticity, the elasticity of the right tail gives a quantifiable measure of the incentive cost of imposing discipline and hence of the need for flexibility. This is the main insight from the

[8]This intuition echoes the bang-bang result of Halac and Yared (2022a): an optimal rule enforced by symmetric penalties necessarily relies on extreme—bang-bang—penalties. The foa-wedge (7) generalizes the function $Q(\theta) = -(1-\beta)\theta f(\theta) + 1 - F(\theta)$ studied in Halac and Yared (2022a) to study environments with asymmetric penalties.

theory that the measurement exercise below exploits (Sublet (2026)). For $\rho > 0$ and a Pareto distribution with tail parameter γ, the foa-wedge is constant because the elasticity of the tail is constant, $\frac{\theta f(\theta)}{1-F(\theta)} = \gamma$. For $\rho = 0$, the foa-wedge is also constant and equal to the degree of present bias, which is reminiscent of a Pigouvian tax aligning the government's incentives with the society's objective.

Proposition 1 (Optimal fiscal rule: Pigouvian benchmark)**.** *Suppose that $\rho = 0$. An optimal fiscal rule implements the* first-best *allocation g_{fb} defined by the constant wedge $\Delta(g_{fb}(\theta), \theta) = 1 - \beta$.*

Unlike the other propositions in this paper, this proposition follows from first principles. Without concern for the societal cost of meting out a penalty on the government, it suffices to verify that the first-best allocation is compatible with incentives and that the associated penalty schedule satisfies the non-negativity constraint on penalties. The first-best allocation, which satisfies society's Euler equation $\theta U'(g_{fb}(\theta)) = W'(T - g_{fb}(\theta))$ for $\theta \in \Theta$, is increasing and hence compatible with incentives. The marginal penalty is strictly positive for $g \geq g_{fb}(\underline{\theta})$, and for $g < g_{fb}(\underline{\theta})$, $P(g) = 0$.

Three constraints are relaxed in the first-order approach: the non-negativity constraint on government spending, the non-negativity constraint on penalties, and the monotonicity constraint. First, I use restrictions that the two non-negativity constraints place on the level of the foa-wedge to define three different degrees of present bias (relative to the incentive cost) depending on whether the foa-wedge is bigger than 1, smaller than 0, or in between 0 and 1.

The foa-wedge is bigger than 1, which is incompatible with non-negative public spending, if the *degree of present bias is high* in the sense that Assumption H holds.

Assumption H. $\rho \frac{1-F(\theta)}{\theta f(\theta)} \leq \rho - \beta$ *for* $\theta \in \Theta$*.*

If Assumption H is not satisfied at θ, denote by $g_n(\theta)$ the spending associated with the foa-wedge (i.e., $\Delta(g_n(\theta), \theta) = \Delta_n(\theta)$).

The *degree of present bias is low* at θ if the foa-wedge is smaller than 0, which is the case under the following assumption.

Assumption L. $1 - \beta \leq \rho \frac{1-F(\theta)}{\theta f(\theta)}$.

If Assumption L is not satisfied at θ, then the foa-wedge is positive and $g_n(\theta) < g_f(\theta)$. If the degree of present bias is neither high nor low at θ, then it is *intermediate* at θ.

Assumption I. $\rho - \beta < \rho\frac{1-F(\theta)}{\theta f(\theta)} < 1 - \beta$.

Assumption I implies that the foa-wedge is in between 0 and 1 at θ.

The monotonicity constraint places restrictions on the slope of the foa-wedge, which I use to define the (virtual) need for flexibility. The following measure of the slope of an allocation g is indicative of the flexibility granted at $g(\theta)$: $\frac{d}{d\theta}\theta(1-\Delta(g(\theta),\theta))\frac{1}{\beta}$. For instance, implementing the first-best allocation grants constant flexibility 1, which is lower than the constant flexibility $1/\beta$ associated with the flexible allocation. Following Myerson (1981), the term *virtual* qualifies a concept augmented by the incentive cost. I define the (virtual) *need for flexibility* as the flexibility granted by the foa-wedge as follows: $\frac{d}{d\theta}\theta(1-\Delta_n(\theta))\frac{1}{\beta}$. The foa-wedge reveals that the slope of the inverse hazard rate governs the (virtual) need for flexibility (need for flexibility for short).

Lemma 2 (Monotonicity and the need for flexibility)**.** *Suppose that $g_n(\theta)$ is well-defined for $\theta \in (\theta_*, \theta^*)$. Then, g_n is non-decreasing at $\theta \in (\theta_*, \theta^*)$ if and only if the derivative of $\rho\frac{1-F(\theta)}{f(\theta)}$ is not smaller than $\rho - \beta$.*

The companion theory paper, Sublet (2026), contains the proof. Lemma 2 shows that for on-equilibrium penalties to be part of a fiscal rule, the tail of the distribution of fiscal needs must be sufficiently thick, as measured by the slope of the inverse hazard rate. The threshold for the thickness of the tail depends on the degree of asymmetry and the degree of present bias. For $\rho = 0$, the threshold is trivially satisfied independently of the degree of present bias. For $\rho = \beta$, it suffices to check whether the inverse hazard rate is non-decreasing, which is the case if and only if $1 - F$ is log-convex. For instance, for $\rho = \beta$, the foa-wedge is a valid building block of an optimal fiscal rule for Pareto-distributed fiscal needs, not for normally distributed fiscal needs. Intuitively, the thick tail associated with Pareto-distributed fiscal needs calls for a blend of flexibility and discipline that only on-equilibrium penalties can achieve. The (virtual) *need for flexibility is high* at θ if the derivative of $\rho\frac{1-F(\theta)}{f(\theta)}$ is not smaller than $\rho - \beta$.

2.1 The optimal fiscal rule

The foa-wedge is a building block of an optimal fiscal rule. The companion theory paper, Sublet (2026), uses global Lagrangian methods to arrange this building block with the flexible allocation and bunching into a complete characterization of the optimal fiscal rule, and to determine the conditions under which each arrangement is optimal. This section states the main result, which motivates the measurement exercise, and refers to Sublet (2026) for the proofs and the full case-by-case characterization.

The main result shows that, for a broad class of environments with asymmetric money burning, a hybrid fiscal rule featuring no penalties below a threshold and a graduated schedule of penalties conforming with the foa-wedge above the threshold up to a cap on the deficit is optimal. Two insights emerge from this characterization. The first relates to the optimal stringency of the cap. The optimal stringency of the cap depends on the balance between the need for discipline and the need for flexibility. As suggested by the foa-wedge, the thickness of the tail of the distribution of shocks to fiscal needs governs the need for flexibility. If the tail is sufficiently thin, large fiscal needs are relatively unlikely, and the optimal fiscal rule imposes a cap on the deficit. In contrast, if the tail is thick, large fiscal needs are relatively likely, and the optimal rule does not impose a cap. Moreover, for a sufficiently thick tail, the absence of a fiscal rule is optimal because the need for flexibility outweighs the need for discipline.

The second insight is specific to fiscal rules featuring on-equilibrium penalties. If the foa-wedge is positive and implements a monotonically increasing allocation, then it is optimal to exempt low levels of spending from on-equilibrium penalties. Because penalties cannot be negative, the level of the penalty schedule can only be shifted upward, which is undesirable. The exemption consists of setting the marginal penalty schedule at zero—instead of that conforming with the foa-wedge—below a threshold as a partial substitute for shifting the level downward. The benefit of the exemption is that it lowers the level of the penalty schedule above the exemption threshold while preserving the same discipline above the threshold. The exemption causes a kink in the penalty schedule, which is in contrast with the notch in the penalty schedule of the Stability and Growth Pact.

The common thread of these insights is that the elasticity of the right tail of the distribution of shocks to fiscal needs gives a quantifiable measure of the incentive cost of imposing discipline and hence of the need for flexibility (Sublet (2026)). This elasticity is therefore a key input to the design of an optimal fiscal rule, which is the motivation for the measurement exercise developed in the rest of the paper.

3 Measurement

The need for flexibility, governed by the elasticity of the right tail of the distribution of shocks, is a rich object that is not directly observable. This section introduces a two-step method to measure the need for flexibility. A first step uses a positive model to infer the unobserved past fiscal needs from government finance data. A second step uses tools from heavy-tail analysis to infer the thickness of the right tail of the distribution of fiscal needs measured in the first step.

3.1 A tractable positive model of government deficit

In this section, I enrich the normative model of Section 1 to obtain a positive model from which unobserved fiscal needs are exactly identified from government finance data.

I enrich the model by fleshing out the dynamics, and, because the yearly frequency of government finance data tends to be shorter than the cyclicality of government budgets, I allow for shocks to the fiscal needs to be persistent.[9] The government budget constraint governs the dynamics of the debt, $g = T - Rb + b'$. The problem of the government is

$$\begin{aligned} V(w,\theta) = \max_{g,b' \in \mathbb{R}_+ \times (-\infty,\bar{b}]} \quad & \theta U(g) \ + \ \beta\delta E\big[V(w',\theta')|\theta\big] \\ \text{s.t.} \quad w &= g - b', \\ w' &= T - Rb', \\ \theta' &= \theta_e + \varphi\theta + \epsilon', \ \text{and } \epsilon' \sim F, \end{aligned} \tag{8}$$

where, for analytical tractability, the borrowing capacity $\bar{b}$ is the natural borrowing limit, and $-1 < \varphi < 1$.

The taste shock θ is a catchall for all sources of fluctuations in fiscal needs, irrespectively of whether it originates in a fluctuation in spending needs, in the cost of servicing the debt,

[9] Halac and Yared (2014) show that in a world with persistent shocks, although the sequentially optimal fiscal rule is static, the ex ante optimal fiscal rule is dynamic. Because the normative analysis above abstracts from issues of persistence in the shocks, it restricts attention to sequentially optimal fiscal rules.

or in revenue. Hence, government revenue is constant in the model precisely to attribute any fluctuation in government revenue in the data to a fluctuation in measured fiscal needs θ.[10] This approach has the advantage of yielding a tractable model in the spirit of recent developments in the structural identification of uninsurable shocks in the labor market and at home (e.g., Heathcote, Storesletten, and Violante (2014) and Boerma and Karabarbounis (2022)).

Given that shocks may be serially correlated, I assume $U(g) = \ln(g)$ to preserve analytical tractability (more on the role of the elasticity of intertemporal substitution below). Also, although the formulation above assumes a constant gross interest rate R for simplicity, it is without loss of generality with log utility (see Barro (1999)).

Lastly, to simplify the exposition, the government discounts the future geometrically instead of quasi-hyperbolically. This is without loss for the positive model because, with log preferences, a hyperbolic discounter behaves like a geometric discounter (see Barro (1999)). The discount factor of the society is δ, whereas the present-biased government has discount factor $\beta\delta$.

At this stage, the task of using a dynamic model to infer the distribution of shocks may appear intractable because of the simultaneity in determining the value function and the distribution of spending needs. This is precisely the motivation to keep the model tractable. The model admits an analytical solution because, as shown below, the value function depends only on the first moment of the distribution of taste shocks, which can conveniently be normalized because taste shocks are in utils, $\theta_e = (1 - \beta\delta)(1 - \varphi\beta\delta)$.

The solution method is to guess and verify the solution. The guess is

$$V(w, \theta) = a(\theta) \ln(w + \bar{b}) + \nu(\theta), \tag{9}$$

where $a(\theta)$ and $\nu(\theta)$ capture the dependence of the continuation value on the realized fiscal need and on the first moment of the distribution of fiscal needs. In turn, the guess for the policy functions is linear in effective wealth,

$$g(w, \theta) = (1 - s(\theta)) \; (w + \bar{b}), \tag{10}$$

and $b'(w, \theta) = -s(\theta) \; (w + \bar{b}) + \bar{b}$, where $s(\theta)$ denotes the savings rate. Note that the savings rate $s(\theta)$ refers to the ratio of the stock of unused borrowing capacity $\bar{b} - b'(w, \theta)$ over the stock of

[10] This interpretation finds support in Section 5.4 in Amador, Werning, and Angeletos (2006). With a CARA utility index $U(g) = e^{-\alpha g}$, the mapping of additive shocks to the government revenue $\tilde{T}$ into taste shocks has a closed-form expression: $\theta = e^{-\alpha \tilde{T}}$. With a log utility index, however, the mapping is not in closed form.

effective wealth $w + \bar{b}$ (unlike other common definitions of the savings rate based on the ratio of the flow of savings over the flow of income).

Proposition 2. *The value function* (9) *satisfies the Bellman equation. The policy function* (10) *solves the government problem with the value function* (9), *and the savings rate is*

$$s(\theta) = \frac{\beta\delta\left(1 + \frac{\varphi}{1-\varphi\beta\delta}\theta\right)}{\theta + \beta\delta\left(1 + \frac{\varphi}{1-\varphi\beta\delta}\theta\right)}. \tag{11}$$

The proof is in Appendix C.1. The formula for the savings rate is only a function of the discount factors, the persistence in the distribution of fiscal needs, and, importantly, of the realized fiscal needs. In response to a persistent shock, the government behaves as if it was more patient than it would be if shocks were iid. As a result, the savings rate is higher, and, more importantly, the θ-elasticity of the savings rate is lower with persistent shocks.

Corollary 1 (Elasticity of savings rate and persistence)**.** *The elasticity of the government's savings rate to fiscal needs is lower with persistent shocks than that with iid shocks,*

$$\left|\frac{\partial \ln s(\theta; \varphi > 0)}{\partial \ln(\theta)}\right| \leq \left|\frac{\partial \ln s(\theta; \varphi = 0)}{\partial \ln(\theta)}\right|.$$

Corollary 1 implies that abstracting from the persistence in the shocks would give a conservative measure of the variation in fiscal needs as a function of the variation in the savings rate.[11]

Exact identification of the need for public spending. The key to the identification of fiscal needs obtains from inverting (11) to express θ as a function of s,

$$\theta(s) = \beta\delta\left(\frac{1-s}{s}\right)\left(\frac{1-\varphi\beta\delta}{1-\varphi\beta\delta\left(\frac{1}{s}\right)}\right). \tag{12}$$

Equation (12) identifies the unobserved fiscal needs that rationalize the savings rates observed in government finance data.

The degree of persistence φ used for the measurement ought to be consistent with the autocorrelation in the measured fiscal needs. I let the data determine φ by finding a fixed point of the function that maps φ to the estimated autocorrelation in the measured fiscal needs with φ

[11] Appendix D contains a sensitivity analysis of the measurement with respect to the persistence of the shocks.

(iterating over φ worked well in the application below). The persistence introduces a distinction between the measured fiscal needs θ and the shock to the fiscal needs ϵ. I use the residuals $\hat{\epsilon}$ from fitting an AR(1) process on the measured fiscal needs as an estimate of the shocks.

3.2 Measuring the thickness of $1 - F$

In theory, kernel density estimation is a natural approach to infer the distribution from a sample. It amounts to smoothing out the histogram by placing a kernel density on each data point and summing them up to obtain the estimated density. In practice, however, because government finance datasets are finite samples containing rarely occurring large observations—say, during recessions—the kernel density estimator performs poorly in this context. A complementary approach draws from heavy-tail analysis to infer whether the tail of the distribution behaves asymptotically like a power function. That is, $1 - F$ is *heavy* if there exists $\gamma > 0$, the exponent of variation, such that $\lim_{t \to \infty} \frac{1-F(\theta t)}{1-F(t)} = \theta^{-\gamma}$. The inverse of the exponent of variation governs the tail thickness at the top (i.e., for $\theta \to \infty$). Interestingly, estimation methods from heavy-tail analysis are also informative about the behavior of the tail below the top (i.e., for θ above some threshold). The Pareto distribution is a notable example of a heavy tail because the tail is a power function $1 - F(\theta) = \theta^{-\gamma}$.

I use two methods to investigate the behavior of $1 - F$. The appeal of the first method—the tail empirical distribution—is that it is graphical and intuitive. The appeal of the second method—the Hill estimator—is that it is well-suited for serially correlated observations. A plot of the tail empirical distribution depicts the log of the rank on the log of the size of the observations. Intuitively, the motivation comes from taking the log on both sides of $1 - F(\theta) = \theta^{-\gamma}$, which gives a linear relationship whose slope is the tail exponent γ. The tail empirical distribution plots

$$\{\ln(\theta_{j:N}), \ln(1 - \hat{F}(\theta_{j:N})), j = 1, \dots N\}, \tag{13}$$

where $\theta_{j:N}$ refers to the j^{th}-order statistics from $\theta_{1:N} \leq \theta_{2:N} \leq \dots \theta_{N:N}$, and $\hat{F}(\theta_{j:N}) = \frac{j}{N}$ denotes the rank of the observation.

If the plot of the log of the rank on the log of the size of the measured θ depicts a linear relationship, the tail exponent can be estimated by OLS (see Gabaix and Ibragimov (2011)). Using the closed-form expression (12) makes the identification of the thickness of the tail from government finance data on savings rate $(s_t)_{t=1}^N$ transparent,

$$\ln(1-\hat{F}(\theta_{j:N})) = \text{constant} - \gamma \ln\left(\frac{1-s_{j:N}}{s_{j:N}}\right) + \gamma \ln\left(1 - \hat{\varphi}\beta\delta \frac{1}{s_{j:N}}\right), \tag{14}$$

which shows that for independent data, variations in the savings rate alone identify γ. The discount factors matter for the evaluation of the estimation of γ only for the term correcting for the serial dependence in fiscal needs.

The role of the elasticity of intertemporal substitution (EIS) is hidden in (14) because with a log utility index, the EIS equals 1. Appendix D shows that the estimation identifies $\frac{\gamma}{EIS}$. Hence, the measured tail thickness is inversely related to the EIS. Intuitively, with a smaller EIS, a larger variation in fiscal needs rationalizes the observed variation in savings rates.

The second method consists of estimating the tail parameter using the Hill estimator for serially correlated data. Intuitively, for iid Pareto-distributed data, the Hill estimator corresponds to the maximum likelihood estimator of the tail parameter. The Hill estimator remains a consistent estimator of the exponent of variation γ even if the the data are not Pareto distributed and exhibit serial correlation (see Resnick and Stărică (1995)). Although the Hill estimator is consistent irrespectively of whether the measured θ or the residual $\hat{\epsilon}$ from fitting an autoregressive process is used, for finite samples, Resnick and Stărică (1995) recommend using the residuals. The Hill estimator of $1/\gamma$ based on $N-k$ upper-order statistics is

$$H_k = \frac{1}{N-k} \sum_{t=k+1}^{N} \ln\left(\frac{\hat{\epsilon}_{t:N}}{\hat{\epsilon}_{k:N}}\right). \tag{15}$$

Computing the Hill estimator for different upper-order statistics is informative about the behavior of the tail above different thresholds.

4 Application: the case of the European Union

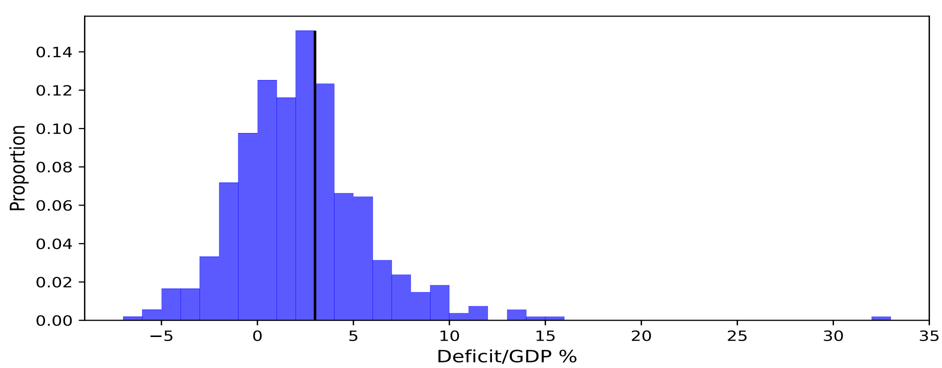
In this section, I use the two-step methodology to measure the need for flexibility of members of the European Union (EU) over the past 28 years. I find evidence of a Pareto tail of the distribution of shocks.

4.1 Government finance data

The dataset is from EuroStat. It contains data on government expenditure, revenue, debt, and interest payable on government debt for the 27 current EU members between 1995 and 2022. EuroStat harmonizes the data across EU members and computes interest payable on an accrual basis (unlike the interest payments reported by the US government (see Hall and Sargent (2011)). Because the model abstracts from default, I do not include data covering periods of high spreads for the countries concerned (Italy, Greece, Portugal, and Spain between 2007 and 2014, and Bulgaria and Romania before 2010). The resulting sample has $N = 631$ observations.

Abstracting from the Excessive Deficit Procedure (EDP) of the European fiscal rule in the positive model, despite the EDP being in place from 1995 to 2019, makes sense for two reasons. First, note that it accords with the historical evidence that EU members largely disregarded the EDP. Figure 1 shows that as much as 34% of the deficits/GDP of EU members between 1995 and 2019 exceeded the 3% threshold above which the EDP would, in theory, impose penalties.[12] Second, if the EDP in place between 1995 and 2019 did incentivize EU members to have lower government deficits, then abstracting from the EDP gives a conservative measure of the spending needs.

Figure 1: EU deficits, 1995-2019



[12]Also, although the lack of enforcement of the EDP is commonly interpreted as evidence of the non-enforceability of supra-national fiscal rules, note that it does not mean that a better-designed fiscal rule is not enforceable (see Halac and Yared (2023) for a theory of self-enforcing fiscal rules and Dovis and Kirpalani (2020, 2021) for reputational concerns in the design of a fiscal rule with limited commitment).

Application to an economic union. The measurement outlined in Section 3 applies without change to an economic union of homogeneous members. The member countries of the EU, however, are not homogeneous. Following the literature on rare events in macroeconomics, the key assumption in using the panel of measured shocks for heterogeneous members to infer the thickness of the tail of the distribution of shocks ϵ is that the distributions have the same exponent of variation (see, for instance, Barro and Jin (2011)). The member countries can, however, be heterogeneous in their other characteristics including their government revenues, the degree of present bias of their government, their cost of borrowing, the mean of their fiscal needs, and the persistence of shocks to their fiscal needs.

4.2 Step 1: measurement of past fiscal needs

The key input to measure fiscal needs is the empirical counterpart to the savings rate. Using (10), the savings rate is the complement to 1 to the spending rate, and the spending rate is the ratio of public spending as a fraction of the net present value of the government's revenue net of servicing the debt,

$$s_{it} = 1 - \frac{g_{it}}{\bar{b}_i + T_{it} - R_{it} d_{it-1}}, \tag{16}$$

where g_{it}, T_{it}, and d_{it} denote government i spending, revenue, and debt in year t. The natural borrowing limit $\bar{b}_i$ is set at $\frac{T_i}{R-1}$, where T_i and R denote averages over the sample period for the revenue of government i and the interest rate payable on the debt.[13] For a given country i, the savings rate varies over time because of variations in government spending, in the cost of servicing the debt, and in government revenues.

I obtain a panel of measured fiscal needs θ_{it} by using s_{it} from (16) in (12) with $\delta = 0.96$, and the present bias β_i is calibrated, for each country, to match the expected government spending in the model to the average government spending in the data. The average β is 0.86. The Greek government has the lowest β at 0.73 relative to the government of Luxembourg, which has the highest β at 1. Further details on the calibration and a sensitivity analysis with respect to the present bias are in Appendix D. The shocks $\hat{\epsilon}_{it}$ are the residuals from fitting country-specific AR(1) processes to the measured fiscal needs, with the location minimally shifted by adding $\min_{it} \hat{\epsilon}_{it}$ so that the shocks are positive (it is without loss because it amounts to a normalization

[13] Appendix D includes a sensitivity analysis to alternative measures of the borrowing limit.

of θ_{ei}). The shift in location does not affect the tail index. The Hill estimator, however, is not location invariant (Appendix D addresses this shortcoming of the Hill estimator with a sensitivity analysis).

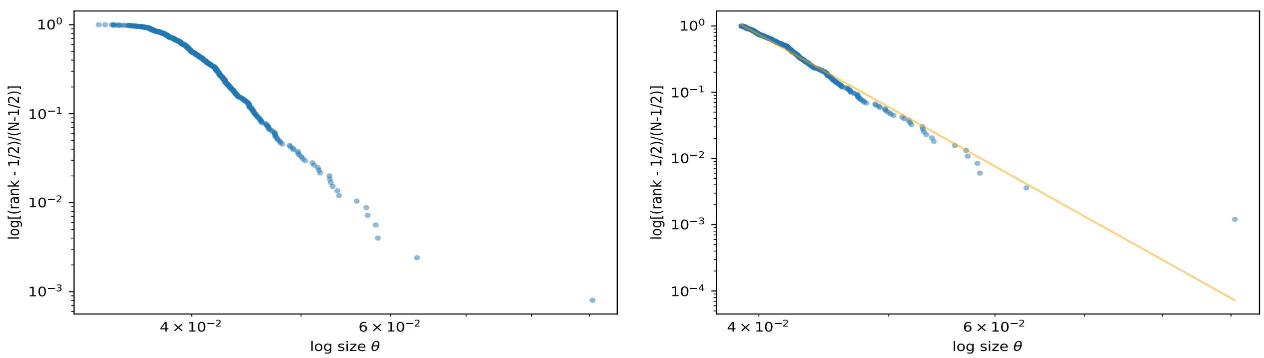
4.3 Step 2: measurement of the tail thickness

In this section, I present evidence of a thick (Pareto) tail of the distribution of shocks on the need for government spending of EU members.

Figure 2: Tail empirical distribution of fiscal needs of EU members, 1995-2022

(a) All measured fiscal needs

(b) Largest two-thirds of measured fiscal needs



The tail empirical distribution. Figure 2a depicts the tail empirical distribution (13). The correction of the rank by -1/2 reduces a finite sample bias in estimating the tail exponent (see Gabaix and Ibragimov (2011)).

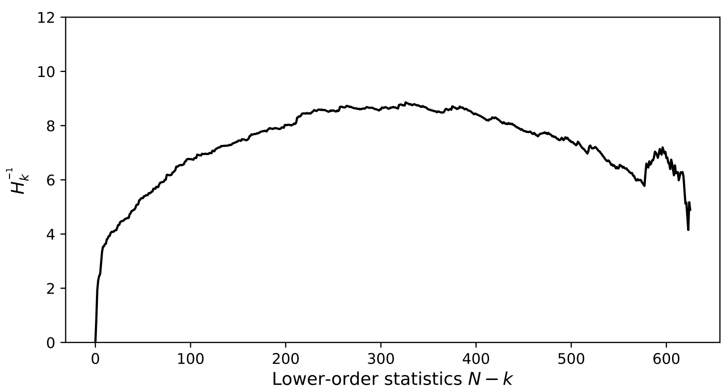
Figure 2b depicts evidence that the tail of the distribution of fiscal needs for EU members follows a Pareto distribution above a threshold. Figure 2b displays the top two-thirds of the measured fiscal needs, which amounts to truncating Figure 2a by retaining the 420 largest observations. The alignment of the log rank on log size indicates that a Pareto distribution fits the upper tail of the distribution of spending needs remarkably well.[14]

The slope of the line fitted to the top two-thirds of the observations is an estimate of the thickness of the Pareto tail parameter. The ordinary least squares estimate from a regression

[14]Figure 4 in Appendix D depicts each country's contribution to the empirical tail distribution in Figure 2.

of the log of rank corrected by a half on the log of size gives $\hat{\gamma} = 11.3$, with standard error $0.78 = \sqrt{\frac{2}{(2/3)N}}\hat{\gamma}$.

Figure 3: Hill plot for the tail of the distribution of shocks to fiscal needs of EU members



The Hill plot. Figure 3 depicts the Hill plot $\{(k, H_k^{-1}), 1 \leq k \leq N\}$, where H_k are the Hill estimates of $1/\gamma$ computed with the Hill estimator (15) and the shocks $\hat{\epsilon}$.

The Hill plot substantiates the finding from the tail empirical distribution. There is evidence of a heavy tail above a threshold, which corresponds to the top two-thirds of the observations. At the left end of Figure 3, the Hill estimator based on all the residuals indicates no evidence of a Pareto tail, which is consistent with the curvature on the left end of the tail empirical distribution. At the opposite end, the Hill estimator becomes arbitrarily imprecise because it is based on a decreasing number of observations. In between, the Hill plot stabilizes with estimates of γ in a range between 5 and 7. The lower estimate from the Hill plot relative to the estimate from the tail empirical distribution accords with the implications of Corollary 1; that is, not accounting for persistence in the shock process biases the measure of the tail thickness downward.

5 Conclusion

To conclude, I use the findings from this paper to evaluate the Excessive Deficit Procedure (EDP) of the Stability and Growth Pact and to propose avenues for reforms. For some context, the following quote reflects the current penalty schedule of the EDP:

> A non-interest-bearing deposit of 0.2% of GDP may be requested from a euro area country that is placed in EDP. [...] In case of non-compliance with the initial recommendation for corrective action, this non-interest-bearing deposit will be converted into a fine. *—European Commission, "EU Economic Governance 'Six Pack' State of Play," Memo/11/647, September 28, 2011.*

First, there is a threshold below which flexibility prevails and above which the country is placed in EDP. This paper lends support to this feature of the Stability and Growth Pact.

Second, the EDP features a jump in the level of penalties—a notch point—from 0 to a non-interest-bearing deposit of 0.2% of GDP. Initially, the penalty is the forgone interest on the deposit. A notch can be part of the optimal penalty schedule if the distribution of shocks has a sufficiently thin tail. In this case, the notch needs to be prohibitive to implement a cap. With France and Germany violating the EDP, history has taught us that the notch did not implement a cap. Besides, doubt remains about the enforceability of the penalties because the violations were left unpunished (see Dovis and Kirpalani (2020, 2021) and Halac and Yared (2022a, 2023)).

This paper suggests that the lack of enforcement may partly be due to a poor design of the penalty schedule in the EDP. Under the conditions outlined in Section 2.1, and given the novel evidence of a Pareto tail of the distribution of shocks to the fiscal needs of EU members, a fiscal rule featuring a graduated schedule of mild penalties is a promising avenue for reform. Although the EDP features some gradualism with the possibility to convert the deposit into a fine (i.e., the penalty would then be the deposit instead of the forgone interest on this deposit), the theory does not lend support to gradualism with two notches. Expressing the penalty as a percentage of the deficit above a threshold, instead of a percentage of GDP, would turn the notch into a kink or a smooth schedule.

Lastly, the global Lagrangian method is widely applicable to solve mechanism design problems with limited transfers in other contexts. Werning (2007) studies Pareto efficient income taxation, where the requirement that a reform be Pareto improving limits transfers. The method also applies to designing the cost of verifying the state to determine whether escape clauses apply (see Halac and Yared (2020b) for the design of rules with costly state verification). Another promising application is the study of the optimal illiquidity of retirement savings accounts for households who undersave for their retirement (see Laibson et al. (1998) and Beshears et al. (2020)).

References

Aguiar, M., Amador, M., Farhi, E. and Gopinath, G. (2015). Coordination and Crisis in Monetary Unions, Quarterly Journal of Economics, 130(4): 1727-1780

Aguiar, M., Amador, M., and Fourakis, S. (2020). On the Welfare Losses from External Sovereign Borrowing. IMF Economic Review, 68(1): 163-194.

Alesina, A., and Passalacqua, A. (2016). The Political Economy of Government Debt. Handbook of Macroeconomics, 2: 2599-2651. North Holland.

Alfaro, L. and Kanczuk, F. (2019). Fiscal Rules and Sovereign Default. Working paper.

Alonso, R., and Matouschek, N. (2008). Optimal Delegation. Review of Economic Studies, 75: 259-293

Amador, M., and Bagwell, K. (2022). Regulating a Monopolist With Uncertain Costs Without Transfers. Theoretical Economics, 17(4):1719-1760

Amador, M., and Bagwell, K. (2020). Money Burning in the Theory of Delegation. Games and Economic Behavior, 112: 382-412

Amador, M., Bagwell, K., and Frankel, A. (2018). A Note on Interval Delegation. Economic Theory Bulletin, 6(2): 239-249

Amador, M. and Bagwell, K. (2013). The Theory of Optimal Delegation With an Application to Tariff Caps. Econometrica, 81: 1541-1599

Amador, M., Werning, I. and Angeletos, G.-M. (2006). Commitment vs. Flexibility. Econometrica, 74: 365-396

Ambrus, A., and Egorov, G. (2017). Delegation and Nonmonetary Incentives, 171: 101-135

Ambrus, A., and Egorov, G. (2013). Comment on "Commitment vs. Flexibility", Econometrica, 81(5): 2113-2124

Athey, S., Atkeson, A., and Kehoe, P. (2005). The Optimal Degree of Discretion in Monetary Policy, Econometrica, 73(5): 1431-1475

Atkeson, A. and Lucas, R. E. (1992). On Efficient Distribution with Private Information. Review of Economic Studies, 59(3): 427-453

Azzimonti, M., Battaglini, M., and Coate, S. (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

Bassetto, M. and Sargent, T. J. (2006). Politics and Efficiency of Separating Capital and Ordinary Government Budgets. Quarterly Journal of Economics, 121(4):1167-1210

Barro, R. J. (1999). Ramsey Meets Laibson in the Neoclassical Growth Model. Quarterly Journal of Economics, 114(4): 1125-1152

Barro, R. J. and Jin, T. (2011). On the Size Distribution of Macroeconomic Disasters. Econometrica, 79(5): 1567-1589

Beetsma, R. and Uhlig, H. (1999). An Analysis of the Stability and Growth Pact, Economic Journal, 109: 546-571

Beshears, J., Choi, J. J., Clayton, C., Harris, C., Laibson, D., and Madrian, B. C. (2020). Optimal Illiquidity. Working paper.

Boerma, J., and Karabarbounis, L. (2021). Inferring Inequality with Home Production. Econometrica 89(5): 2517-2556

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

Clayton, C., and Schaab, A. (2022). A Theory of Dynamic Inflation Targets. Working paper.

Cooper, R., and Kempf, H. (2004). Overturning Mundell: Fiscal Policy in a Monetary Union. Review of Economic Studies, 71: 371-396

Diamond, P. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. American Economic Review, 88: 83-95

Dovis, A., and Kirpalani, R. (2020). Fiscal Rules, Bailouts, and Reputation in Federal Governments. American Economic Review 110(3): 860-888

Dovis, A., and Kirpalani, R. (2021). Rules without Commitment: Reputation and Incentives. Review of Economic Studies, 88(6): 2833-2856

Felli, C., Piguillem, F., and Shi, L. (2021). Fiscal Rules and Discretion with Risk of Default. Working paper.

Galperti, S. (2015). Commitment, Flexibility, and Optimal Screening of Time Inconsistency. Econometrica, 83: 1425-1465

Gabaix, X., and Ibragimov R. (2011). Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. Journal of Business Economics and Statistics. 29(1): 24-39

Hall, G. J., and Sargent, T. J. (2011). Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics. American Economic Journal: Macroeconomics, 3: 192-214

Halac, M., and Yared, P. (2014). Fiscal Rules and Discretion under Persistent Shocks. Econometrica, 82: 1557-1614

Halac, M., and Yared, P. (2018). Fiscal Rules and Discretion in a World Economy. American Economic Review, 108: 2305-2334

Halac, M., and Yared, P. (2020a). Inflation Targeting under Political Pressure. in Independence, Credibility, and Communication of Central Banking, edited by Ernesto Pastén and Ricardo Reis, Santiago, Chile: Central Bank of Chile, 7-27

Halac, M., and Yared, P. (2020b). Commitment vs. Flexibility with Costly Verification. Journal of Political Economy, 128: 4523-4573

Halac, M., and Yared, P. (2022a). Fiscal Rules and Discretion under Limited Enforcement. Econometrica, 90: 2093-2127

Halac, M., and Yared, P. (2022b). Instrument vs. Target-Based Rules. Review of Economic Studies, 89: 312-345

Halac, M., and Yared, P. (2023). A Theory of Fiscal Responsibility and Irresponsibility. *working paper*

Harstad, B. (2007). Harmonization and Side Payments in Political Cooperation. American Economic Review, 97(3): 871-889

Hatchondo, J.-C., Martinez, L., and Roch, F. (2020). Fiscal Rules and the Sovereign Default Premium. American Economic Journal: Macroeconomics, 14(4): 244-273

Heathcote, J., Storesletten, K., and Violante, G. (2014). Consumption and Labor Supply with Partial Insurance: An Analytical Framework. American Economic Review 104(7): 2075-2126

Holmström, B. (1977). On Incentives and Control in Organizations. Ph.D. thesis, Stanford Graduate School of Business

Kováč, E. and Mylovanov, T. (2009). Stochastic mechanisms in settings without monetary transfers: The regular case. Journal of Economic Theory, 144: 1373-1395

Laibson, D., I. Repetto, A., Tobacman, J., Hall, R., Gale, W. and Akerlof, G., A. (1998). Self-Control and Saving for Retirement. Brookings Papers on Economic Activity 91-196

Luenberger, D. G. (1969). *Optimization by Vector Space Methods.* Series in Decision and Control. New York: Wiley

McKibbin W. J., and Wilcoxen, P. J. (2002). The Role of Economics in Climate Change Policy. Journal of Economic Perspectives, 16(2): 107-129

Melumad, N. D., and Shibano, T. (1991). Communication in Settings with No Transfers. The Rand Journal of Economics, 22: 173-198

Myerson, R. B. (1981). Optimal Auction Design. Mathematics of Operations Research, 6: 58-73

Resnick, S. (2007). Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer-Verlag, New York

Resnick, S., and Stărică, C. (1995). Consistency of Hill's Estimator for Dependent Data. Journal of Applied Probability, 32(1): 139-167

Sublet, G. (2026). The Optimal Degree of Flexibility in Policy Rules. *working paper*

Weitzman, M. (1974). Prices vs. Quantities. Review of Economic Studies, 41(4): 477-491

Werning, I. (2007). Pareto Efficient Income Taxation. working paper.

Yared, P. (2019). Rising Government Debt: Causes and Solutions for a Decades-Old Trend. Journal of Economic Perspectives, 33: 115-140

Appendix

A Design of a fiscal rule for an economic union

This section introduces the additional notation needed to design a fiscal rule for an economic union.[15] The design a fiscal rule for an economic union that is not fiscally integrated separates into the design of country-specific penalty schedules as studied in Sections 1 and 2.

The economic union has $\mathcal{N}$ countries indexed by $i = 1, \ldots, \mathcal{N}$. Each country has its own government choosing its spending according to (2), where θ denotes the idiosyncratic shock to the country's fiscal needs.

A fiscal rule is a tuple of penalty schedules $(P_i(\cdot))_{i=1}^{\mathcal{N}}$. In the context of an economic union, the non-negativity constraint on penalties captures the lack of fiscal integration because a negative penalty could model a transfer across members. The lack of fiscal integration also implies that the budget constraints of the different members are independent due to the absence of transfers, $g + \frac{x}{R_i} = T_i$. The country-specific interest rates are assumed to be exogenous which implicitly assumes that each country is small and the economic union is small as well. Halac and Yared (2018) show that for a large economic union, however, optimal coordinated and uncoordinated fiscal rules differ due to the disciplining and redistributive effects of the endogenous response of the interest rate to the fiscal rule.

The three key determinants of an optimal fiscal rule, namely the distribution of shocks F_i, the degree of present bias $1 - \beta_i$, and the degree of asymmetry in the cost of meting out a penalty $1 - \rho_i$, can be country specific. The welfare of the economic union is the aggregation of the welfare of each country:

$$\sum_{i=1}^{\mathcal{N}} \int_{\Theta_i} \left[\theta U(g_i(\theta)) + W_i(T_i - g_i(\theta)) - \rho_i P_i(g_i(\theta))\right] dF_i(\theta),$$

where the country-specific interest rates are subsumed into the continuation values $W_i(T_i - g(\theta)) = W(R_i(T_i - g(\theta)))$. Setting equal weights on the welfare of each country is without loss

[15] A long line of research studies fiscal rules in economic unions with limited commitment (see Beetsma and Uhlig (1999), Cooper and Kempf (2004), Chari and Kehoe (2007), Chari and Kehoe (2008), Aguiar, Amador, Farhi, and Gopinath (2015), Dovis and Kirpalani (2020), and Dovis and Kirpalani (2021)). Yared (2019) surveys the literature. As mentioned in the introduction, some of the literature provides the micro-foundations for modeling the deficit bias on the part of members of an economic union with a present bias and focuses on the optimal stringency of a cap on deficit. This paper focuses on the optimal structure of a fiscal rule.

of generality precisely because, for strictly positive weights, the problem for an economic union separates into $\mathcal{N}$ country-specific problems independently of the weights.

B Solution method for the design of rules

The non-negativity constraint on the penalty schedule sets program (4) apart from mechanism design problems in which transfers are possible. Unlike the incentive compatibility constraints, the non-negativity constraint on the penalty schedule cannot be easily summarized in the objective function to be maximized point-wise without resorting to Lagrangian methods. This section outlines how I use the first-order conditions of the Lagrangian method to identify a candidate solution and to find conditions for global optimality.

The first step uses a Lagrange multiplier function to combine the non-negativity constraint on the penalty schedule and the objective in a Lagrangian. Let $\Lambda : \Theta \mapsto [0,1]$ be a non-decreasing function such that $\lim_{\theta \to \bar{\theta}} \Lambda(\theta) = 1$ and $1-\Lambda$ is integrable. A valid Lagrange multiplier function is non-decreasing, which is the infinite dimensional analog of a non-negative Lagrange multiplier for the Kuhn-Tucker theorem with finitely many inequality constraints. Define the Lagrangian, with Lagrange multiplier function Λ, as a functional on $\Phi \equiv \{(u, \underline{t}) \mid u : \Theta \mapsto \mathbb{R}_+$ is non-decreasing, and $\underline{t} \in \mathbb{R}_+\}$ as follows:

$$\mathcal{L}(u, \underline{t}|\Lambda) \equiv \int_\Theta \left[\theta u(\theta) + W(T - U^{-1}(u(\theta))) - \rho t(\theta, u, \underline{t})\right] dF(\theta) + \int_\Theta t(\theta, u, \underline{t}) d\Lambda(\theta), \tag{17}$$

where $t(\theta, u, \underline{t})$ is the schedule associated with allocation $U^{-1}(u(\cdot))$ in Lemma 1 and $t(\underline{\theta}) = \underline{t}$.

The Gateaux derivative in the direction $(h, h_t) \in \Phi$ is defined as follows:[16]

$$\partial\mathcal{L}(u, \underline{t}, h, h_t|\Lambda) \equiv \lim_{\alpha \downarrow 0} \frac{1}{\alpha}\left[\mathcal{L}(u + \alpha h, \underline{t} + \alpha h_t|\Lambda) - \mathcal{L}(u, \underline{t}|\Lambda)\right]. \tag{18}$$

The next lemma gives optimality conditions in terms of the Gateaux derivative evaluated at the solution. The optimality conditions are that the Gateaux derivative evaluated at the solution is null in the direction of the solution and non-positive in any non-decreasing direction.

Lemma 3 (Lemma of optimality)**.** *If there exists a non-decreasing $u^* \equiv U(g^*)$ and $\underline{t}^*$ in the convex cone Φ and a non-decreasing function $\Lambda^* : \Theta \mapsto [0,1]$ such that $\lim_{\theta \to \bar{\theta}} \Lambda^*(\theta) = 1$ and $1 - \Lambda^*$ is integrable, and if*

$$\partial\mathcal{L}(u^*, \underline{t}^*, u^*, \underline{t}^*|\Lambda^*) = 0, \quad \textit{and} \quad \partial\mathcal{L}(u^*, \underline{t}^*, h, h_t|\Lambda^*) \leq 0 \quad \textit{for all } (h, h_t) \in \Phi,$$

[16] The existence of the Gateaux differential follows from Lemma A.1 p. 390 of Amador, Werning, and Angeletos (2006).

then $g^* \equiv U^{-1}(u^*)$ and the associated money-burning schedule t^* characterized by (5) *with $t^*(\underline{\theta}) = \underline{t}^*$ solve the mechanism design problem* (4)*.*

The proof is an application of the global theory of constrained optimization (Chapter 8 in Luenberger (1969), Lemma 1 p.227, and Theorem 1 p.220), as used in Lemma A.2 in Amador, Werning, and Angeletos (2006) and Theorem 1 in Amador and Bagwell (2013). A part of the proof shows that the degree of concavity of the Lagrangian depends on ρ. The Lagrangian is strictly concave for $\rho \in [0, 1)$ and a non-decreasing Lagrange multiplier function (see the proof of Lemma 3). The Lagrangian is linear for $\rho = 1$.

Proof of Lemma 3. Lemma A.2 in Amador, Werning, and Angeletos (2006) implies that if the Lagrangian with Lagrange multipliers Λ^* is concave, then the equality and inequality conditions in terms of Gateaux derivatives imply that the Lagrangian is maximized at $u^*, \underline{t}^*$:

$$\mathcal{L}(u^*, \underline{t}^*|\Lambda^*) \geq \mathcal{L}(u, \underline{t}|\Lambda^*) \quad \text{for all } (u, t) \in \Phi.$$

To show the concavity of the Lagrangian with Lagrange multipliers Λ^*, it is convenient to spell out the Lagrangian (17) and factorize the non-linear terms as follows:

$$\begin{aligned}
\mathcal{L}(u, \underline{t}|\Lambda^*) \equiv & \int_\Theta \left[u(\theta) \ \left((1 - F(\theta)) - \theta \tfrac{\rho-\beta}{\rho} f(\theta)\right)\right] d\theta \\
& - \int_\Theta [u(\theta)(1 - \Lambda^*(\theta))] \, d\theta + (\underline{\theta} \ u(\underline{\theta}) - \underline{t})\Lambda^*(\underline{\theta}) + \int_\Theta [\theta \ u(\theta)] \, d\Lambda^*(\theta) \\
& + \int_\Theta \left[\beta W(T - U^{-1}(u(\theta)))\right] d\left(\tfrac{1-\rho}{\rho} F(\theta) + \Lambda^*(\theta)\right) \\
& + \beta W(T - U^{-1}(u(\underline{\theta})))\Lambda^*(\underline{\theta}).
\end{aligned}$$

The integrands for the integrals in the first two lines are linear in u. For the terms in the remaining two lines, to show that the function $u \mapsto W(T - U^{-1}(u)))$ is concave, note that the utility index U is strictly increasing and concave so its inverse U^{-1} is strictly increasing and convex ($U^{-1'}(U(x)) = 1/U'(x)$ and $U^{-1''}(U(x)) = -U^{-1'}(U(x))U''(x)/U'(x)^2$). Since W is increasing and concave and $-U^{-1}$ is concave, the composition $u \mapsto W(T - U^{-1}(u))$ is concave. A sufficient condition for the Lagrangian to be concave is that the function $\frac{1-\rho}{\rho} F(\theta) + \Lambda^*(\theta)$ be non-decreasing, which is the case since $0 \leq \rho \leq 1$ and F and Λ^* are both non-decreasing.

It remains to show that the maximizer of a concave Lagrangian at a valid Lagrange multiplier is the solution to the constrained optimization problem of interest. This is precisely what the global theory of constrained optimization does for us.

The following notation maps the environment studied in this paper to Theorem 1 in Amador and Bagwell (2013) p.1575: $X = \{u, \underline{t} \mid u : \Theta \mapsto \mathbb{R}, \ \underline{t} \in \mathbb{R}\}$, $Z = \{z \mid z : \Theta \mapsto \mathbb{R}\}$ with norm $||z||= \sup_{\theta \in \Theta}|z(\theta)|$, $\Omega = \{(u, \underline{t}) \in X| \ u \text{ is non-decreasing}, \ \underline{t} \geq 0\}$, and $P = \{z \in Z \mid z(\theta) \geq$

0 for $\theta \in \Theta\}$. The objective is a functional $f : \Omega \mapsto \mathbb{R}$ defined as follows:

$$f(u, \underline{t}) = -\int_{\Theta} \left[u(\theta)\rho\frac{1-F(\theta)}{f(\theta)} + \beta(1-\beta)W(T - U^{-1}(u(\theta))) \right] dF(\theta)$$
$$- (\rho - \beta) \int_{\Theta} \left[\theta u(\theta) + \beta W(T - U^{-1}(u(\theta))) \right] dF(\theta).$$

The constraints on limited transfers are defined as follows: $G : \Omega \mapsto Z$,

$$G(u, \underline{t}) = -\left(\underline{t} + \theta\, u(\theta) + \beta W(T - U^{-1}(u(\theta))) - \underline{\theta}\, u(\underline{\theta}) - \beta W(T - U^{-1}(u(\underline{\theta}))) - \int_{\underline{\theta}}^{\theta} u(\tilde{\theta})\, d\tilde{\theta} \right),$$

and their contributions to the Lagrangian are given by $T : Z \mapsto \mathbb{R}$,

$$T(z) = \int_{\Theta} z(\theta) d\Lambda^*(\theta),$$

which satisfies $T(z) \geq 0$ for all $z \in P$ since Λ^* is non-decreasing. Since $\mathcal{L}(u|\Lambda^*) = -f(u) - T(G(u))$, Theorem 1 from Amador and Bagwell (2013) implies that $(u^*, \underline{t}^*)$ solves

$$\min_{(u,\underline{t}) \in \Omega} \{ f(u, \underline{t}) | -G(u, \underline{t}) \in P \}.$$

Inverting the above mapping from the environment of this paper to Theorem 1 in Amador and Bagwell (2013) and using $t(\cdot)$ defined in (5) as a function of $g(\cdot)$, the allocation $g^* = U^{-1}(u^*)$ and the initial level $\underline{t}^*$ solve the optimization problem:

$$\max_{g \in \Omega,\ \underline{t} \geq 0} \int_{\Theta} \left[\theta U(g(\theta)) + W(T - g(\theta)) - \rho t(\theta) \right] dF(\theta)$$

s.t. for all $\theta \in \Theta$:

$$\beta t(\theta) = \beta \underline{t} + \theta U(g(\theta)) + \beta W(T - g(\theta)) - \underline{\theta} U(g(\underline{\theta})) - \beta W(T - g(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} U(g(\tilde{\theta})) d\tilde{\theta}$$

g is non-decreasing

$t(\theta) \geq 0$.

The characterization of incentive compatible allocations in Lemma 1 implies that (g^*, t^*) in which

$$\beta t^*(\theta) \equiv \beta \underline{t}^* + \theta\, U(g^*(\theta)) + \beta W(T - g^*(\theta)) - \underline{\theta}\, U(g^*(\underline{\theta})) - \beta W(T - g^*(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} U(g^*(\tilde{\theta}))\, d\tilde{\theta}$$

solve the mechanism design problem (4). □

B.1 Sufficient optimality conditions for an unbounded support

The Lemma of optimality 3 is stated for a support Θ whose supremum $\bar{\theta}$ may be infinite. It requires only that the first moment of F be finite, that $\lim_{\theta \to \bar{\theta}} \Lambda(\theta) = 1$, and that $1 - \Lambda$ be integrable. This section makes explicit why these conditions, which are slack on a compact support, are exactly what is needed to extend the global optimization argument to a distribution with unbounded support. This extension matters for the measurement exercise because the evidence below points to a Pareto—hence unbounded-support—distribution of fiscal needs.

Two steps in the argument rely on these conditions when $\bar{\theta} = \infty$. The first is the substitution of the money-burning schedule into the objective using Lemma 1, which yields the objective (6). The substitution integrates the envelope term by parts, and the boundary term at the top vanishes because the first moment is finite, so that $\lim_{\theta \to \bar{\theta}} \theta(1 - F(\theta)) = 0$, and because $1 - \Lambda$ is integrable, so that the term $\int_\Theta u(\theta)(1 - \Lambda(\theta))\, d\theta$ in the Lagrangian (17) is well defined. The second step is the exchange of the limit defining the Gateaux derivative (18) with the integral over Θ. For a compact support this exchange is immediate because the integrand is bounded. For $\bar{\theta} = \infty$, the dominated convergence theorem applies, with the candidate solution—which is integrable because the first moment of F is finite and $1 - \Lambda$ is integrable—serving as the dominating function.

Hence the sufficient optimality conditions of Lemma 3, and therefore the characterization of optimal fiscal rules, hold for distributions with unbounded support so long as the first moment is finite. This is the sense in which the optimal fiscal rule, and the role of the thickness of the right tail of the distribution of shocks in shaping it, are well defined for the Pareto distributions documented in the measurement exercise.

C Proofs

C.1 Proposition 2 on the measurement of fiscal needs

Proof of Proposition 2. To verify that the guess is a solution, it suffices to check that the value function is a fixed point of the Bellman equation and that the policy functions solve the optimization problem with the value function.

First, to check that the policy functions solve the optimization problem with the continuation value (9), take the first-order condition to get $\theta(w + b')^{-1} = \beta\delta R E[a(\theta')|\theta]\big(T - Rb' + \bar{b}\big)^{-1}$. Substituting the policy function (10) for spending on the left-hand side and the policy function for borrowing on the right-hand side gives $\theta \left((1 - s(\theta))\ (w + \bar{b})\right)^{-1} = \beta\delta R E[a(\theta')|\theta](s(\theta)(w+\bar{b})R)^{-1}$. The terms involving the effective wealth $w+\bar{b}$ cancel out. Rearranging terms to isolate the savings rate gives $s(\theta) = \frac{\beta\delta E[a(\theta')|\theta]}{\theta+\beta\delta E[a(\theta')|\theta]}$.

Having verified that the policy functions (10) with the savings rate (11) solve the maximization problem in the Bellman equation given the value function (9), it only remains to verify that the value function solves the Bellman equation.

To check that the value function is a fixed point of the Bellman equation, substitute the value function (9) on both sides of the Bellman equation (8) and, after substitution of the policy functions, the terms involving the effective wealth $w + \bar{b}$ cancel out (because $a(\theta) = \theta + \beta\delta E\left[a(\theta')|\theta\right]$) to give $\nu(\theta) = \theta \ln(1 - s(\theta)) \ + \beta\delta E\left[a(\theta')|\theta\right] \ln\big((s(\theta)R)\big) + \beta\delta E[\nu(\theta')|\theta]$. The term $\nu(\theta)$ solves the following recursion:

$$\nu(\theta) = \theta \ln\left(\frac{\theta}{\theta + \beta\delta E\left[a(\theta')|\theta\right]}\right) \ + \beta\delta E\left[a(\theta')|\theta\right] \ln\left(\frac{\beta\delta E\left[a(\theta')|\theta\right]}{\theta + \beta\delta E\left[a(\theta')|\theta\right]}\right) + \beta\delta \ln\big(R\big) + \beta\delta E[\nu(\theta')].$$

I guess that the conditional expectation is affine in θ (and verify the guess below), $E\left[a(\theta')|\theta\right] = a_0 + a_1\theta$. Then, $a(\theta) = \theta(1+\beta\delta a_1) + \beta\delta a_0$. To verify the guess that the conditional expectation of $a(\cdot)$ is affine, substituting in the expectation, $E\left[a(\theta')|\theta\right] = \theta_e(1+\beta\delta a_1) + \beta\delta a_0 + \varphi(1+\beta\delta a_1)\theta$, which confirms that the conditional expectation is affine and the parameters are $a_1 = \varphi/(1-\varphi\beta\delta)$ and $a_0 = 1$, where the last equality follows from the normalization. Substituting the linear conditional expectation simplifies the formula for the savings rate $s(\cdot)$ which gives (11).

This confirms that the value function (9) with $a(\theta)$ solves the Bellman equation. □

D Sensitivity analysis of the measurement

In this section, I study the sensitivity of the measurement to the elasticity of intertemporal substitution, the present bias of the governments, the persistence of shocks to fiscal needs, the borrowing limit, and the location parameter for the residuals.

Elasticity of Intertemporal Substitution. To elicit the role of the Elasticity of Intertemporal Substitution (EIS) in the measurement, I relax the assumption of a unit EIS and, to preserve analytical tractability, I assume that the interest rate is constant. Let the utility index be CRRA with EIS $1/\eta$.

Proposition 3. *The value function* $V(w,\theta) = a(\theta)\frac{\left(w+\bar{b}\right)^{1-\eta}}{1-\eta}$ *satisfies the Bellman equation* (8) *for* $a(\theta) = (\theta^{1/\eta} + (\beta\delta R^{1-\eta})^{1/\eta})^{\eta}$. *The policy function* (10) *solves the government problem with the value function and the savings rate is only a function of the intertemporal elasticity of substitution, the interest rate, the discount factor, and the realized fiscal needs as follows:*

$$s(\theta) = \frac{(\beta\delta)^{\frac{1}{\eta}} R^{\frac{1-\eta}{\eta}}}{\theta^{\frac{1}{\eta}} + (\beta\delta)^{\frac{1}{\eta}} R^{\frac{1-\eta}{\eta}}}. \tag{19}$$

Proof of Proposition 3. Note that $E[a(\theta)] = E[(\theta^{1/\eta} + (\beta\delta R^{1-\eta})^{1/\eta})^{\eta}] = 1$, where the first inequality uses the guess for $a(\theta)$ and the second equality follows from the normalization of the public spending needs.

First, to check that the policy functions solve the optimization problem with the continuation value V, take the first-order condition to get $\theta(w+b')^{-\eta} = \beta\delta RE[a(\theta)]\left(T-Rb'+\bar{b}\right)^{-\eta}$. Substituting the policy function (10), $\theta\left((1-s(\theta))\ (w+\bar{b})\right)^{-\eta} = \beta\delta RE[a(\theta)](s(\theta)(w+\bar{b})R)^{-\eta}$. Rearranging terms to isolate the savings rate gives (19) because $E[a(\theta)] = 1$.

To check that the value function is a fixed point of the Bellman equation, substitute the value function on both sides of the Bellman equation (8) to get, after substituting the policy function,

$$a(\theta)\frac{\left((w+\bar{b})\right)^{1-\eta}}{1-\eta} = \theta\frac{((1-s(\theta))\ (w+\bar{b}))^{1-\eta}}{1-\eta} + \beta\delta E\left[a(\theta')\right]\frac{\left((s(\theta)\ (w+\bar{b})R-\bar{b}+\bar{b})\right)^{1-\eta}}{1-\eta}.$$

Substituting the saving rate gives and $a(\theta) = \left(\theta^{\frac{1}{\eta}} + R^{\frac{1-\eta}{\eta}}(\beta\delta)^{\frac{1}{\eta}}E\left[a(\theta')\right]^{\frac{1}{\eta}}\right)^{\eta}$, as claimed. □

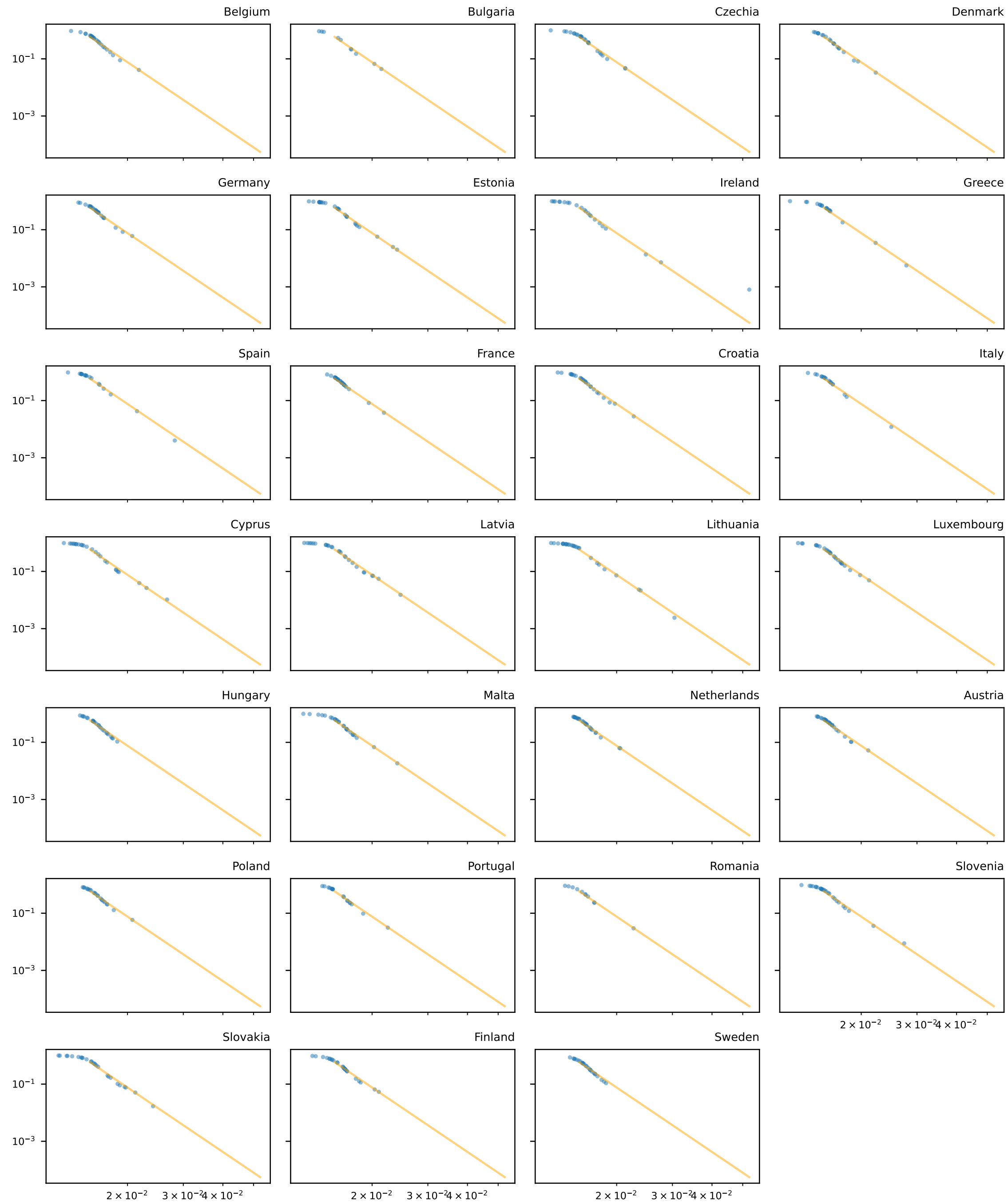
Rewriting (19) to express the fiscal needs as a function of the saving rate gives $\theta(s) = \beta\delta\left(\frac{1-s}{s}\right)^{\eta}R^{1-\eta}$. The empirical framework identifies the product of the tail parameter and the inverse of the elasticity of intertemporal substitution from variations in the savings rate across countries and time,

$$\ln(1-F(\theta(s_{it}))) = -\gamma\eta\ln\left(\frac{1-s_{it}}{s_{it}}\right) - \gamma(1-\eta)\ln(R) - \gamma\ln(\beta\delta).$$

Hence, the measured tail parameter γ is inversely related to η. The intuition is the following. A larger elasticity of intertemporal substitution—a lower η—implies that a smaller variation in fiscal needs rationalizes the observed saving rates, and hence the tail is thinner (i.e., a larger γ).

The degrees of present bias of the governments. The calibration of the degree of present bias is such that, each country's average spending in the model matches the average spending in the data. Because the model does not separately identify δ, β, and the average fiscal needs, I set the discount factor of the society $\delta = 0.96$ and normalize the level of the average fiscal needs such that the country with the highest estimated β, namely Luxembourg, has a $\beta = 1$. This calibration attributes heterogeneity in average government spending across countries to heterogeneity in the degrees of present bias of the governments. The calibrated discount factors β are, 0.73 for Greece, 0.77 for Italy, 0.79 for Portugal, 0.80 for Hungary, 0.81 for Romania, 0.81 for Malta, 0.81 for Cyprus, 0.82 for Spain, 0.82 for Slovakia, 0.83 for Belgium, 0.83 for Ireland, 0.83 for France, 0.83 for Poland, 0.85 for Croatia, 0.85 for Austria, 0.86 for Slovenia, 0.87 for Germany, 0.87 for Czechia, 0.87 for Lithuania, 0.88 for the Netherlands, 0.88 for Latvia, 0.91 for Bulgaria, 0.92 for Finland, 0.93 for Sweden, 0.97 for Denmark, 0.97 for Estonia, and 1 for Luxembourg.

Figure 4: Tail empirical distribution by country

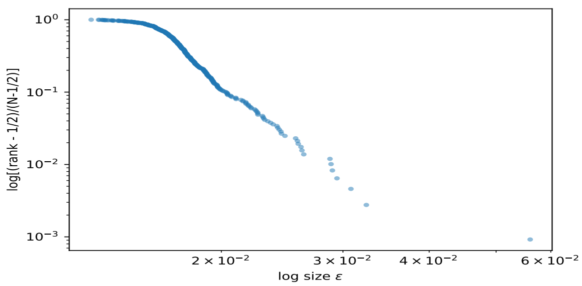


Notes: This figure decomposes the tail empirical distribution depicted in Figure 2a by country. The orange lines are the same as the orange line in Figure 2b.

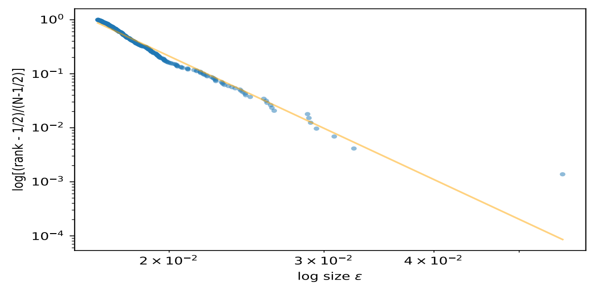
To elicit the sensitivity of the measurement to heterogeneity in the degrees of present bias, consider an alternative calibration in which the governments have the same degree of present bias (set at the average of the estimates above). This alternative calibration attributes heterogeneity in average government spending across countries to heterogeneity in the average fiscal needs of the countries. Figures 5a and 5b show that the the tail of the distribution of shocks to fiscal needs also display evidence of a heavy tail with thickness $\hat{\gamma} = 7.59$ and a standard error of 0.56.

Figure 5: Sensitivity analysis of the tail empirical distribution

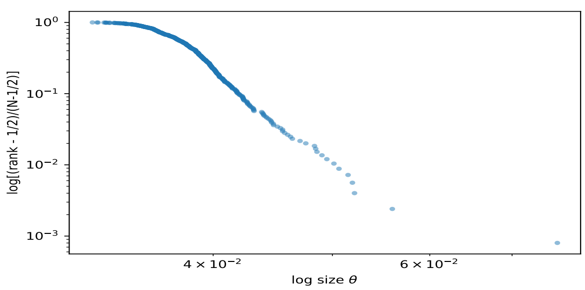
(a) All measured shocks to fiscal needs ($\beta_i = \beta_j$)



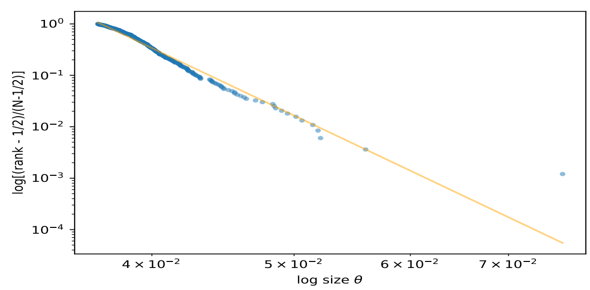
(b) Top two-thirds of shocks to fiscal needs ($\beta_i = \beta_j$)



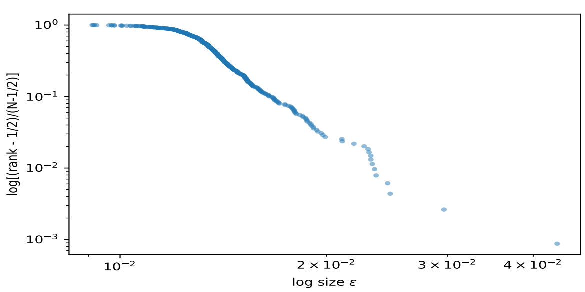
(c) All measured fiscal needs ($\varphi = 0$)



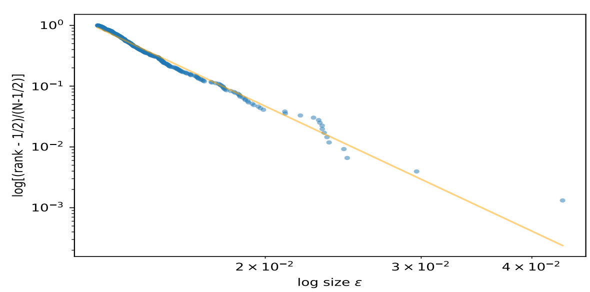
(d) Top two-thirds of fiscal needs ($\varphi = 0$)



(e) All measured shocks to fiscal needs ($R_i \neq R_j$)



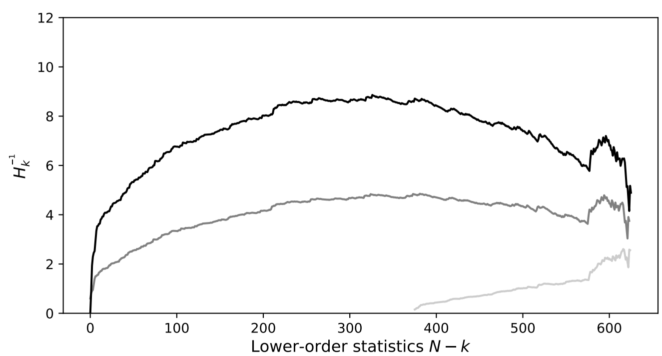
(f) Top two-thirds of shocks to fiscal needs ($R_i \neq R_j$)



Persistence of shocks. The analysis in the main text jointly infers the fiscal needs and the associated persistence of the fiscal needs. Corollary 1 shows that persistence in the shocks to the fiscal needs amplifies the variation in the measured fiscal needs. To evaluate the sensitivity of the results to this amplification mechanism, Figures 5c and 5d depict the tail empirical distribution of fiscal needs measured with $\varphi = 0$. In line with Corollary 1, not modeling the persistence in the shock biases the measurement of the tail thickness downward, $\hat{\gamma} = 13.54$ and a standard error of 0.94.

Borrowing limit. The analysis in the main text computes each government's borrowing limit based on the average government revenues for each country and on the average interest rate for the EU. Because small differences in interest rates can lead to large differences in the discounted value of future cash flows, Figures 5e and 5f depict the tail empirical distribution based on a measure of the borrowing limit that uses each countries average interest rate instead of the EU average. The estimate is $\hat{\gamma} = 6.83$, with a standard error of 0.50.

Figure 6: Sensitivity of the Hill estimates to the location of the distribution



Notes: The black line depicts the same Hill plot as in Figure 3. The light grey line depicts the Hill plot for the residuals centered around 0. The dark grey line depicts the Hill plot for the residuals with location shifted by half of the shift of location for the black line.

Location shifter of the residuals. In the main text, because the Hill estimator uses the log of the shocks, the residuals are minimally shifted to have all estimates of the shocks be positive. Although this does not affect the thickness of the tail of a distribution in theory, it does affect the Hill estimator. Figure 6 shows that the analysis in the main text gives a conservative estimate of the tail thickness.