Fencing Off Part of the Budget: Expiring Individual Designation Rights and the Pay-As-You-Go Margin

Zhang Bin, Independent Scholar[1]

August 2026

Abstract

Long-term deficits are studied through institutions written on the government—balanced-budget rules, tax and expenditure limits, supermajority requirements for taxes or borrowing. They name totals, and their enforcement is assigned to courts, procedures, and politics. This paper studies a fiscal institution written on the citizen and asks what it does to the deficit margin. Under Personal Tax Holding (PTH), every citizen may direct a fixed fraction α of her own tax liability to certified public-goods organizations before a deadline; the right renews with each fiscal year, and accounts not directed expire—they lapse to the treasury under a rule fixed in advance. Because unexercised claims expire, the designated share of the budget is pay-as-you-go by an accounting identity: no debt can accumulate on it, and α is the width of the budget withdrawn from deficit finance. Scaled against the Congressional Budget Office's February 2026 baseline for fiscal year 2027, that width is $29 billion, or 0.09 percent of GDP, at the one percent of Hungary's percentage-philanthropy statute, and $1,474 billion, or 4.4 percent of GDP—four fifths of the projected deficit—at one half of individual income tax liability, where the revenue base left to carry cyclical shocks falls from 17.7 to 13.2 percent of GDP. Aggregate lapsing is an observable statistic of the demand for public expenditure, but a rule that cuts next year's levy by realized lapsing extinguishes the levy in every admissible environment, because lapsing conflates absent demand with search friction and inattention; the friction-adjusted rule contracts only on evidence no friction can produce. And the coalition of never-designating citizens that enforces the constitutional ceiling—sustained by fiscal privacy, with a computable opacity threshold—constrains debt only where the ceiling is written on revenue plus net borrowing; the American state record on debt limits and tax limitation is the record of what the narrower clause invites. Because a fiscal constitution is written once, behind a veil of ignorance about the populations it will govern, we evaluate each rule by what it guarantees against every distribution of citizens' participation costs the framer cannot rule out; the same criterion fixes the deadline (fixed and interior), the lapse rule (constant marginal rate), and certification (margin caps), and it is the composition margin—the heterogeneity that survives Tiebout sorting because sorting operates on bundles, not on their composition—that the mechanism was first built

---

[1] Prepared for the NBER conference "Long-Term Budget Deficits: Economic Effects and Potential Responses," Washington, DC, May 20–21, 2027. In keeping with NBER guidelines, the paper characterizes the properties of alternative fiscal rules and makes no policy recommendations. This draft integrates and supersedes two earlier manuscripts by the author: a formal treatment of Personal Tax Holding, and an informal companion essay on individually held taxes from which the annual expiration rule, the menu of lapse destinations, and the revealed-demand adjustment of the levy descend.

to reach. We locate PTH among balanced-budget rules, expiring appropriations, earmarking, FOCJ, charitable tax concessions, and participatory budgeting, and derive positive implications for the percentage statutes. The conference version adds three empirical components, specified in Section 13: a panel of the European percentage-designation statutes, state-level estimates of the borrowing clause, and illustrative debt-path scenarios to 2056 under CBO's extended baseline, indexed by the displacement parameter.

JEL codes: D02, D72, D82, E62, H11, H41, H62, H63, H71

Keywords: budget deficits; fiscal rules; fiscal constitution; tax designation; percentage philanthropy; pay-as-you-go; expiring appropriations; constitutional political economy; Leviathan; robust mechanism design; tax and expenditure limitations

1. Introduction

The United States is projected to run structural primary deficits indefinitely. In the Congressional Budget Office's February 2026 baseline, the deficit is 5.8 percent of GDP in fiscal year 2026 and 6.7 percent in 2036, and federal debt held by the public rises from 101 percent of GDP to 120 percent over that decade and to 175 percent two decades later (Congressional Budget Office 2026a, 2026b). One line of research on that path studies the effect of fiscal institutions—automatic stabilizers, balanced-budget rules, tax and expenditure limits, supermajority requirements for taxes or borrowing, deficit targets—on the trajectory of debt and deficits, reading the experience of the US states and of other nations alongside the federal record; this paper belongs to that line. The institutions it has studied so far share a design grammar: they are rules about totals, written on the government, and enforced through courts, procedures, and electoral politics. The evidence that such institutions matter is real (Poterba 1994, 1996), and so are their recorded weaknesses. Rules on totals invite substitution toward the instruments the rule does not name, deficit finance being the classic exit from a tax constraint (Buchanan and Wagner 1977); their enforcement is assigned to actors with fiscal stakes of their own; and their optimal stringency depends on primitives—the severity of the present bias, the persistence of shocks—that the rule-writer rarely knows (Alesina and Tabellini 1990; Halac and Yared 2014; Yared 2019). This paper takes up a branch of that line that has not been studied: a fiscal institution with a different grammar, a rule written on the citizen rather than on the government—the citizen as fiscal principal—whose enforcement arises as a by-product of individual behaviour rather than as a task assigned to anyone, and whose evaluation is conducted—deliberately—without the knowledge its fine-tuning would require. Its working ancestors are the percentage-designation statutes of Central and Eastern Europe and of Italy, and the record it reads for the borrowing margin is that of the American states. The question we put to it is the one this literature puts to the others: what does it do to the deficit, and how much of the budget does the answer cover?

Two things are studied: the provision itself and, through it, a methodological claim about how fiscal constitutions should be evaluated. The provision is Personal Tax Holding (PTH): every citizen holds a constitutional right to direct a fixed fraction α of her own tax liability to certified public-goods organizations (PGOs) before a constitutional deadline; the right renews with each fiscal year, and accounts not directed by the deadline expire—

they lapse to the treasury under a redistribution rule fixed in advance; entry into the certified sector is free subject to a certification standard; and the designation right is coupled to a tax ceiling that can be lifted only with the consent of a coalition of specified size. The instrument has working ancestors. The “percentage philanthropy” statutes of Central and Eastern Europe—Hungary’s one-percent law of 1996, Slovakia’s two-percent scheme, Italy’s cinque per mille, and analogues elsewhere in the region—already let taxpayers assign a small slice of their income tax to eligible organizations (Bullain 2004; Török and Moss 2004).[2] PTH constitutionalizes the instrument and treats its parameters—the fraction, the deadline, the fate of lapsed funds, the certification standard, the surrounding tax ceiling—as objects of deliberate constitutional design rather than legislative accident.

Three results follow, and they are the paper’s business here. First, because unexercised claims expire with the fiscal year, the designated share of the budget is pay-as-you-go by an accounting identity: outlays on that margin cannot exceed the year’s designated revenue, no debt can accumulate there, and the primary deficit is the imbalance of the residual budget alone (Proposition 9). The constitutional fraction α is therefore the width of the budget withdrawn from intertemporal finance, and it prices that width against the base over which tax smoothing must operate (Barro 1979). Section 10.1 scales the trade against the Congressional Budget Office’s February 2026 baseline. At the one or two percent of the working percentage statutes the pay-as-you-go margin would be $29 billion to $59 billion—under two tenths of one percent of GDP, against a deficit projected at $1,887 billion for fiscal year 2027. At one half of individual income tax liability it would be $1,474 billion, or 4.4 percent of GDP, some four fifths of that deficit, and the revenue base left to absorb cyclical shocks would fall from 17.7 to 13.2 percent of GDP. The identity does not close a deficit. It fixes how much of the budget a deficit can no longer reach, and the fractions at which that width becomes a fiscal object are large ones.

Second, aggregate lapsing, observed at each year’s close, is a statistic of the demand for public expenditure—the first such statistic an involuntary fiscal system produces—and it is natural to let the next year’s designated levy respond to it. Here the single-cycle analysis earns its keep twice. A naive revealed-demand rule, which cuts the levy in proportion to realized lapsing, extinguishes the designated budget in every admissible environment, because lapsing never falls below the attrition hazard even when every citizen values the mechanism: friction masquerades as absent demand. The robust correction subtracts the worst-case lapsing floor that Section 6 characterizes, contracting the levy only when realized lapsing exceeds what friction alone can produce; the corrected rule never contracts under universal demand, and detects genuine non-demand exactly when its mass exceeds a computable bound. Third, the enforcement coalition constrains revenue; whether it constrains debt is a drafting fact—whether the constitutional ceiling is written on revenue alone or on revenue plus net borrowing—and the American states have already recorded what the narrower clause invites.

---

[2] Rates, eligible recipient classes, and administrative details differ across these statutes and have been revised repeatedly since their adoption; we invoke them as existence proofs for the administrative feasibility of individual tax direction, not as calibrations of the model.

Constitutional limits on general-obligation borrowing are followed by unrestricted borrowing, by public authorities, and by local debt rather than by less debt (von Hagen 1991; Bunch 1991; Kiewiet and Szakaly 1996), and restrictive tax and expenditure limits are associated with more state debt and with substitution toward non-tax revenue (Shadbegian 1999; Deller et al. 2013). Section 10.3 prices the mechanism's own version of that escape, which is the escape Buchanan and Wagner (1977) diagnosed.

The mechanism was not designed for that margin, which is why the identity is a by-product of its accounts rather than a premise of its design. The problem it was built for is older. Wicksell's (1896) complaint against the public finance of his day was that it judged taxation apart from the expenditure it finances, as though the two sides of the budget could be evaluated separately. Buchanan carried the complaint to the constitutional level: the productive question is not which fiscal outcome is best but which fiscal rules citizens would accept before knowing their own places under them (Buchanan and Tullock 1962; Brennan and Buchanan 1980, 1985). Modern fiscal constitutions answer half of that question. They constrain how much the state may take—bases, rates, procedures, ceilings—and interjurisdictional competition disciplines how much government citizens are willing to buy: a resident who dislikes the size or broad character of her jurisdiction's budget can move to one she likes better (Tiebout 1956; Oates 1972). What remains essentially unconstrained is what any given budget buys. Within a jurisdiction, public expenditure is a single bundle assembled by majoritarian and bureaucratic politics, and a citizen who agrees with her neighbours about how large the public household should be may still disagree with every one of them about its furniture. We call this the within-bundle problem: the component of preference heterogeneity that survives Tiebout sorting because sorting operates on bundles, not on their composition.

A deficit is the intertemporal form of Wicksell's complaint: expenditure detached not from the willingness to finance it in general but from the willingness to finance it now, with the difference booked to fiscal years whose taxpayers were not asked. The within-bundle problem and the deficit problem are therefore not neighbours by accident. Both arise because the fisc aggregates before it listens—across purposes in the one case, across time in the other—and an institution that restores the individual tie between payment and purpose turns out, once its accounts are made to expire, to restore part of the tie between payment and period as well. That second margin is this paper's addition to an analysis of the first.

The methodological claim concerns the criterion by which such a design should be judged. A constitution is written once, behind a veil of ignorance about the populations it will govern, and it cannot be re-optimized as those populations are revealed; costly revision is not a defect of constitutions but the source of their commitment value. The framer of a fiscal mechanism therefore cannot know the datum on which its fine-tuning would depend—here, the distribution of citizens' costs of engaging with the mechanism: the attention, time, and effort it takes to evaluate organizations and decide. A mechanism certified "optimal provided the cost distribution is G" is, to a framer who cannot verify the proviso, not certified at all. We accordingly replace the designer's prior with an uncertainty set—the class of cost distributions consistent with what a framer can credibly claim to know—and judge each candidate rule by its worst-case performance over that set. This is the maxmin discipline of robust mechanism design, descending from the

Wilson (1987) doctrine through Bergemann and Morris (2005) to the worst-case contracting of Carroll (2015). Our point of departure is that the discipline's native home may be constitutional. Elsewhere, the write-once, no-recourse premise that makes full worst-case evaluation seem extreme is a modelling stipulation; for a constitution it is a description of the object. Read this way, the maxmin criterion is not imported into constitutional political economy from contract theory: it is the operational form of the veil of ignorance that constitutional political economy has always used (Buchanan and Tullock 1962; Rawls 1971), applied to a framer who must write fiscal rules for citizens she will never meet.

Four design lessons result, each a statement about what a fiscal constitution can guarantee rather than about what an omniscient planner would choose. First, the designation deadline. Two forces pull against each other: a short window cuts search short and raises non-participation, while a long window keeps citizens searching and exposes them to attrition—the everyday inattention that makes people miss windows altogether. The worst-case lapsing rate is therefore non-monotone in the deadline, and under a condition we characterize its minimum is interior: a fixed, finite deadline set from structural primitives beats both a tight window and an open-ended one. The citizen this deadline protects is, perhaps surprisingly, the most patient one—the citizen for whom evaluation is costless—because patience is what feeds deferral, and under the robust deadline every other citizen lapses less often than she does. The deadline optimal for any particular realized population, by contrast, cannot be implemented by a framer who does not know the population; and if the framer can credibly assert one moment of the cost distribution —its mean—the constitution must instead balance the most patient against the most cost-burdened citizen, in proportions the mean pins down. Second, lapsed funds. The rule returning lapsed revenue to participating citizens should have a constant marginal rate—it should be linear—because it presses against an enforcement ceiling that is linear in realized lapsing; a schedule fine-tuned to participation elasticities the framer cannot estimate is exactly what a constitution should not attempt, a conclusion in the spirit of the flat, uniform tax rules Brennan and Buchanan (1980) urge on Leviathan for kindred reasons. Third, enforcement. Citizens who never designate hold a stake in the constitutional tax ceiling, and they form a coalition the government cannot dissolve by targeted side-payments—not because they are organized, but because the government cannot observe who has designated and so cannot aim its offers. Fiscal privacy is thereby revealed as constitutional infrastructure, playing a role analogous to the secret ballot in markets for votes: what cannot be observed cannot be cheaply bought. The protection is not a knife-edge; it degrades continuously as the government's information improves, and we compute the precision threshold at which it fails—the amount of opacity the ceiling needs. Fourth, certification. On the supply side the framer should cap the margin organizations may retain and leave quality to competition, rather than fix a quality target calibrated to costs the framer cannot observe; the logic is the delegation-by-caps logic of Frankel (2014) and Amador and Bagwell (2013), arrived at from constitutional premises.

Because the mechanism competes with existing institutions rather than with an unstructured status quo, we also locate PTH among its relatives: the functional, overlapping and competing jurisdictions (FOCJ) of Frey and Eichenberger (1996, 1999); the earmarking of tax revenues (Buchanan 1963); charitable deductions and credits;

participatory budgeting; balanced-budget rules and expiring appropriations, the deficit institutions nearest in spirit and in evidence (Poterba 1996; Liebman and Mahoney 2017); and the percentage-philanthropy statutes themselves, for which the analysis yields concrete positive implications—on how the worst-case-optimal designation window is set and then left alone, on the fate of undesignated shares, on why eligibility rules take the form of margin caps rather than performance targets, and on why the confidentiality of individual designations is an enforcement asset rather than a data-protection nicety.

Three candors govern the exercise. The first concerns what the worst-case criterion delivers. On the coarsest information set—nothing about the cost distribution beyond its support—the worst case turns out to be a single, knowable type, so a maxmin framer, a Bayesian framer with any full-support prior, and a framer who simply designs for the most patient citizen select the same deadline; the criterion begins to do distinctive work only when the framer can assert a moment such as the mean, and we develop both cases explicitly rather than let one masquerade as the other. The second concerns novelty: the contribution is less a new technique than a new domain for an existing one. The robustness method was built for principals, sellers, and information designers; we argue that the constitutional fisc—where its premises are simply true—is where it belongs, and we show what it yields there. The third concerns register. The exercise is positive throughout: we characterize what alternative rules guarantee and compare institutions by those guarantees; nothing in the paper recommends the adoption of any rule, and the comparative statements—this rule is worst-case-undominated, that one is not—are theorems about a model, offered as such.

1.1 Related literature

On the political economy side, the within-bundle problem descends from the fiscal-constitution program of Brennan and Buchanan (1980, 1985): if the constitution's task is to constrain a revenue-seeking state while still letting it produce what citizens want, then the composition of expenditure is a constitutional margin, not merely an administrative one. Tiebout (1956) supplies the disciplining benchmark whose boundary defines our problem; the polycentric tradition of Ostrom, Tiebout and Warren (1961) supplies the vision of provision unbundled from production that PTH pushes down to the individual level; and Hirschman (1970) supplies the vocabulary in which the mechanism is best described—designation is a third channel beside exit and voice, exit within the fisc, moving money rather than people. Buchanan's (1963) reconstruction of earmarking as piecewise fiscal choice is the mechanism's nearest collective ancestor; Frey and Eichenberger's FOCJ its nearest jurisdictional one; the percentage-philanthropy statutes surveyed by Bullain (2004) and Török and Moss (2004) its nearest working one. Section 11 develops these comparisons in detail.

On the deficit side, the paper joins the literature on fiscal institutions and fiscal rules. The American state evidence establishes that budget institutions are not veils: stricter balanced-budget requirements and tighter limits on deficit carryover are associated with faster fiscal adjustment (Poterba 1994, 1996), and Bohn (1998) supplies the sustainability benchmark—a corrective response of the primary surplus to debt—against which any institution's contribution can be read. The political-economy accounts of why deficits arise—strategic debt under alternating governments (Persson and Svensson 1989; Alesina

and Tabellini 1990), the constitutional diagnosis of Buchanan and Wagner (1977), and the survey of Yared (2019)—motivate rules but also explain why rules on totals leak: an instrument the rule does not name absorbs the pressure the rule creates. Halac and Yared (2014) put fiscal rules on the delegation footing this paper's methods share—rules as caps, chosen under private information about the temptation to spend—and Liebman and Mahoney (2017) document the behavioural signature of expiration inside government: agencies facing use-it-or-lose-it budgets spend heavily, and measurably badly, at the year's end. PTH relocates the expiring claim from the agency to the citizen; Section 10 asks what the relocation does to the deficit margin, and Section 11.5 develops the institutional comparison.

On the methods side, the worst-case discipline follows Carroll (2015, 2019). Chassang (2013) reaches detail-freeness by a different route—calibration through repeated interaction—that is instructive precisely because it is unavailable to a one-shot framer; that learning over time and worst-casing at a single stroke arrive at the same destination of simple rules is part of why the constitutional setting is a natural place to insist on the static version. Frankel (2014) shows that maxmin delegation collapses rich menus onto interval and cap rules, which is exactly the shape of our redistribution and certification results; Amador and Bagwell (2013) supply the general theory of such caps. Garrett and Pavan (2015) stand as the Bayesian dynamic foil: their optimal schemes depend finely on stochastic primitives the analyst is assumed to know, and our deadline analysis asks which coarse object survives when that knowledge is withdrawn. Dworczak and Pavan (2022) trace the entire spectrum between Bayesian and maxmin evaluation; in their terms, the reversibility of a decision indexes where it belongs on the spectrum, and a constitution sits at the fully robust pole. Yang (2022) is representative of the broader move of optimizing over a class of distributions consistent with coarse constraints, of which our Lemma 1 is the constitutional-fiscal instance. Bergemann and Morris (2005), Hurwicz (1972), and Wilson (1987) are the foundations; Carroll and Meng (2016) study how robust contracts degrade under noise, the analogue of our enforcement threshold.

The citizen's problem is one of sequential search with attrition in the tradition of Lippman and McCall (1976), and the supply side adapts Salop (1979), with Wolinsky (1986) and Anderson and Renault (1999) marking the search-theoretic reconciliation the supply side still owes and the conference version takes up. The paper proceeds as follows. Section 2 states the within-bundle problem. Section 3 lays out PTH as a constitutional provision. Section 4 develops the evaluation criterion. Section 5 presents the model. Sections 6 through 9 derive the four design lessons—deadline, lapsed funds, enforcement, and supply. Section 10 develops the deficit margin: the pay-as-you-go identity, the revealed-demand rule for the levy, and the borrowing clause. Section 11 compares PTH with neighbouring institutions—now including balanced-budget rules and expiring appropriations—and draws positive implications for the percentage statutes. Section 12 weighs what robustness buys and costs, Section 13 specifies the components of the conference version, and Section 14 concludes. All proofs are in Appendix A.

2. The within-bundle problem

2.1 The boundary of Tiebout sorting

A unit mass of citizens $i \in [0,1]$ each hold an ideal composition of public purposes—a point θi on a circle of unit circumference (Salop 1979)—and these ideal points are spread over the circle with positive density. A government collects revenue τY per citizen and provides one public expenditure bundle. In the Tiebout framework citizens sort across jurisdictions offering distinct bundles. Such sorting reveals preferences over the size and broad type of the bundle, but not over its composition, the dimension the circle records. Citizens may thus agree on how large the public sector should be yet disagree on its composition, and bundle-level sorting cannot resolve that disagreement. With K separable public-good dimensions and finitely many jurisdictions, a within-bundle welfare gap $\Delta WB > 0$ persists. The heuristic is that relocation cost is essentially flat in K while the dimensions along which a citizen is mismatched are not, so the gap should grow with K; we do not formalize this, and the one-dimensional circle used below does not establish it. Nothing that follows depends on the claim.

2.2 Exit, voice, and designation

Hirschman's (1970) two channels are poorly matched to the within-bundle margin. Exit disciplines through relocation, and relocation is whole-bundle by construction: a citizen cannot move out of the police budget and into the parks budget. Voice disciplines through collective decision, and collective decision aggregates to a single bundle by construction: however the votes are counted, one composition emerges, and the losers along every dimension remain mismatched. Designation is a third channel. It is exit without relocation—money moves, people do not—and it aggregates not through a collective choice rule but through the distribution of individual choices, so that minority compositions can be funded in proportion to the citizens who hold them. It is also the only channel on which Wicksell's consent principle operates literally rather than through representation: each directed dollar is public expenditure to which its payer has individually consented, and each lapsed dollar is expenditure to which consent was offered and not given—an accounting of consent that Section 10 converts into an accounting of demand. In the terms of the polycentric tradition, designation unbundles provision from production inside a single fisc (Ostrom, Tiebout and Warren 1961): the treasury remains one, the tax remains compulsory and uniform, but a constitutional slice of its use becomes polycentric.

That such a channel is administratively feasible is not conjecture. The percentage statutes have operated for decades: tax authorities receive designations with annual returns, verify recipient eligibility against a certified register, and transfer the designated slices (Bullain 2004; Török and Moss 2004). What the statutes have not been is designed. Their deadlines, lapse rules, eligibility standards, and fractions were fixed by legislative bargaining and revised by the same; none was chosen with a criterion in view, and none is constitutionally protected. The remainder of the paper asks what happens when they are treated as objects of constitutional design—and, because the design is constitutional, what criterion the designing should answer to.

3. Personal Tax Holding as a constitutional provision

The provision has six elements.[3] First, a designation right: each citizen may direct a fraction $\alpha\tau Y$ of her tax liability, with $\alpha \in (0,1)$ fixed constitutionally, to a certified PGO whose composition lies near her ideal point θ_i; the residual $(1-\alpha)\tau Y$ funds general purposes under ordinary oversight. Second, a certified sector: PGOs are formed by free entry of citizen-founders subject to a certification standard carrying a fixed cost F, and each may retain at most a margin m of the directed funds it receives—the certification instrument studied in Section 9. Third, a deadline and a default: designation must occur within a window of T periods; a citizen who has not directed by T, or whose attention lapses before then, sees her account revert to the general fund and obtains the default payoff L—the value to her of the undesignated composition, which the constitution can normalize and which lies below the best attainable match, so lapsing is a real loss. Fourth, a lapse rule: realized lapsed revenue R is returned to citizens who did designate according to a rule ψ fixed in advance, with the residual $R - \psi(R)$ reverting to the general fund—the object of Section 7. Fifth, an enforcement clause: the constitutional tax ceiling may be exceeded only with the consent of a coalition of mass at least μ, the object of Section 8. Sixth, an annual cycle: the right renews with each fiscal year, and claims not exercised within the year expire—there is no carry-forward of designation rights and no designation against future liabilities, so the designated margin is levied, directed, and settled period by period.[4]

Three contrasts fix what the provision is not. It is not a charitable deduction or credit: nothing is given out of net income, no price of giving is changed, and the directed funds never leave the fisc—the citizen allocates public money, within a public-purpose sector, under public certification. It is not a voucher: the fraction α is uniform and fixed, the eligible uses are public purposes rather than private consumption, and the general fund retains the complement. And it is not participatory budgeting: no assembly meets, no collective decision is taken, and the aggregation is statistical rather than deliberative.

---

[3] PTH is most naturally read as a constitutional provision: the fraction α, the deadline T, the redistribution rule ψ, and the coalition threshold μ are fixed prior to, and binding on, ordinary legislation. Nothing in the analysis requires a formal constitution rather than a durable statute; the operative premises are that the parameters are set before the relevant population is observed and that revision is costly enough to make the guarantee, rather than the expectation, the appropriate criterion. The percentage statutes themselves illustrate the point: their designation rights have proved difficult to revise once enacted, precisely because commitment is part of their design.

[4] The annual expiration rule, the menu of lapse destinations examined in Section 7, and the revealed-demand adjustment of Section 10 descend from the companion essay, which motivated individually held taxes through the consent principle of the Wicksellian tradition and through the observation that existing fiscal practice already treats a year's taxes as claims that expire with the fiscal year. The expiration of an unexercised right has a familiar constitutional analogue, also noted there: a ballot not cast by the close of polls is void, and it is the deadline that makes the electoral term—and with it the office's accountability—determinate. Sections 6 and 10 are the formal treatment of both observations.

Section 11 returns to each comparison; here they serve only to mark the mechanism's place—an individual, rule-based, intra-fisc instrument for the composition margin.

Reading the parameters as design objects yields the paper's agenda. How long should the designation window be, when the framer cannot know how costly citizens will find it to use (Section 6)? What should happen to the funds that lapse (Section 7)? Can the tax ceiling that accompanies the right be made to enforce itself (Section 8)? And by what standard should organizations be certified, when the framer cannot know how they are produced (Section 9)? And what does an expiring, individually exercised claim on the fisc guarantee about deficits—about the margin on which debt can and cannot accumulate, and about the response of the levy to what citizens' behaviour reveals (Section 10)?

4. Constitutional choice behind the veil: from priors to guarantees

Buchanan and Tullock (1962) located the normative force of constitutional agreement in uncertainty: a citizen who does not know her future place under a rule evaluates the rule by how it treats whoever she might turn out to be. Rawls (1971) pressed the same veil toward maxmin; Harsanyi (1955) toward expected utility under an equiprobability prior. For ordinary, revisable policy the Harsanyi position is compelling—priors can be formed, tested, and corrected as evidence accumulates. A constitution forfeits the correction. It is written before the relevant data exist; it binds successor populations its framer will never meet; and it is costly to revise by design, because costly revision is what makes commitment credible. For such a rule the natural evaluation is the guarantee: the worst performance over every environment the framer cannot rule out. This is the constitutional reading of the Wilson (1987) doctrine that a mechanism is more compelling the less it leans on common knowledge of fine details. For a constitution the doctrine is closer to a feasibility constraint than a desideratum: a certificate reading "optimal provided the cost distribution is G and log-concave" is, to a framer who can verify neither clause, no certificate at all.

The discipline this imposes is on what the framer may be assumed to know. We let her know the structure of the problem—the search technology by which citizens learn PGO quality, the fact that evaluation costs are nonnegative and bounded above by a known constant, and qualitative properties of the value of quality—but not the distribution of evaluation costs. She then chooses the mechanism to maximize its worst-case performance over every cost distribution consistent with that coarse knowledge. The criterion is deliberately conservative, and Section 12 weighs its price; its justification is that it asks of the constitution only what a constitution can deliver—a guarantee that does not require the framer to have guessed the governed correctly. One virtue deserves early notice: under this discipline, whatever the framer does assume is visible in the uncertainty set rather than hidden in a prior. When, in Section 6.2, we let her assert one moment of the cost distribution—its mean, identifiable from aggregate participation—the assertion is explicit, and the guarantee sharpens by an amount the reader can inspect. In Dworczak and Pavan's (2022) terms, ordinary policy may sit anywhere on the spectrum between Bayesian and worst-case evaluation; the constitution's irreversibility places it at the robust pole, not as a temperament but as a matching of criterion to object.

5. A model of designation

Before notation, the core trade-off. A citizen has a window of T periods to find a suitable PGO. Each evaluation costs her c in attention and effort. If she finds one good enough, she directs her account and is done. If she has not directed by the deadline, or if her attention lapses through attrition before then, her account reverts to the general fund. More time helps her find a better match—the search channel; more time also means more exposure to the attrition hazard that makes her miss the window altogether—the deferral channel. The model formalizes this trade-off and asks which constitutional parameters minimize the worst-case probability that accounts lapse unproductively.

5.1 Citizens and match values

There is a unit mass of citizens $i \in [0,1]$, each with income Y and tax liability τY, of which the PTH right covers $\alpha\tau Y$. A citizen who directs her account to a PGO of quality q at location ℓ on the circle obtains the designation value $v(q) - \Delta(\theta i, \ell)$, where v is strictly increasing and strictly concave and $\Delta(\theta i, \ell)$ is the within-bundle mismatch cost, increasing in the arc-distance between the citizen's ideal composition θi and the PGO's composition ℓ. The equilibrium configuration of PGOs—their locations and qualities—gives each citizen a distribution of attainable designation values. By the rotational symmetry of the free-entry configuration this distribution, written Φ with support [umin, umax] and density φ, is common across citizens, while θi still determines which PGO each citizen ultimately approaches. Sections 6–8 take Φ as given; Section 9 derives it from free entry and closes the model.

5.2 Search, attrition, and lapsing

Quality is an experience attribute: a citizen learns a PGO's designation value only by evaluating it. Time is discrete, $t = 1, 2, \ldots, T$, where T is both the constitutional deadline and the terminal period. Citizen i has a per-period evaluation cost ci, her type, drawn from a distribution G with support in $[0, \bar{c}]$; the bound $\bar{c}$ is known to the framer, but G is not. In any period before she has directed, she may evaluate one previously unexamined PGO at cost ci, observing an independent draw U distributed as Φ, with full recall. A citizen who has not yet directed leaves the process each period with hazard $\delta \in (0,1)$—attrition, the lapse of attention that makes a citizen miss the window—upon which her account lapses. A citizen who has still not directed by T also lapses. Lapsing yields the constitutionally specified payoff $L \in (\text{umin}, \text{umax})$, so that directing to a sufficiently poor match is worse than lapsing and a citizen directs only to a value above L. Let x denote her best option in hand, $x = \max\{L, \text{highest value evaluated so far}\}$, and let $s = T - t + 1$ be the number of periods remaining. The value of a type-c citizen holding x with s periods left satisfies

$$V_0(x; c) = x, \quad V_s(x; c) = \max\{\, x,\ -c + (1-\delta)\, E[V_{s-1}(\max\{x, U\}; c)] + \delta L \,\}, \quad s \geq 1: \tag{1}$$

she may direct now and obtain x, or pay c, survive attrition with probability $1-\delta$, draw U, and continue—lapsing at L if attrition strikes. The continuation value is nondecreasing in x with slope below one, so for each (s, c) there is a unique reservation $r_s(c)$: the citizen directs in the current period if and only if $x \geq r_s(c)$. A type-c citizen lapses if attrition

removes her before she directs, or if she reaches the deadline having never drawn a value above L. Write p(c; T) for the probability that a type-c account lapses, and

$$\Lambda(T; G) = \int[0, \bar{c}]\, p(c; T)\, dG(c) \quad (2)$$

for the aggregate lapsing rate under deadline T and cost distribution G.

5.3 The mechanism and welfare

A mechanism is a pair M = (T, ψ): a deadline T and a rule ψ for redistributing lapsed funds.[5] Realized lapsed revenue, R = ατY · Λ(T; G), is an observable aggregate. The rule returns an aggregate ψ(R) to citizens who directed their accounts, with the residual R − ψ(R) reverting to the general fund; directed and redistributed funds are valued at γd per unit, the general fund at γg < γd.[6] The mechanism also specifies a constitutional tax ceiling and the conditions under which it may be exceeded—the enforcement structure of Section 8.

5.4 The framer's ignorance and the criterion

The framer's ignorance is captured by an uncertainty set $\mathscr{G}$ of cost distributions consistent with coarse knowledge. The baseline class is shape-free on the known support,

$$\mathscr{G} = \{\, G : G \text{ is a probability distribution on } [0, \bar{c}] \,\}, \quad (3)$$

which encodes the single credible claim—that evaluation costs are nonnegative and bounded by c̄—and says nothing about their shape. The framer evaluates a mechanism by its worst-case welfare guarantee V(M) = min$_{G \in \mathscr{G}}$ W(M; G), where W(M; G) is per-capita welfare under M when the cost distribution is G, and chooses M to maximize V(M). Two criteria appear below and we distinguish them rather than let one stand in for the other. Section 7 uses the welfare functional W(M; G) = W0(G) + γgR + (γd − γg)ψ(R), where R = ατY·Λ(T; G) is realized lapsed revenue and W0 collects the terms ψ does not affect. Section 6 instead ranks deadlines by the worst-case lapsing rate. We do not claim that the two orderings coincide: a longer deadline improves the quality of realized matches as well as the exposure to attrition, and the lapsing rate does not price the former. Proposition 1 is therefore a statement about lapsing and not about welfare,

---

[5] The fraction α is fixed constitutionally and is not optimized here. Under our welfare specification the choice would be degenerate: with constant weights γd > γg and citizen behaviour independent of α, welfare is monotone in α and the robust choice would be the corner α = 1. An interior α requires curvature in the value of the general fund, which we do not model; Section 10.1 records the tax-smoothing motive of Barro (1979) as exactly such a curvature, operating across fiscal years, and treats α positively as the parameter of a discipline–smoothing trade. We therefore keep α an institutional constant and record its endogenous choice as open.

[6] The premium γd − γg is a primitive. Its informal rationale is the taxonomy of Friedman and Friedman (1980): funds a citizen directs toward a purpose she has chosen are, at the margin, one's own money spent on one's own ends, while the general fund is other people's money spent on others; the model prices the difference and does not derive it. Nothing below depends on the premium's size, only on its sign.

and we say so; giving conditions under which the two rank deadlines identically is left open. We impose the following throughout.

Assumption 1 (Interior default). *$L \in (u_{min}, u_{max})$; a citizen directs her account only to an evaluated value exceeding L.*

Assumption 2 (Universal participation). *$\bar{c} \le (1-\delta)\, E[(U - L)+]$; even the highest-cost type finds a single evaluation worthwhile, so every type participates for every $T \ge 1$.*

Assumption 3 (Regular match values). *Φ admits a continuous, strictly positive density φ on $[u_{min}, u_{max}]$.*

Assumption 4 (Nontrivial attrition). *$\delta \in (0,1)$.*

Assumptions 1 and 3 make lapsing an interior decision and keep the value functions and reservations continuous; Assumption 2 rules out the degenerate case in which the worst type never participates, so the deadline still has work to do; Assumption 4 is the friction that gives the deadline its role. Two remarks on Assumption 2. The numerical illustration of Section 6.1 deliberately sits close to its boundary, so that both channels of the deadline problem are active. And if the assumption fails for a set of types of positive measure, those types lapse with probability one regardless of T and the worst-case lapsing rate acquires a floor equal to their mass; the results of Sections 6 and 7 survive verbatim on the participating subpopulation, with the higher floor making an interior deadline less likely and pushing the robust deadline toward T = 1 when non-participation is widespread. We maintain Assumption 2 to isolate the two-channel tension cleanly.

6. The robust deadline: whom the constitution must protect

Fix ψ and α and study the deadline. Two opposing forces shape the lapsing rate. When T is small, a citizen has too few periods to secure a value above L, and lapsing is high whatever the cost distribution—the search channel. When T is large, the option value of further search keeps citizens looking, exposing them to attrition over a longer window, and lapsing rises again—the deferral channel, which the deadline exists to curb. Define the worst-case lapsing rate $\Lambda w(T) = \max_{G \in \mathcal{G}} \Lambda(T; G)$. The first result reduces this worst case to a maximization over types; the second identifies the maximizing type.

Lemma 1 (Worst-case reduction). *Under Assumptions 1–4, for every T the worst-case lapsing rate equals the largest type-specific lapsing probability, $\Lambda w(T) = \max_{c \in [0, \bar{c}]} p(c; T)$, and the maximum is attained: the worst case is a single-type (degenerate) cost distribution.*

The reason is simple. The map from G to $\Lambda(T; G)$ is linear; the set of distributions on the compact support $[0, \bar{c}]$ is convex and compact; and a continuous linear functional on such a set attains its maximum at an extreme point—here a point mass, a population made of one type. The consequence is that the worst-case lapsing rate depends on the support $[0, \bar{c}]$ and the structural primitives (δ, L, Φ) alone, never on the shape of G.

Lemma 2 (The worst type is the most patient type). *Under Assumptions 1–4, $p(c; T)$ is nonincreasing in the evaluation cost c for every T, and strictly decreasing at any c at which some finite-horizon reservation strictly falls. The maximum in Lemma 1 is*

therefore attained at c = 0: the worst-case lapsing rate equals the zero-cost type's own lapsing probability, p(0; T), at every T.

The proof is a coupling argument. A higher evaluation cost lowers the continuation value of search and hence every finite-horizon reservation; facing the same sequence of draws and the same attrition shock, a higher-cost citizen therefore directs her account no later than a lower-cost one, so the event that she lapses is contained in the event that the lower-cost citizen lapses. Both channels run the same way: patience raises the search-failure component by making a citizen pickier, and raises the deferral component by keeping her exposed to attrition longer. The citizen the constitution must protect is the one for whom evaluation is costless.

Lemma 2 disciplines what the worst-case criterion can be said to deliver here, and we state the consequence plainly. The maximization in Lemma 1 ranges over the whole support, but its solution is the left endpoint at every T; the support bound $\bar{c}$ does not enter the worst-case lapsing rate at all and appears only through Assumption 2. Equivalently, the robust deadline solves a single known type's finite-horizon search problem: a maxmin framer, a Bayesian framer with any full-support prior, and a framer who simply designs for the most patient citizen all select the same deadline, and the ignorance of G does no work. Read constitutionally, this is a generality result in the spirit of Buchanan and Congleton (1998): bound to write one deadline for all types, the framer writes the rule for a nameable, extreme, and knowable citizen—the one most exposed to the mechanism's characteristic risk—and under that rule every other citizen lapses less often than she does. What remains substantive is not a hedge against distributional ignorance but three things: the two-channel non-monotonicity of a single type's lapsing curve, a detail-free certificate for when the optimal deadline is interior, and the non-implementability of the per-type optima. Section 6.2 shows where the maxmin criterion does bind, and Section 12 returns to what this does and does not vindicate about the criterion.

To characterize the robust deadline, let

$$\bar{M}(n) = 1 - (1-\delta)n + \Phi(L)n, \ \Lambda S = \delta + (1-\delta)\Phi(L), \quad (4)$$

$$\Lambda D = \delta / [\ \delta + (1-\delta)(1 - \Phi(r\infty(0)))\ ], \quad (5)$$

where $r\infty(0)$ is the stationary reservation of the most patient (zero-cost) type. Here $\bar{M}(n)$ is an upper bound on the worst-case lapsing rate at deadline n; ΛS is the worst-case lapsing rate at T = 1; and ΛD is its limit as $T \to \infty$. The Deadline Paradox—an interior optimal deadline—holds exactly when some intermediate number of evaluation rounds beats both extremes in the worst case.

Proposition 1 (Robust Deadline Paradox). *Under Assumptions 1–4 the worst-case lapsing rate satisfies $\Lambda w(1) = \Lambda S$ and $\lim T\to\infty \Lambda w(T) = \Lambda D$, and both Λw and its minimizer depend only on (δ, L, Φ), the support bound $\bar{c}$ entering only through Assumption 2. The worst-case-optimal deadline TR is interior, $1 < TR < \infty$, if and only if the Interior Deadline Condition (IDC) holds:*

$$\min T \geq 1\ \Lambda w(T) < \min\{\ \Lambda S, \Lambda D\ \}. \quad (6)$$

When IDC holds the minimizer is finite and at least 2; when it fails a corner is optimal—the shortest deadline $T = 1$ if $\Lambda S \le \Lambda D$, and arbitrarily long deadlines if $\Lambda D < \Lambda S$; both corners are live. At $c = 0$ the stationary reservation lies strictly below the top of the match-value support, so the long-horizon limit is strictly less than one, and it may fall below or above the one-shot rate depending on the attrition hazard and the match-value distribution. By contrast the per-type optimal deadline $T^*(c) = \arg\min_T p(c; T)$ is increasing in the cost type over the range where the interior optimum obtains, so no per-distribution optimum is implementable by a framer who does not know G.

A condition sufficient for IDC, verifiable from primitives without solving any search problem, is $\min_{n \ge 1} \bar{M}(n) < \min\{\Lambda S, \Lambda D\}$, since $\Lambda w(n) \le \bar{M}(n)$ for every n. In the uniform benchmark that Section 9 delivers, where $\Phi(L) = \lambda$, the bound can be evaluated in closed form at $n = 2$. Writing $a = 1 - \delta$, $\bar{M}(2)$ falls short of the one-shot rate precisely when $\lambda^2 - a\lambda + \delta a < 0$, a quadratic in λ with real roots if and only if $\delta < 1/5$; in that case the inequality holds for every λ strictly between $(a - \sqrt{(a(a - 4\delta))})/2$ and $(a + \sqrt{(a(a - 4\delta))})/2$—a condition on the attrition hazard and the match-value distribution alone, checkable before any search problem is solved. It is only half of the certificate, however: the bound must also fall below the long-horizon floor ΛD, and it may not. At the configuration of Section 6.1, $(\delta, \lambda) = (0.15, 0.35)$, the interval is $(0.1945, 0.6555)$ and $\bar{M}(2) = 0.40 < \Lambda S = 0.4475$; but $\Lambda D = 0.3457 < 0.40$, so the certificate fails there. It is not vacuous—for the same λ it holds at every $\delta \le 0.05$—but at that calibration interiority must be verified by solving the search problem, and we report this rather than let the bound appear to do work it does not do.

IDC can be read directly. The bound $\bar{M}(n)$ adds a cumulative-attrition term, rising in the number of rounds, to a search-failure term, falling in it; its minimizer balances the two channels, and the condition asks that this balanced worst case beat both the one-shot search-failure rate ΛS and the long-horizon attrition floor ΛD. It holds when the attrition hazard δ is moderate and the probability $\Phi(L)$ of an unacceptable draw is bounded away from one—exactly the regime in which neither channel dominates. The point is not that an interior deadline is always optimal, but that when it is, its location is fixed by primitives a framer can know, while the deadline optimal for any particular realized population is not.

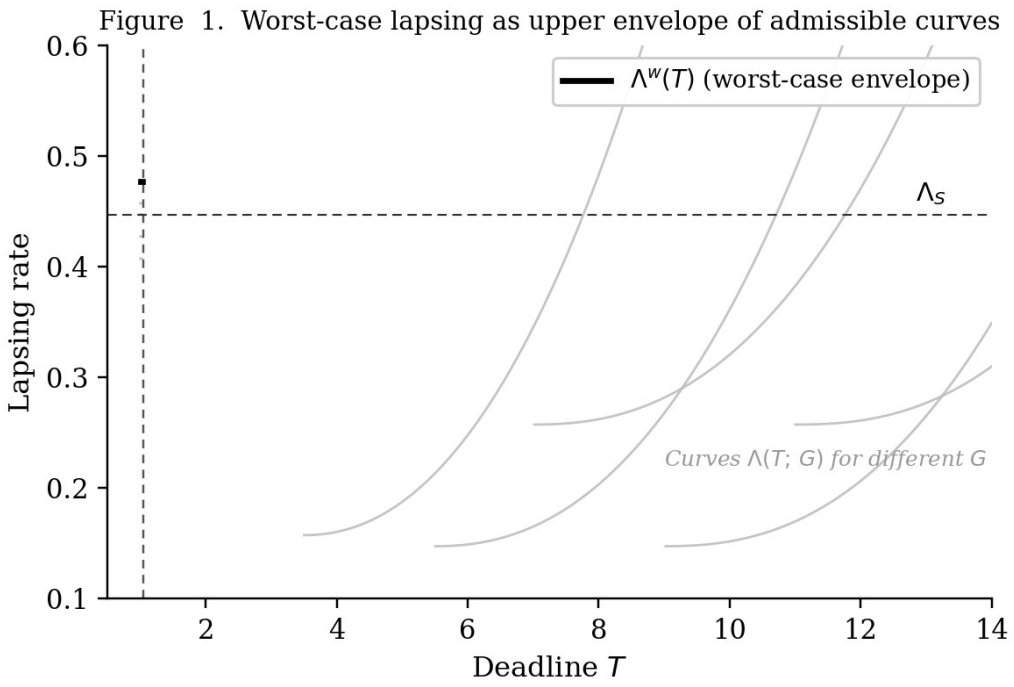


Figure 1. Worst-case lapsing Λw(T) = maxG ∈ 𝒢 Λ(T; G) (bold) as the upper envelope of the admissible distributions' lapsing curves (thin). Each cost distribution induces a U-shaped curve with its own optimum; these optima vary with the unknown distribution and are not implementable. Under IDC the worst-case curve attains its minimum at an interior, detail-free deadline TR determined by the structural primitives. The existence of an interior optimum is the robust object; the optimizer's location under any particular distribution is not.

The positive reading follows. A rule that shortens the deadline when lapsing is observed to be high fails on two counts: shortening can raise lapsing, since the worst-case rate is non-monotone, and chasing the per-distribution optimum is itself non-robust. The worst-case-optimal object is a fixed interior deadline set from structural primitives, not a deadline that reacts to realized participation. A remark on responsive rules belongs here. Within the single-generation model, no deadline rule can condition on realized lapsing, because lapsing is observed only after the deadline has passed; the class of responsive rules is empty and the claim that the fixed TR cannot be improved upon by responsiveness is, strictly, vacuous. A responsive rule requires an intergenerational reading with an explicit information flow; under that reading, calibration in the sense of Chassang (2013) may improve on a fixed deadline, and we flag the comparison as open rather than settled.

6.1 A numerical illustration

It is useful to watch the robust deadline emerge from explicit primitives. Fix the per-period attrition rate at δ = 0.15 and let the match-value distribution be the uniform law that Section 9 delivers under linear mismatch. By the location-invariance of Proposition 7 we normalize its support to [0, W] with W = t/2 = 1, so t = 2; the lapse payoff is L = λW

with $\lambda = \Phi(L) = 0.35$, the constitutional fraction of citizens whose best attainable value falls short of the standard. The only cost parameter is the upper end of the support, $\bar{c} = 0.17$. These numbers are not arbitrary: since $(1 - \delta)(1 - \lambda)^2 t/4 = 0.85 \cdot 0.65^2 \cdot 0.5 \approx 0.180$, the bound $\bar{c} = 0.17$ sits just inside the region Assumption 2 carves out—close enough to the boundary that both channels of the deadline problem are active.

For each deadline T we solve the citizen's finite-horizon search problem—backward induction for the reservations rs(c), then a forward pass that records the probability p(c; T) that a type-c account lapses—and read the worst case off Lemma 1 as $\Lambda w(T) = \max_c p(c; T)$. Two reference points frame the picture. At T = 1 the mechanism is a single draw and every type lapses at the one-shot rate $\Lambda S = \delta + (1 - \delta)\lambda = 0.4475$, independent of cost. As T grows without bound, the option value of further search keeps the most patient citizens looking, and a fixed fraction of them is claimed by attrition before they stop: $\Lambda D = 0.3457$, strictly below one, because even the most patient citizen stops once her reservation is met. The detail-free bound does not settle interiority at this calibration: $\bar{M}(1) = 0.50$, $\bar{M}(2) = 0.40$, $\bar{M}(3) \approx 0.429$, so $\min_n \bar{M}(n) = 0.40$ at n = 2, which does not fall below $\min\{\Lambda S, \Lambda D\} = 0.3457$. Interiority does hold, and must be verified by solving the search problem: the worst-case curve attains 0.3093 at T = 3, below both 0.4475 and 0.3457.

Solving for the envelope itself locates the optimum at TR = 3, where the worst-case lapsing rate falls to $\Lambda w(TR) \approx 0.309$—a reduction of nearly thirty-one percent from the one-shot rate, secured with no knowledge of the cost distribution beyond its support. Figure 2 plots $\Lambda w(T)$, the numerical counterpart of the envelope sketched in Figure 1. The curve drops steeply over the first two periods, as extra rounds let citizens clear the threshold—the search channel—reaches its floor at the third, and then climbs slowly as the deferral channel reasserts itself, drifting toward the limit $\Lambda D = 0.3457$ that it approaches only asymptotically. The minimizer of the detail-free bound, n = 2, need not equal the true optimum, and here it does not: $\bar{M}$ certifies cheaply but is not the objective, and the gap between n = 2 and TR = 3 is the price of that economy.

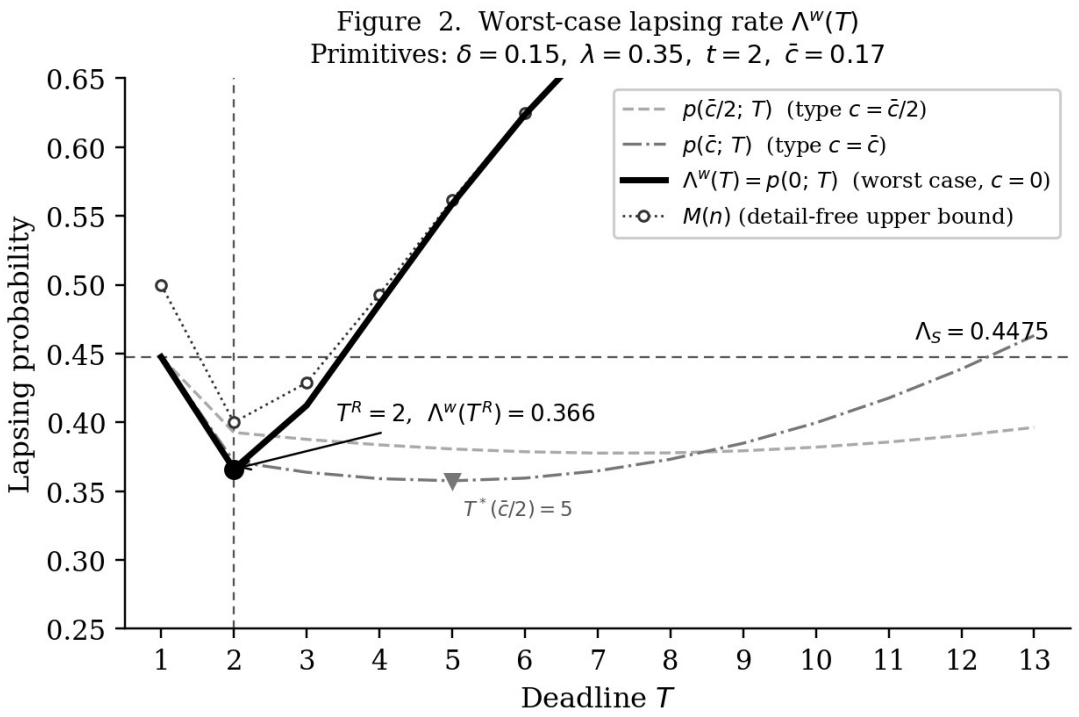


Figure 2. The worst-case lapsing rate Λw(T) for the primitives δ = 0.15, λ = Φ(L) = 0.35, t = 2, and c̄ = 0.17—a numerical instance of the envelope of Figure 1. The bold curve is Λw(T) = maxc p(c; T); it falls from the one-shot rate ΛS = 0.4475 to its minimum Λw(TR) ≈ 0.309 at the interior robust deadline TR = 3, then rises toward the long-horizon limit ΛD = 0.3457 as deferral reasserts itself. The maximizing type is the patient one, c = 0, at every T, so the envelope coincides with that type's lapsing curve; the thin curves are two higher-cost types, whose private optima lie to the right and cannot be implemented jointly. Faint markers trace the detail-free bound M̄(n), whose minimum 0.40 at n = 2 lies above ΛD, so the closed-form certificate fails at this calibration and interiority is verified numerically.

The same computation names the citizen the constitution must protect. At every deadline the maximum in Lemma 1 is attained at the most patient type, c = 0, so the worst-case envelope is that type's own lapsing curve—patience feeds deferral, and deferral is exactly what exposes a citizen to attrition over a long window. Here the non-implementability of Proposition 1 turns concrete. Could the framer set the deadline type by type, the patient type would receive T∗(0) = 3, the median-cost type T∗(c̄/2) = 5, and the most cost-burdened type a longer window still—each type tolerating a later deadline than the last, because a higher cost of waiting makes deferral self-limiting and the residual risk worth trading for a better match. Solving the search problem confirms the profile in this benchmark: T∗(c) = 3, 4, 5, 6 and 11 at c = 0, 0.04, 0.085, 0.11 and 0.14. (We do not prove monotonicity of T∗ in general; Lemma 2 establishes monotonicity of p in c, which is what the worst-case analysis needs, and a single-crossing argument for the minimizer is a separate exercise we leave open.) No single deadline implements that profile. The constitution, bound to one number, is pinned by the worst case at TR = 3—shorter than

any positive-cost type would choose alone, and optimal not for a type but for the population the framer cannot rule out.

6.2 Making the worst case bind: a known mean

Lemma 2 shows that on the shape-free class the maxmin criterion selects the zero-cost type, so ignorance of G does no work. The remedy is not to abandon the criterion but to let the uncertainty set encode a claim about G that a framer might credibly make beyond its support. The weakest such claim is a mean: aggregate participation identifies the average cost of engaging with the mechanism even when its dispersion is unknown, and Section 4 admits exactly this kind of coarse, verifiable knowledge. Let $\mathscr{G}(m)$ denote the set of distributions on $[0, \bar{c}]$ whose mean equals m.

Proposition 2 (Mean-constrained worst case). *Fix T and $m \in [0, \bar{c}]$. The worst-case lapsing rate over $\mathscr{G}(m)$ equals the value at m of the concave envelope of $p(\cdot; T)$ on $[0, \bar{c}]$, and the maximizing distribution is supported on at most two points. If $p(\cdot; T)$ is convex in c, that worst case equals $(1 - m/\bar{c})\cdot p(0; T) + (m/\bar{c})\cdot p(\bar{c}; T)$, attained by the two-point law placing mass $1 - m/\bar{c}$ at $c = 0$ and mass $m/\bar{c}$ at $c = \bar{c}$. Hence for $m \in (0, \bar{c})$ the worst-case lapsing rate depends on both endpoints of the cost support, the robust deadline is no longer the zero-cost type's optimum, and the maxmin criterion binds.*

The economics is transparent. A mean pins how much probability the adversary may move away from the zero-cost type, but not where it goes; because lapsing is decreasing in cost, the adversary sends that probability as far as it can, to $\bar{c}$. The framer must therefore guard simultaneously against the most patient and the most cost-burdened citizen, in the proportions the asserted mean permits, and the support bound $\bar{c}$ does real work. In the uniform benchmark $p(\cdot; T)$ is convex and decreasing in c at every T, so the two-point characterization applies. Solving the search problem, the robust deadline lengthens as the asserted mean rises: it equals 3, 3, 4 and 5 at m = 0, 0.04, 0.085 and 0.13, with worst-case lapsing rates 0.3093, 0.2923, 0.2705 and 0.2445. At m = 0 the shape-free case of Proposition 1 is recovered; at $m = \bar{c}$ only the point mass at $\bar{c}$ has mean $\bar{c}$ and the problem degenerates again at the other end. Between these extremes the guarantee is strictly better than the shape-free one and the deadline that delivers it is not any single type's optimum. This—and not the shape-free class—is where we would locate the paper's methodological claim: the constitution's guarantee improves exactly in step with what its framer can credibly assert about the governed, and the improvement is inspectable.

7. Lapsed funds: a constant marginal rate

Funds that lapse must return to the fisc under a rule the constitution fixes in advance, and the framer cannot know how future participation will respond, because that response runs through the same unknown cost distribution. Realized lapsed revenue $R = \alpha\tau Y \cdot \Lambda$ is observable, and the rule ψ maps it into an aggregate redistribution $\psi(R)$ credited to citizens who directed—raising the value of directing, and so bearing on participation through the unknown distribution. Redistribution is socially valuable at the margin, since redistributed funds carry weight γ_d and the residual the lower weight γ_g. But it is constrained: per-designator redistribution cannot exceed the size of the lapsing penalty,

$\psi(R)/(1 - \Lambda) \le |L|$, for otherwise the penalty that sustains the enforcement coalition of Section 8 is undone. Since $\Lambda = R/(\alpha\tau Y)$, this constraint is the affine ceiling

$$\psi(R) \le B(R) := |L| \cdot \left(1 - R/(\alpha\tau Y) \right). \quad (7)$$

Before solving, it is worth fixing the design space, because a constitution could in principle send lapsed revenue to four destinations. It may let lapses revert to the treasury ($\psi = 0$), the silent rule of every existing percentage statute. It may credit them to the citizens who did designate—the family the ceiling (7) constrains and this section studies. It may recycle them as designation rights: lapsed revenue augments a later cycle's designated pool, spread over all citizens, taxpayers and non-taxpayers alike—which is not a choice of ψ within the year but a choice of the levy path across years, and belongs to the dynamics of Section 10. Or it may rebate them as private income, an equal per-capita transfer out of the fisc. The last destination is not innocent: a rebate that reaches non-designators gives them a claim that rises with realized lapsing, and thereby erodes the homogeneous stake w on which the enforcement coalition of Section 8 stands—the equalitarian destination is purchased with enforcement. The analysis below accordingly restricts attention to rules that keep lapsed funds inside the fisc and credit only designators, with the treasury as residual claimant; within that family, the question is the shape of ψ.

Structurally this is the problem of Carroll (2015): the designer knows a bound—here the ceiling B, the counterpart of limited liability—but not how the population responds to the instrument, and seeks the rule with the best guarantee. The answer, here as there, is a rule of constant marginal rate; the reason, as we are explicit below, is shallower here than there.

Proposition 3 (Linearity of the robust redistribution rule). *Under Assumptions 1–4, among all rules ψ that are feasible ($0 \le \psi(R) \le R$) and enforcement-compatible ($\psi(R) \le B(R)$) for every $G \in \mathcal{G}$, the worst-case welfare guarantee $V(\psi) = \min_G W(\psi; G)$ is maximized by $\psi^*(R) = \min\{ R, B(R) \}$, which is concave and piecewise linear and, on the region where the enforcement ceiling binds, has the constant marginal rate $d\psi^*/dR = -|L|/(\alpha\tau Y)$. Every rule with a non-constant marginal rate on the realizable range of R is worst-case dominated, and ψ^* is the unique feasible, enforcement-compatible rule that is worst-case undominated. Indeed ψ^* maximizes $W(\psi; G)$ pointwise, for every $G \in \mathcal{G}$ separately, so it is optimal under every prior; the qualifications below spell out what this does and does not establish.*

The reason is clear once the structure is laid bare. Because the uncertainty set is rich, realized lapsing—and hence R—ranges over a nondegenerate interval as G varies, so the constraints bind pointwise across that interval. Because welfare rises in $\psi(R)$ at each realization, the optimum sets ψ as high as feasibility and the ceiling allow, namely $\psi^* = \min\{R, B(R)\}$, which inherits a constant marginal rate from the affine ceiling on the binding region. Any rule with a non-constant rate either breaches the ceiling at the realization that maximizes its slope or forgoes welfare at the one that minimizes it—and nature selects whichever realization is worst. The constitutional characterization is concrete and detail-free: the rule that hard-wires the return of lapsed funds at a constant marginal rate is the unique worst-case-undominated member of the feasible class, while

every schedule calibrated to participation elasticities the framer cannot estimate is dominated. The kinship with Brennan and Buchanan (1980) is not decorative. Their case for flat, uniform tax rules is that fine-tuned fiscal instruments presuppose knowledge the constitutional stage does not have and confer discretion the post-constitutional stage will abuse; the constant-rate lapse rule is the same argument run on the expenditure side of the mechanism.

Three qualifications belong in the text rather than a footnote. First, $\psi*$ is pointwise optimal—it maximizes $W(\psi; G)$ for each admissible G separately—so the maxmin criterion does no work here. This is unlike Carroll (2015), where linearity is not pointwise optimal and is earned precisely by the worst-case criterion; the present result is a dominance argument, and we do not claim the deeper correspondence. Second, the constant marginal rate is inherited from the ceiling B, which is imposed rather than derived; deriving B from the enforcement coalition of Section 8—endogenizing the stake w as the net loss a non-designator suffers from a higher tax, and reading the ceiling off the coalition's incentive constraint—would make the appearance of $|L|$ a conclusion rather than a notation, and is worth doing. Third, the motivation invokes redistribution feeding back into participation, but the decomposition $W(\psi; G) = W0(G) + \gamma gR + (\gamma d - \gamma g)\psi(R)$ treats R as exogenous given G: no participation feedback is present in the formal model. A two-generation or steady-state formulation, in which ψ raises the value of directing, shifts $p(c; T)$, and makes R endogenous to ψ, is the natural next step; under it linearity is either earned or refuted, and either outcome is informative. Finally, Assumption 1 permits $L > 0$, in which case lapsing is not a penalty and the reading of $|L|$ as its magnitude fails; Sections 7 and 8 maintain $L < 0$ as a normalization.

8. Leviathan and the passive coalition

It remains to ask whether the fiscal constraint enforces itself—the question Brennan and Buchanan (1980) made central and left to constitutional culture to answer. The constitution sets a tax ceiling that may be exceeded only with a coalition's consent: the ceiling is lifted if and only if the mass of consenting citizens is at least μ. Non-designators—citizens whose accounts lapsed—hold a stake $w > 0$ in the ceiling and consent to lifting it only for compensation of at least w; designators consent at no cost. Their mass is Λ, so the consenting mass absent payments is $1 - \Lambda$. We treat μ as a constitutional parameter fixed before the population is realized, set in the region $\mu \in (1 - \Lambda, 1]$ that makes non-designators pivotal—the interesting enforcement region—and study how the mechanism performs there; the endogenous choice of μ, which would require specifying the framer's preferences over tax levels, is left for future work. To breach the ceiling, a government that gains $\Delta\pi$ in revenue from a higher tax must buy $\zeta := \mu - (1 - \Lambda)$ additional consents from non-designators.

Two points of discipline are needed before the results, because $\Lambda = \Lambda(T; G)$ depends on the very distribution the framer does not know. First, the framer cannot verify $\mu > 1 - \Lambda$, or the calibration below, distribution by distribution; both must hold uniformly over $\mathscr{G}$. By Lemma 2 the lapsing rate ranges over $[\Lambda\min, \Lambda\max]$ with $\Lambda\min = p(\bar{c}; T)$ and $\Lambda\max = p(0; T)$, so the requirement is $\mu > 1 - \Lambda\min$. Second, the coalition's per-capita stake $(\mu - 1 + \Lambda)/\Lambda$ is increasing in Λ, so it is smallest at the smallest admissible lapsing rate: the binding calibration is $w(\mu - 1 + \Lambda\min)/\Lambda\min \geq \Delta\pi$, and Proposition 5 should be read as

reporting the worst-case threshold, the infimum of $\rho*(\Lambda)$ over admissible Λ. Propositions 4 and 5 are stated for a given Λ; the uniform requirement is recorded here.

Proposition 4 (Robust enforcement under unobservability). *Suppose designation status di is not among the variables on which the government can condition side-payments before the deadline, and the constitution is calibrated so that the blocking coalition's aggregate stake exceeds the prize, $w\cdot\zeta/\Lambda \geq \Delta\pi$. Then for every budget-feasible side-payment scheme the coalition remains blocking and the ceiling holds. The conclusion covers the entire class of schemes because unobservability prevents the government from raising the non-designator share of any compensated set above the population rate Λ; the dilution factor $1/\Lambda$ is intrinsic to the informational constraint.*

The substantive question is whether this survives when the government is not perfectly blind. Replace perfect unobservability with a binary signal $\sigma i \in \{ND, D\}$ of designation status, of precision $\rho = Pr(\sigma i = di) \in [1/2, 1]$: $\rho = 1/2$ is the uninformative case of Proposition 4, and $\rho = 1$ is perfect observation. Enforcement degrades continuously, and the breaking point can be located.

Proposition 5 (Continuous degradation and the precision threshold). *Let the government observe a signal of designation status with precision ρ. The blocking power of the coalition—the largest revenue gain $\Delta\pi$ it deters—equals the cost $K(\rho) = w\cdot\zeta/\pi(\rho)$ of dissolving it, where $\pi(\rho) = \rho\Lambda / [\rho\Lambda + (1-\rho)(1-\Lambda)]$ is the posterior non-designator share among flagged citizens; K is continuous and strictly decreasing in ρ. Provided $w\cdot\zeta < \Delta\pi < w\cdot\zeta/\Lambda$, there is a threshold $\rho* \in (1/2, 1)$, solving $\pi(\rho*) = w\cdot\zeta/\Delta\pi$, such that the coalition blocks for all $\rho < \rho*$ and fails for $\rho > \rho*$. The unobservability result is therefore robust in a neighbourhood of $\rho = 1/2$ rather than a knife-edge, and enforcement degrades gracefully rather than collapsing at the first leak of information.*

Figure 3. Blocking power as a function of government precision

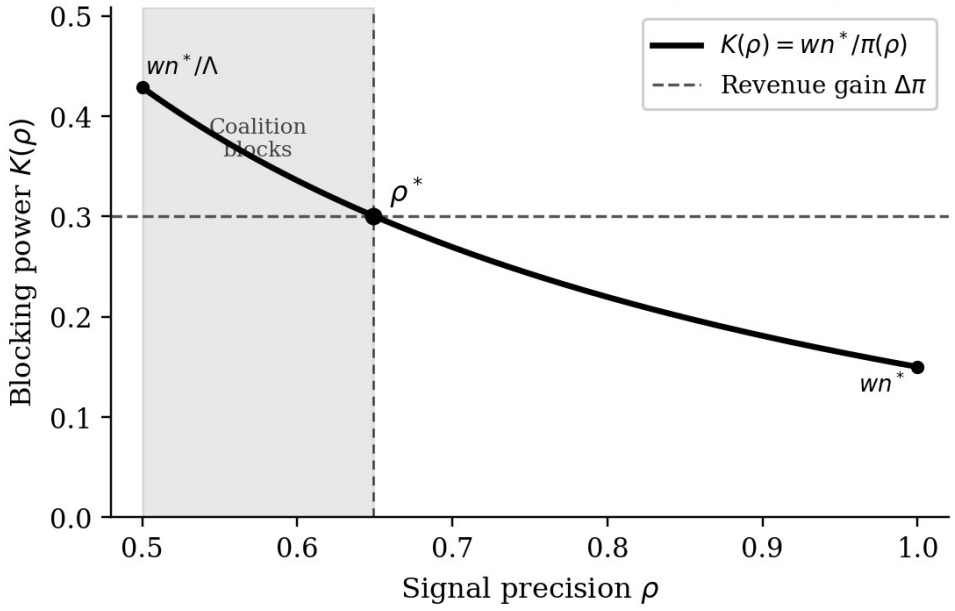


Figure 3. Blocking power K(ρ), the cost of dissolving the enforcement coalition, as a function of the government's signal precision ρ. At ρ = 1/2 (no information) the cost is w·ζ/Λ; at ρ = 1 (perfect observation) it is w·ζ. The coalition blocks while K(ρ) exceeds the revenue gain Δπ, that is, for ρ below the threshold ρ∗ (shaded region); above ρ∗ the government can profitably breach the ceiling. The degradation is continuous, paralleling the behaviour of robust contracts under additive noise (Carroll and Meng 2016).

The result assigns constitutional value to an unglamorous feature: the government's inability to see who has designated. Fiscal privacy here is not a data-protection nicety but enforcement infrastructure, playing a role analogous to the secret ballot in markets for votes—what cannot be observed cannot be cheaply bought. A treasury that cannot identify non-designators cannot buy them out of the blocking coalition at less than the diluted price w·ζ/Λ, and Proposition 5 prices the asset: it tells the framer not merely that opacity protects the ceiling but how much opacity the ceiling needs—the actionable quantity. The Leviathan reading is direct. Brennan and Buchanan's constitution constrains the revenue-maximizing state with rules whose enforcement is left offstage; here a by-product of the designation mechanism itself—a mass of passive citizens with a homogeneous stake, shielded by informational opacity—supplies the enforcement, without organization, without coordination, and without anyone intending it.

Two modelling gaps are recorded rather than papered over. The side-payment game is not formalized: the government's strategy space, the timing of offers, and the origin of the homogeneous stake w are left unspecified, so "every budget-feasible side-payment scheme" is an informal class. And designators are assumed to consent at no cost although they too pay (1 − α)τY and therefore have a stake in the ceiling; that asymmetry drives Propositions 4 and 5 entirely. It can be micro-founded—for instance if the redistribution ψ exactly offsets a designator's loss from a higher tax—but we have not derived it, and it should be read as an assumption.

9. The supply side: certifying organizations the framer cannot know

The analysis so far has treated the distribution of attainable designation values, Φ, as a primitive. It is not exogenous: it is the equilibrium output of the organizations citizens choose among, whose number, locations, and quality are set by entry and competition. This section closes the model by deriving Φ from a free-entry equilibrium of citizen-founded PGOs—in a complete-information benchmark whose status Section 9.1 states exactly—and then puts to the supply side the question the earlier sections put to the demand side: which features of the configuration, and which certification instrument, a framer can fix without knowing the primitives. Two messages emerge. First, the equilibrium Φ takes a simple form—uniform, under linear mismatch—and satisfies the maintained assumptions, so the demand-side guarantees carry over to the endogenous Φ; and because the uniform form rests on symmetric locations and undirected sampling rather than on the information structure, the guarantees survive the search-consistent extension of the supply side, as Section 9.1 records. Second, the robust certification instrument is a constant cap on the share of directed funds an organization may keep, with quality left to competition, rather than a quality target calibrated to costs the framer cannot observe—the same conclusion, and for the same reason, as the linearity of the lapse rule.

9.1 Entry, location, and quality competition

Assumption 5 (Uniform ideal points). *Citizens' ideal compositions θi are distributed uniformly on the circle of unit circumference.*

A PGO is a citizen-founded organization that meets a certification standard at fixed cost F, chooses a location on the composition circle, and selects quality at cost κ(q)—strictly convex with Inada conditions. The constitution permits each PGO to retain a margin m of the directed funds it receives; this margin cap is the supply-side instrument, analogous to the lapse rule ψ. With linear mismatch Δ(θ, ℓ) = t·d, a standard Salop quality-competition argument (Appendix A.4) delivers the symmetric free-entry equilibrium. Writing β = (1 − Λ)b for the effective margin per assigned citizen—each PGO retains b = mατY on the fraction 1 − Λ of assigned citizens who direct—the common quality q∗ and number of organizations n∗ jointly satisfy

$$\beta \cdot v'(q^*) / t = \kappa'(q^*) , \; n^* = \beta / [\, \kappa(q^*) + F \,] , \quad (8)$$

and the induced match-value distribution is uniform,

$$\Phi = \text{Uniform}[\, v(q^*) - t/2 , \, v(q^*) \,] , \; u_{min} = v(q^*) - t/2 , \; u_{max} = v(q^*). \quad (9)$$

The lapse payoff is fixed by the constitutional fraction rule

$$L = u_{min} + \lambda (u_{max} - u_{min}) , \; \lambda \in (0,1) \text{ fixed}, \quad (10)$$

under which Φ(L) = λ: the constitution sets, once and for all, the fraction of the match-value range that counts as unacceptable, rather than a nominal threshold whose bite would depend on equilibrium objects the framer cannot know.

Assumption 6 (Observable quality in the supply-side benchmark). *In computing the demand facing a deviating organization, its quality is treated as observable prior to evaluation, so that patronage is determined by the Salop indifference condition.*

This assumption is inconsistent with Section 5.2, where quality is an experience attribute learned only by paying the evaluation cost, and we state the inconsistency rather than suppress it. Under the search model, the demand facing a deviator would be governed by the acceptance behaviour of the citizens who happen to evaluate it, not by an indifference condition among citizens who observe the deviation in advance, and the first-order condition in (8) would take a different form. Two further tensions run the same way: with finitely many evenly spaced organizations the induced match-value law is discrete, whereas Assumption 3 requires a continuous density, and the sequential search of Section 5.2 samples without replacement and permits deadlines exceeding the number of organizations. Expression (9) should accordingly be read as the large-n limit, and this section as a complete-information benchmark for the supply side: Propositions 6–8 are conditional on Assumption 6, and Sections 6–8 do not depend on them, since there Φ is a primitive satisfying Assumption 3 and the benchmark serves only to show that a natural free-entry configuration can generate such a Φ. Reconciling the two sides is the paper's largest acknowledged debt; it has a known form—the framework of Wolinsky (1986) and Anderson and Renault (1999), with a continuum of organizations whose locations and

qualities are both unobserved before evaluation—and the conference version takes it up. What can be recorded now is why the demand side does not wait on it.

Assumption 6 enters the construction of Φ only through the level of the common quality. The derivation in Appendix A.4 obtains Φ from undirected sampling—a citizen who evaluates draws a uniformly random organization, precisely because quality is an experience attribute—so for any configuration with locations uniform on the circle and a common quality q, linear mismatch delivers a sampled value $v(q) - t \cdot d$ with d uniform on $[0, 1/2]$, and hence Φ uniform on $[v(q) - t/2, v(q)]$: the shape and the spread $t/2$ of (9), with only the location free. Passing to the continuum makes the uniform law exact rather than a large-n limit and settles the other two tensions at a stroke: the sampled law is continuous, so Assumption 3 holds without rounding, and sampling without replacement coincides with independent draws, so no deadline outruns the number of organizations. What the information structure does move is the deviation calculus—under search a deviator's demand runs through the acceptance behaviour of the citizens who sample it, the first-order condition replacing (8) prices quality through acceptance rather than through the Salop indifference condition, and the equilibrium quality, and with it the location of Φ, will in general differ from $q*$. But under the fraction rule (10) the lapse payoff moves with the support, so for the uniform law just derived $\Phi(L) = \lambda$ at any location; the sufficient condition for Assumption 2 in Proposition 6, $\bar{c} \le (1 - \delta)(1 - \lambda)^2 t/4$, involves no location; and by Proposition 7 the robust deadline, the lapse rule, and the enforcement threshold depend on Φ only through (t, λ). Every demand-side guarantee is therefore invariant to the information structure of the supply side, and the recursive order of Proposition 6—Λ first, then the configuration—survives with a search-based first-order condition in the last step. What the extension must still deliver is the supply side itself: existence and uniqueness of the symmetric search equilibrium, and the analogues of Propositions 6 and 8 under acceptance-based demand. Those are the conference version's objects; the guarantees of Sections 6–8 do not depend on them.

Proposition 6 (Endogenous configuration and induced match values). *Maintain Assumptions 4–6, with v strictly increasing and concave, linear mismatch $\Delta(\theta, \ell) = t \cdot d$, and κ strictly increasing and strictly convex with $\kappa(0) = 0$ and the Inada conditions. Then the symmetric free-entry equilibrium exists and is characterized by (8): an interior common quality $q*$ and a number of organizations $n*$. The induced match-value distribution is (9). Under the fraction rule (10), Assumptions 1 and 3 hold for this Φ with $\Phi(L) = \lambda$; if in addition $\bar{c} \le (1 - \delta)(1 - \lambda)^2 t/4$ then Assumption 2 holds, and, under Assumption 6, Lemmas 1–2 and Propositions 1–5 apply to the endogenous Φ. The location $u_{max} = v(q*)$ is strictly increasing in the effective margin and the spread $t/2$ in the mismatch intensity. Under (10) the lapsing rate is fixed by $(T, \bar{c}, \delta, t, \lambda)$ alone, independently of the supply side, so the configuration is recursive—Λ first, then $\beta = (1 - \Lambda)b$, then $(q*, n*)$; under a fixed nominal penalty it is instead a fixed point in (β, Λ), which exists and is unique under the margin bound of Lemma A.1.*

9.2 The robust certification instrument

The framer who writes the certification standard faces on the supply side the difficulty she faced on the demand side: the welfare-relevant quantities depend on primitives—the cost of quality κ, the mismatch intensity t, the value of quality v—that a constitution

cannot presuppose. Two instruments could in principle regulate quality. A quality target certifies only organizations whose quality reaches a constitutionally fixed floor qmin. A margin cap fixes the retained fraction m and certifies any quality, leaving the level to competition through (8). The comparison between them is the supply-side counterpart of the linearity result, and it rests on two observations.

The first is that endogenizing supply does not smuggle detail-dependence back into the demand-side characterizations. By Proposition 6 the equilibrium Φ enters the robust deadline, the lapse rule, and the enforcement threshold only through two coarse summaries: the spread t/2, which the framer can bound from the observable dispersion of compositions, and the fraction $\lambda = \Phi(L)$, which she sets directly through (10). The fine primitives (κ, F, v) move the location $u_{max} = v(q^*)$ of Φ, but the robust objects depend on Φ only through λ and the relative positions of the reservations within the support—not through its location—so they are invariant to (κ, F, v) given (t, λ).

Proposition 7 (Robust supply is detail-free). *Maintain the hypotheses of Proposition 6. The robust deadline of Proposition 1, the linear lapse rule of Proposition 3, and the enforcement threshold of Proposition 5 depend on the endogenous match-value distribution Φ only through the support spread t/2 and the constitutional fraction $\lambda = \Phi(L)$. They are invariant to the fine supply primitives (κ, F, v), which enter the equilibrium solely through the location $u_{max} = v(q^*)$ of Φ and wash out of every robust characterization. Endogenizing the supply of organizations therefore reintroduces no detail-dependence: a framer who can bound the dispersion of compositions and fix the penalty fraction can implement all of Sections 6–8 without knowing how organizations are produced.*

The second observation concerns the instrument. Measure the framer's supply-side objective by the worst-case misallocation of quality—the largest gap, over an admissible class of environments $E = \{(\kappa, t, v)\}$, between the induced quality and the efficient quality q_e for that environment. A quality target is a single number applied to every environment: when it binds it pins quality at qmin regardless of the environment, so its worst-case gap is exactly half the range of efficient qualities across E and cannot be reduced by any choice of floor. A margin cap, by contrast, induces through (8) a quality that moves with the environment—the same fixed cap yields higher quality where quality is cheaper or more valuable. When the responsive quality can be calibrated to stay within the efficient range (Condition T below), the margin cap's worst-case gap is strictly smaller.

Proposition 8 (A margin cap dominates a quality target; conditional on Condition T). *Let efficient quality vary non-trivially across the admissible class E, with range of width $\Delta q > 0$, and suppose the framer minimizes worst-case quality misallocation. Then every fixed quality target has worst-case misallocation at least $\Delta q/2$. If Condition T holds—there is a single margin cap $\bar{m}$ whose induced quality $q^*(\bar{m}; \cdot)$ stays within distance $A < \Delta q/2$ of efficient quality for every environment in E—then the best margin cap attains worst-case misallocation at most A and strictly dominates every quality target.*

The result is marked as conditional, and deliberately so. Condition T assumes precisely the tracking property that dominance requires; Lemma A.2 replaces it with a

responsiveness condition—the competitive quality at least half as responsive as the efficient quality on each side of a midpoint calibration—that is verifiable in the linear-Salop benchmark but is itself high-level rather than derived from primitives. The efficient quality must also be defined relative to an explicit welfare criterion—whether evaluation and entry costs are counted changes the target—and we take it to be the quality maximizing total surplus in a given environment at the free-entry number of organizations. Establishing Condition T from primitives, and relating the cap to the interval-and-cap optimality of Amador and Bagwell (2013), is left open. Even so, the direction of the result is the one the rest of the paper travels: robustness collapses a rich instrument space onto caps and constant slopes, on the supply side as on the demand side.

Endogenous entry also pins down a tax-rate floor. Certifying and sustaining n∗ organizations consumes real resources n∗F per capita in certification cost alone, which the constitution must finance; a tax rate that raises less than n∗F cannot support the equilibrium supply, so there is a positive lower bound τmin on the constitutional tax rate, set by n∗F. The certification cost the framer can least do without thus becomes a floor under the tax rate, closing the last element the demand-side analysis took as given.

Figure 4. Supply-side equilibrium and induced match-value distribution

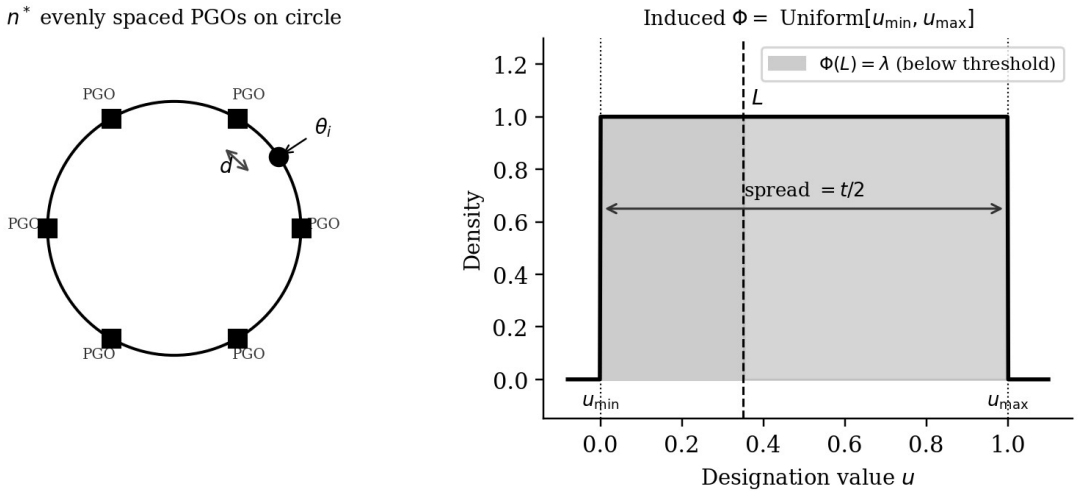


Figure 4. Left: the symmetric free-entry configuration of n∗ evenly spaced organizations on the circle of compositions; a citizen at ideal point θ patronizes the nearest. Right: the induced match-value distribution Φ, uniform on [v(q∗) − t/2, v(q∗)] by (9). Quality q∗ sets the upper edge umax = v(q∗) and the mismatch intensity t sets the width t/2. The lapse payoff fixed by the fraction rule (10) sits at L = umin + λ(umax − umin), so the shaded mass to its left—draws too poor to accept—is exactly Φ(L) = λ, the search-failure probability that enters the worst-case lapsing rate of Proposition 1.

10. The deficit margin: expiration, revealed demand, and the borrowing clause

The mechanism was introduced for the composition of public expenditure. This section develops its consequences for the level—for deficits, and for the path of the levy—which

follow from the one element of the provision that the earlier sections used only within a single cycle: the annual expiration of unexercised claims. Three results organize the section. The designated margin is pay-as-you-go by an accounting identity, so the fraction α prices a trade between fiscal discipline and tax smoothing (Section 10.1). Aggregate lapsing is an observable statistic of the demand for public expenditure, and the levy can be made to respond to it—but only through the worst-case lapsing floor of Section 6, without which the response confuses friction with absent demand and extinguishes the levy in every admissible environment (Section 10.2). And the coalition of Section 8 constrains debt only if the constitutional ceiling names it: a ceiling on revenue alone invites the substitution into borrowing that is the recorded fate of tax limitation (Section 10.3).

10.1 Expiration and the pay-as-you-go margin

Read the mechanism in fiscal-year time. Let P_t denote the year-t designated pool per citizen, with $P_1 = \alpha\tau Y$ at adoption. A fraction $1 - \Lambda_t$ of the pool is directed to PGOs before the year's deadline and a fraction Λ_t lapses; of lapsed revenue $R_t = P_t\Lambda_t$, the rule returns $\psi(R_t)$ to designators and the residual $R_t - \psi(R_t)$ reverts to the general fund. Because claims expire with the year and cannot be pledged against future years (Section 3), outlays on the designated margin in year t satisfy

$$(1 - \Lambda_t)P_t + \psi(R_t) \le P_t, \quad (11)$$

by the feasibility constraint $\psi(R_t) \le R_t = P_t\Lambda_t$, with strict inequality whenever lapsing is positive and $\psi(R_t) < R_t$. Nothing on this margin can be spent before it is levied, and no liability on it survives the year. The observation is one of accounting rather than of behaviour, and it is stated at the level of the mechanism's accounts because the rest of this section rests on it.

Proposition 9 (The designated margin is pay-as-you-go). *In every year t and for every cost distribution $G \in \mathcal{G}$, designated-margin outlays satisfy (11): they cannot exceed the year's designated revenue. Debt attributable to the designated margin is identically zero along every path, and the primary deficit equals the imbalance of the residual budget alone, $D_t = E_t - [(1 - \alpha)\tau Y + (R_t - \psi(R_t))]$, where E_t is general-fund expenditure. The identity is mechanical—it uses no assumption on G, on behaviour, or on the government's objective—and α is the width of the margin it protects.*

The content of the identity is what it excludes, and the exclusion has a price the literature has named. Barro (1979) established the efficiency case for deficits: with distortionary taxation and fluctuating expenditure needs, smoothing tax rates over time—running deficits in the bad years—minimizes excess burden, and a rule forcing the whole budget onto a pay-as-you-go basis forfeits that margin. PTH does not force the whole budget. The residual share $(1 - \alpha)$ retains full access to debt finance and carries the smoothing burden; the designated share is the part of the budget removed from intertemporal finance altogether. The constitutional fraction thus prices a trade the deficit literature usually treats one side at a time: each point of α widens the margin on which deficits are impossible—and, through the stake of Section 8, the coalition's hold on the ceiling—

while narrowing the base over which tax smoothing operates.[7] At the percentage statutes' one or two percent, essentially all smoothing capacity is retained and essentially no deficit margin is created; the phase path that has been put forward for the mechanism elsewhere—α rising stepwise from one half toward larger values as the certified sector matures—traces the trade in the opposite direction, and is recorded here as a family of points on the frontier, not proposed.[8] Nothing in the analysis selects a point on the path: α is where the framer's preferences over discipline and smoothing enter, and the model's contribution is to make the terms of the trade explicit.

The terms can be put in the units the American deficit debate uses. Table 1 scales α against the Congressional Budget Office's February 2026 baseline for fiscal year 2027, in which individual income taxes yield $2,947 billion, revenues $5,885 billion and outlays $7,772 billion, the total deficit is $1,887 billion and the primary deficit $779 billion, and nominal GDP is about $33.3 trillion (Congressional Budget Office 2026a). Because the percentage statutes operate on individual income tax liability, α is applied to that base; the designated pool is ατY read at national scale, and the final column reports what is left of federal revenues to carry the smoothing burden. The exercise is static and illustrative: it holds the baseline fixed and asks only how wide the pay-as-you-go margin would be at each α. It is not a projection of what adopting the mechanism would do to receipts, outlays, or the path of debt, which would require behavioural responses the model does not estimate; Section 13 specifies the projections the conference version adds.

Table 1. *The pay-as-you-go margin at alternative values of α: United States federal magnitudes, fiscal year 2027*

α	Designated pool ($ billions)	Percent of outlays	Percent of GDP	Percent of projected deficit	Residual revenue base (percent of GDP)
0.01	29	0.4	0.09	1.6	17.6
0.02	59	0.8	0.18	3.1	17.5
0.05	147	1.9	0.44	7.8	17.2
0.10	295	3.8	0.88	15.6	16.8
0.25	737	9.5	2.21	39.0	15.5

[7] Footnote 3 records that the model does not optimize α, because with constant welfare weights the choice would be degenerate; the tax-smoothing motive is exactly the curvature in the value of the residual budget, operating across years, that an interior choice of α would require. The present section accordingly treats α positively, as the parameter of a discipline–smoothing frontier, and leaves its optimization open.

[8] The companion essay (footnote 1) put forward an explicit path: one half of each citizen's liability designated at adoption, the ordinary budget retaining the complement, with the designated share rising in steps toward nine tenths as the certified sector matures. That path is not proposed here. It is recorded because it fixes the endpoints of Table 1, and the analysis treats it, like any other path, as a family of points on the frontier of Section 10.1 and selects none of them.

α	Designated pool ($ billions)	Percent of outlays	Percent of GDP	Percent of projected deficit	Residual revenue base (percent of GDP)
0.50	1,474	19.0	4.42	78.1	13.2
0.90	2,652	34.1	7.96	140.6	9.7

Note: Fiscal year 2027, CBO February 2026 baseline (Congressional Budget Office 2026a). Designated pool = α × projected individual income tax receipts of $2,947 billion. Columns 3–5 express that pool as a share of projected outlays ($7,772 billion), of nominal GDP (about $33.3 trillion), and of the projected total deficit ($1,887 billion). The residual revenue base is projected total revenues less the designated pool, in percent of GDP: the base over which the tax smoothing of Barro (1979) must operate. Rows 1 and 2 are the fractions used by the Hungarian and Slovak statutes; rows 6 and 7 are the endpoints of the phase path of footnote 8. Illustrative calculation: the mechanism is not scored here; Section 13 specifies the debt-path projections of the conference version.

Three readings. First, the working statutes are invisible on this margin. Hungary's one percent and Slovakia's two would place $29 billion to $59 billion beyond the reach of deficit finance, against a projected shortfall of $1,887 billion; the composition margin they deliver is real, and the deficit margin is smaller than the revisions CBO routinely makes to a single year's baseline. The fifth positive implication of Section 11.6 is this row of the table. Second, the fractions at which the margin becomes a fiscal object are large. The pay-as-you-go margin matches projected nondefense discretionary outlays at $\alpha \approx 0.33$, projected net interest at $\alpha \approx 0.38$, the projected primary deficit at $\alpha \approx 0.26$, and the projected total deficit at $\alpha \approx 0.64$. The phase path recorded in footnote 8 runs from the sixth row of the table to the seventh, and the seventh is a budget whose designated pool alone exceeds two fifths of federal revenues and a third of federal outlays. Third, the price sits in the same table. At $\alpha = 1/2$ the base carrying the smoothing burden falls from 17.7 to 13.2 percent of GDP, and the shocks that base must absorb are not small: federal deficits reached 9.8 percent of GDP in 2009 and 14.9 percent in 2020. A framer who finds the sixth row attractive is choosing to meet the next such shock on a narrower base, and that is the choice, not an incidental cost of it.

Nothing in the analysis selects a row, and the table is not an argument for any α. What it fixes is that the mechanism's deficit content is a function of α alone; that the content is bounded above by $\alpha\tau Y$, so no drafting of the surrounding clauses can make a small α consequential for debt; and that the fractions the working statutes use put the mechanism below the resolution of the American fiscal problem. Readers who take that as the paper's central negative result are reading it correctly.

There is also a behavioural reading of the expiration rule, and it cuts against the rule's reputation. Inside government, expiring budgets are associated with waste: federal agencies spend a disproportionate share of their appropriations in the final week of the fiscal year, on projects of measurably lower quality (Liebman and Mahoney 2017)—the use-it-or-lose-it pathology. PTH relocates the expiring claim from the agency to the citizen, and the relocation changes what "losing it" means. An agency that fails to spend

surrenders budget to the treasury and signals slack to next year's appropriators; its rational response is to spend on anything. A citizen who fails to direct receives L and forgoes her share of ψ; her account lapses to destinations fixed in advance, and the lapse is informative—it is the raw material of the demand statistic of the next subsection—rather than merely wasteful. The year-end margin does not disappear: citizens direct late as agencies spend late, and the deadline analysis of Section 6 is precisely the theory of that margin on the citizen's side. But the object the deadline disciplines is a designation, not an expenditure, and a lapsed designation destroys no resources: it reallocates them under a rule. Whether relocated expiration outperforms agency expiration is an empirical question the model cannot settle; what the model fixes is the accounting difference between the two objects at risk.

10.2 Revealed demand and the path of the levy

An involuntary fiscal system produces no statistic of the demand for what it finances; that is the oldest complaint against it, and the reason the benefit side of Wicksell's ledger has stayed empty. PTH produces one. At each year's close the treasury observes $1 - \Lambda t$, the fraction of the designated pool citizens took the trouble to direct, and Λt, the fraction they let expire. It is natural—and it was the expiration rule's original motivation in the companion essay—to let the next year's designated levy respond: if citizens direct two hundred of a three-hundred-million-dollar pool, the reading goes, the pool was one hundred million too large, and the following year's levy on this margin can fall accordingly.[9] Formally, a levy rule maps the observed lapsing rate into next year's pool. Lapsing each year is generated by the stationary search environment of Section 5, which does not depend on the pool's size, so $\Lambda t = \Lambda(TR; Gt)$ for the year's population Gt, and with a continuum of citizens the aggregate is deterministic given the environment. Consider two rules:

$$\text{naive: } P_{t+1} = (1 - \Lambda t)\, P_t \; ; \quad (12)$$

$$\text{adjusted: } P_{t+1} = \left[\, 1 - \max\{\, 0, (\Lambda t - \Lambda w(TR)) / (1 - \Lambda w(TR)) \,\} \,\right] P_t \; , \quad (13)$$

where $\Lambda w(TR)$ is the worst-case lapsing rate of Section 6 at the constitutional deadline—an object computable from the structural primitives (δ, L, Φ) alone, before any population is observed. The naive rule treats every lapsed dollar as a dollar of absent demand. The adjusted rule treats lapsing up to the friction floor as uninformative—it is what the mechanism itself generates even when every citizen values it—and contracts only on the excess.

Proposition 10 (Friction-adjusted revealed demand). *Maintain Assumptions 1, 3, and 4, fix the constitutional deadline at TR, and write $pmin = p(\bar{c}; TR)$. (i) Under the naive rule (12), $\Lambda t \geq \delta$ for every population, with or without Assumption 2, so the pool*

---

[9] The illustration, in the stylized municipality of the companion essay: a city of seventy thousand residents levies a designated pool of three hundred million dollars; two hundred million is directed by the deadline and one hundred million lapses. All numbers are illustrative. The naive reading—cut the pool by a third—is the one Proposition 10(i) rejects; the adjusted reading nets out first the lapsing the mechanism's own frictions would generate at that deadline, and contracts only on the remainder.

contracts by a factor of at least 1 − δ per year and converges to zero: the naive rule extinguishes the designated margin in every admissible environment, including those in which every citizen strictly values designation. (ii) Under the adjusted rule (13), if Assumption 2 holds for the entire population in every year—universal demand—then Λt ≤ Λw(TR) for every Gt ∈ 𝒢, the correction is zero, and the pool is never reduced. (iii) If instead a mass ν > 0 of citizens strictly prefers never to evaluate, so that Assumption 2 fails for them as in the remark on Assumption 2 in Section 5.4, then Λt ≥ ν + (1 − ν)pmin for every admissible configuration of the remaining types, and the adjusted rule contracts the pool whenever ν exceeds the detection bound ν̄ = (Λw(TR) − pmin)/(1 − pmin); any ν ≤ ν̄ can evade detection for some admissible configuration of the remaining types.

The proposition is the section's version of the paper's method, and its halves should be read together. The naive rule fails not marginally but completely, and for a reason the single-cycle analysis already contains: Λt ≥ δ always—attrition claims its share of even the most eager population—so a rule that reads all lapsing as absent demand compounds a strictly positive contraction forever. The adjusted rule is conservative by construction: it forgives lapsing up to the worst case over the uncertainty set, so pockets of genuine non-demand of mass below ν̄ can hide inside the friction floor and go undetected. That is the price of robustness on this margin, and it is the price paid everywhere in the paper: the guarantee protects against every admissible environment at the cost of power against particular ones. What the framer buys is one-sided safety. The adjusted levy never ratchets down a margin citizens value, under any admissible population; it ratchets down only on evidence inconsistent with universal demand at the assumed primitives.

Three remarks attach. First, the rule conditions on (δ, L, Φ) through Λw(TR), and the asymmetry recorded in Section 12 applies with full force: a framer who misestimates the attrition hazard misplaces the friction floor, and the one-sided safety of part (ii) is only as good as the primitives it is computed from. Second, the rule is responsive, and Section 6 closed with a warning against responsive deadlines; the two positions are consistent, and the difference is worth stating. A responsive deadline conditions on a statistic—realized lapsing—that the unknown distribution contaminates at every level of its range, and chases a per-distribution optimum that Proposition 1 shows is not implementable. The responsive levy conditions only on the excess of that statistic over its worst-case bound, a threshold no admissible friction can cross; it chases nothing, and in the environments the framer cannot rule out it does nothing at all. Responsiveness is safe exactly where the uncertainty set makes a region of the statistic's range unreachable by friction; the deadline has no such region, and the levy does. Third—a limit rather than a remark—the levy rule as written adjusts downward only. A symmetric upward response, raising the pool when lapsing is far below the floor, has no analogous safety: low lapsing is consistent with every admissible environment, since designation is capped at the pool and excess demand leaves no trace in the lapsing statistic, so no excess-over-bound logic is available and an upward ratchet would be exactly the responsive chasing Section 6 warns against. The mechanism reveals over-levy robustly; it does not reveal under-levy at all.

10.3 The borrowing clause: what the coalition constrains

The coalition of Section 8 blocks increases in the constitutional tax ceiling. Whether it thereby constrains debt turns on a drafting question the analysis can pose precisely: is the ceiling written on revenue, or on revenue plus net borrowing? Under a revenue-only ceiling the government's problem has an unpriced instrument. A revenue gain $\Delta\pi$ blocked at the ceiling can be replaced by issuing debt, which requires no coalition's consent, shifts the levy to fiscal years the current coalition does not sit in, and—since a non-designator's stake w is her loss from a higher current tax—meets no stake at all. The prediction is the substitution that tax limitation has historically produced, and that Buchanan and Wagner (1977) made the centre of the constitutional case against unconstrained deficit finance: constrain one instrument, and the pressure exits through the other. The state evidence reads the same way in reverse—the balanced-budget institutions that bind are those that close the deficit-carryover margin, not those that regulate a single year's revenue (Poterba 1994, 1996).

The states supply more than an analogy here, because they have run the experiment on both instruments for a century and a half, and the conference's invitation to read the experience of the US states is an invitation to read it against this clause. Where constitutions restrict general-obligation borrowing, the restriction relocates debt rather than retiring it. Von Hagen (1991) finds substitution out of restricted and into unrestricted debt instruments, with no reduction in debt per capita; Bunch (1991) finds that states whose debt limits bind incorporate more public authorities, assign them more functions, and route more of their infrastructure finance through them; Kiewiet and Szakaly (1996) find that lower state borrowing under constitutional limits is accompanied by higher local borrowing, so that the ceiling moves the liability down one level of government. The revenue side records the mirror image. Tax and expenditure limits—ceilings on revenue in the sense of this paper—are associated with higher state debt where they are restrictive on revenues or on expenditures singly, and with lower debt only where they name both margins together (Deller et al. 2013); at the local level they shift the revenue structure away from taxes and toward charges and miscellaneous revenue rather than reducing what is financed (Shadbegian 1999). The closest institutional relative of the ceiling of Section 8 is the supermajority requirement for tax increases, which does hold tax rates down in the states that adopt it (Knight 2000)—which is precisely the result that leaves the borrowing margin open. This section reads that evidence as a constraint on drafting and offers no estimates of its own; the estimation it calls for is specified in Section 13.

Two things follow for the drafting question, and the second is the more useful one. The first is that a ceiling on revenue alone is not a weaker version of a ceiling on revenue plus net borrowing; it is a different instrument, with an escape route the American record has already mapped. The second is that the coalition of Section 8 does not repair the omission. Its advantage over a legislative supermajority is that it cannot be bought, because it cannot be seen; that advantage is an advantage in enforcing whatever the clause names, and confers nothing on a margin the clause leaves out. A mechanism whose enforcement technology is stronger than the states' and whose clause is drafted as theirs were would substitute into debt exactly as theirs did, and faster.

Write the clause instead on revenue plus net borrowing, and Propositions 4 and 5 apply verbatim with $\Delta\pi$ reinterpreted as the prize from relaxing the combined constraint by either instrument: the blocking condition $w \cdot \zeta/\Lambda \geq \Delta\pi$ and the opacity threshold $\rho*$ carry over unchanged, because nothing in their derivation used the instrument's name. One asymmetry does not carry over, and should be recorded rather than assumed away. A tax increase burdens coalition members now, which is what the stake w prices; a bond issue burdens fiscal years in which the present coalition may not sit, so the stake behind the borrowing half of the clause is the coalition's concern for those years—plausibly smaller than w, and zero for a coalition that discounts fully. The clause is therefore only as strong as the overlap between the citizens it empowers and the citizens the borrowing burdens; a full treatment needs the overlapping-generations formulation the conclusion lists among extensions, and we flag the gap. What the single-generation analysis establishes is the drafting result: under the revenue-only clause the mechanism's discipline does not reach the deficit at all, and under the combined clause it reaches the deficit exactly as far as Propositions 4 and 5 reach, no further.

11. PTH among fiscal institutions

A constitutional provision earns its place by comparison, not by assertion. This section sets PTH beside the institutions it most resembles—including the deficit institutions nearest it in purpose—and then returns what it owes to the percentage statutes: positive implications the model can actually deliver.

11.1 Functional, overlapping and competing jurisdictions

Frey and Eichenberger's (1996, 1999) FOCJ attack the same bundling from the jurisdictional side: let functions be provided by overlapping, competing jurisdictions that citizens or communes join and fund separately, so that exit operates function by function rather than bundle by bundle. The kinship is real—both designs unbundle the budget—but the margins differ. FOCJ unbundle across governments: they require founding, joining, and governing new jurisdictions, with the fixed costs, minimum scales, and collective-choice overhead that entails, and their discipline still runs through membership decisions that carry relocation-like switching costs. PTH unbundles within the treasury at the individual level: no jurisdiction is founded, no membership changes, and the marginal act is a designation on a tax return. The two are complements rather than rivals—FOCJ for functions with scale economies and club structure, PTH for the residual composition margin inside any given fisc—and both descend from the polycentric tradition of Ostrom, Tiebout and Warren (1961). The comparison also locates PTH's distinctive vulnerability: because designation requires no coordination, it also builds no governance, and the quality of what is funded must be disciplined by certification and competition (Section 9) rather than by member voice.

11.2 Earmarked taxes

Buchanan (1963) rescued earmarking from the charge of fiscal illusion by reading it as piecewise fiscal choice: tying a revenue source to a use lets the voter-taxpayer purchase public goods item by item rather than in a single all-or-nothing bundle. PTH is earmarking individualized. Instead of one tie chosen collectively and applied to everyone

—with the majority-cycling and log-rolling that collective choice over earmarks invites—each citizen ties a share of her own liability to a certified use, and the aggregate composition emerges from the distribution of those ties. The collective version buys commitment at the price of uniformity; the individual version keeps the commitment (α is constitutionally fixed, so the general fund is protected in the aggregate) while releasing the uniformity along the one dimension where uniformity has no rationale, the composition of citizens' marginal public consumption.

11.3 Charitable deductions and credits

Tax concessions for giving subsidize private contributions out of net income: the citizen surrenders her own money and the fisc matches at the rate her bracket implies. Three differences follow. PTH redirects the tax itself—no private gift is made, so participation is free in money and costly only in the attention the model prices; the entire design problem of Sections 6–9, deadlines and lapse rules and certification margins, exists precisely because attention rather than generosity is the scarce input, and that problem has no counterpart in deduction design. The deduction's subsidy rate rises with the marginal tax rate, so its allocative weight tilts toward high-bracket itemizers; PTH's fraction α is uniform by construction—the generality norm of Buchanan and Congleton (1998) applied to the instrument itself. And the deduction moves resources out of the fisc into a private nonprofit sector, while PTH keeps them inside a certified public-purpose sector under the constitution's own oversight, which is what makes the enforcement coalition of Section 8 possible at all.

11.4 Participatory budgeting

Participatory budgeting allocates a share of a municipal budget through assemblies and votes (Cabannes 2004). It is a voice institution: deliberative, collective, and capable of surfacing information and legitimacy that no designation mechanism can. Its weaknesses are voice's weaknesses—participation is costly in time and therefore selective, assemblies can be captured by the organized, and the outcome is again a single collective composition. PTH is the mirror image: statistical rather than deliberative aggregation, trivially low participation cost at the margin, capture-resistant in the aggregate because the aggregate is not chosen by anyone—but for the same reasons deliberation-free and legitimacy-thin. A constitution could sensibly run both, on different slices; nothing in the analysis forces a choice.

11.5 Balanced-budget rules, tax limits, and expiring appropriations

The deficit institutions nearest PTH in purpose are furthest from it in grammar. A balanced-budget rule or a tax and expenditure limit is written on the government: it names a total, and its enforcement is assigned—to a court, a comptroller, a supermajority procedure—so its bite varies with the assignee's independence and with the substitution margins the drafting leaves open, which is what the state evidence measures (Poterba 1994, 1996) and what Section 10.3 prices: restricted borrowing reappears as unrestricted borrowing and as local debt, and restricted revenue as charges (von Hagen 1991; Bunch 1991; Kiewiet and Szakaly 1996; Shadbegian 1999; Deller et al. 2013). PTH's deficit margin is written on no total and assigned to no one. The pay-as-you-go property of

Section 10.1 is an accounting identity of expiring individual claims—there is nothing to enforce—and the ceiling of Section 8 is enforced by a coalition that exists as a by-product of individual lapsing and is protected by opacity rather than by office. The comparison also runs through information. A balanced-budget rule uses none: it binds identically whether the budget it balances is wanted or not. The levy rule of Section 10.2 is built from a demand statistic the institution itself generates, and its robust form is exactly as responsive as that statistic's friction floor allows and no more. Halac and Yared (2014) derive fiscal rules as optimal delegation under private information about the temptation to spend—caps, with escape clauses priced; the present analysis reaches cap-shaped rules from the framer's side of the veil, and the two exercises are complements: theirs asks what rule a principal with a prior over temptations writes for a government, ours what rule a framer with no prior over populations writes for citizens. Expiring appropriations, finally, are the mechanism's closest working relative on the time margin, and its cautionary tale: expiration inside the agency produces the year-end spending surge and its quality collapse (Liebman and Mahoney 2017), and Section 10.1 records why citizen-side expiration changes the object at risk—a lapsed designation reallocates under a rule, while a year-end obligation spends. The behavioural family resemblance—deadline effects—is real, and Section 6 is offered as their theory on the citizen's side.

11.6 Positive implications for the percentage statutes

The percentage-philanthropy statutes are PTH's working prototype, and the model returns five positive characterizations to them. First, on deadlines: a designation window—typically the annual filing season—that is fixed structurally and left alone is the worst-case-optimal rule. The legislative practice of shortening the window when undesignated shares are high fails twice over (Section 6): lapsing is non-monotone in the window, and chasing the realized optimum is itself non-robust. Second, on lapsed shares: every existing statute either lets undesignated slices revert silently to general revenue or, as Italy's otto per mille does, distributes them among recipients in proportion to the choices expressed; none returns them to the taxpayers who designated, so each is $\psi = 0$—feasible but not optimal in the model. Returning lapsed revenue to designating taxpayers at a constant marginal rate, capped by (7), would reward participation without undoing the penalty that disciplines the fisc; the constant rate, not the level, is the robust feature. Third, on eligibility: within the model, certification rules of the margin-cap form—limits on the share of directed funds a recipient may absorb in administration—dominate performance targets and quality floors, which presuppose knowledge of the sector's cost structure that no legislature has (Section 9). Existing statutes already gesture at this with use-of-funds restrictions; the model implies the gesture is the right instrument in the worst-case sense, not a second best. Fourth, on confidentiality: the analysis assigns enforcement value—not merely privacy value—to shielding the individual designation record from the spending side of government. Section 8 prices the asset: the ceiling-enforcing coalition survives exactly as long as the government's signal about who designated stays below $\rho*$. Fifth, on scale: at the statutes' actual fractions—one or two percent—the stake w that non-designators hold in the tax ceiling is small and the enforcement channel is correspondingly weak. The constitutional version of the mechanism needs an α large enough for the coalition's aggregate stake $w \cdot \zeta/\Lambda$ to exceed plausible revenue prizes; the statutes as they stand deliver the composition margin but

neither the Leviathan margin nor, by Section 10.1, any appreciable deficit margin: at one or two percent the pay-as-you-go share of the budget is negligible—Table 1 puts the American equivalent at $29 billion to $59 billion, under two tenths of one percent of GDP, against a deficit projected at $1,887 billion. The phase path recorded in Section 10.1 is a positive description of how the three margins would scale together; it is recorded, not selected.

Honesty requires the reverse ledger too. Individual designation fragments funding and can starve goods with concentrated costs and diffuse benefits; the model confines PTH to a slice α precisely so that the general fund continues to carry such goods, and the within-bundle gap it closes is the gap among certifiable, organization-providable purposes, not the whole of public finance. Designation may be expressive rather than informed—citizens direct to what flatters rather than what serves—which is why the mechanism's design leans on certification and competition rather than on the wisdom of any single choice. And the certified sector itself consumes resources: the floor τmin = n∗F of Section 9 is a real cost, and a jurisdiction whose feasible α is tiny may find the administrative floor eats the allocative gain. And the deficit margin of Section 10 is only as wide as α and only as binding as the borrowing clause: a jurisdiction that adopts the revenue-only clause acquires the composition margin and none of the deficit discipline—a drafting difference invisible in the mechanism's description and decisive in its accounts. These are the terms on which the provision would have to be evaluated, and the model is built so that each term is visible.

12. What robustness buys, and what it costs

Worst-case design is conservative, and candor about its price is part of using it well. The robust deadline TR, the constant-rate lapse rule, and the opacity threshold ρ∗ guarantee the worst-case envelope, but for any particular realized population they are beaten by the Bayesian-optimal mechanism tuned to that population. The case for the robust mechanism is not that it does better against the true distribution; it is that the framer cannot identify the true distribution and cannot revise the rule once it is wrong. Where a credible prior exists and the rule can be updated, the Bayesian mechanism is the right one. The robust analysis is the relevant benchmark precisely in the constitutional case, where neither condition holds—which is the complementarity, indexed by reversibility, that Dworczak and Pavan (2022) formalize and Carroll (2019) surveys.

The uncertainty set is itself a modelling choice, and should be defended rather than assumed. The results are stated for the shape-free class on a known support. Widening it—dropping the support bound—weakens every guarantee, since the adversary gains room. Narrowing it—positing a known mean, a known tail, log-concavity—sharpens the guarantees but reintroduces detail-dependence, and each such assumption must be one a framer can credibly claim. The discipline is to let the set encode exactly the coarse, verifiable knowledge a constitution can presuppose, and to exclude the fine knowledge it cannot; under that discipline what the framer assumes is visible in the set rather than hidden in a prior, itself an advantage of the approach. One asymmetry in our set deserves explicit statement: the framer is assumed ignorant of G yet to know the attrition hazard δ and the match-value distribution Φ exactly. An attrition rate is at least as hard to know a priori as a cost distribution, and the robust deadline depends on δ finely, so "detail-free"

in this paper means free of the shape of G, not free of all fine detail. The levy rule of Section 10.2 inherits the same exposure through the floor Λw(TR), and part (ii) of Proposition 10 carries that proviso. A joint uncertainty neighbourhood over (G, δ, Φ) would change the analysis of Section 6 rather than decorate it, and remains open.

Finally, the two halves of Section 6 should be kept distinct. On the shape-free class the maxmin criterion selects a single known type, so the ignorance of G does no work and the substance lies in the two-channel non-monotonicity, the interiority certificate, and the non-implementability of per-type optima. With a known mean the criterion binds: the worst case becomes a two-point law on the endpoints of the cost support, the support bound does real work, and the robust deadline moves with the asserted mean. The methodological claim of the paper lives in the second half; the first half is reported with equal prominence because a criterion defended by its honesty should be seen not binding where it does not.

13. The conference version: what the May 2027 paper adds

The paper as it stands establishes three things about a fiscal institution written on the citizen, each a statement about accounts, guarantees, or drafting rather than about a measured trajectory. Proposition 9 and Table 1 fix the width of the budget that expiring individual claims withdraw from deficit finance, and the width is a static scaling, not a path. Proposition 10 fixes the one levy rule that reads realized lapsing without mistaking friction for absent demand, and it conditions on primitives the paper assumes rather than measures. Section 10.3 fixes the drafting condition under which the enforcement coalition reaches borrowing at all, and it reads the state evidence rather than adding to it. The version of the paper prepared for the May 2027 conference adds one empirical component for each. Their designs are fixed now; their estimates are not, and nothing below anticipates a result. One theoretical component joins them: the search-consistent supply side in the framework of Wolinsky (1986) and Anderson and Renault (1999), which Section 9.1 defers here after recording why the demand-side guarantees do not wait on it.

First, a panel of the percentage-designation statutes. Since Hungary’s Act CXXVI of 1996, taxpayers in Hungary, Slovakia, Lithuania, Poland, and Romania, and in Italy under the otto per mille of 1985 and the cinque per mille of 2006, have been able to direct a fixed share of income tax liability to eligible organizations, and the tax administrations of these countries publish annual statistics on the exercise of the right—in the fullest cases the number of taxpayers entitled and the number designating validly, the amounts designated, left undesignated, and rejected as invalid, and the recipients and their receipts; Hungary’s National Tax and Customs Administration, for instance, reports these series annually to a committee of Parliament. Assembled across countries and years, the record yields the objects the model names and that no American data can supply: the lapsing rate Λt, as the share of the designated pool left to expire; its movement across the changes in designation windows, eligibility rules, and fractions that these statutes have repeatedly undergone, which is the comparative static of Section 6; and the concentration of directed funds across recipients, the object of Section 9. It also yields a decomposition the model asks for. Where an administration reports invalid designations separately from non-designation, one component of friction is observed apart from absent demand, and

the naive and adjusted levy rules of Section 10.2 can be run against a lapsing series in which part of the friction floor is measured rather than assumed. The panel is the paper's contribution to the experience of other nations, and it will be reported as measurement: at the statutes' fractions the mechanism sits below the resolution of the American deficit (Table 1), so what the statutes inform is the behaviour of citizens holding expiring claims, not the trajectory of American debt.

Second, state-level estimates of the borrowing clause. Section 10.3 reads the American state evidence as a constraint on drafting and offers no estimates of its own; the conference version estimates the drafting result directly. Using the Census Bureau's annual state government finance series for debt outstanding, revenue, and their composition, together with the classifications of state tax and expenditure limits, balanced-budget requirements, and constitutional debt limits assembled in the literature (Poterba 1994; Kiewiet and Szakaly 1996; Deller et al. 2013), it classifies each limit by what its clause names—revenue, borrowing, or both—and compares the paths of state and of local debt after adoption and amendment across clause types. The substitution that von Hagen (1991), Bunch (1991), Kiewiet and Szakaly (1996), and Shadbegian (1999) document in cross-section should appear as a divergence of debt paths by clause type, and the drafting result of Section 10.3 predicts its direction: limits written on revenue alone should be followed by more unrestricted, authority, and local debt, and limits written on both margins by less. The design is an event study around clause adoption; its identifying assumption, and the limits of staggered adoption by states that adopted for different reasons at different times, will be stated as such.

Third, debt paths under the extended baseline. Table 1 holds the baseline fixed and reports widths; the conference version reports illustrative paths. For each α in Table 1 it projects debt held by the public as a share of GDP from 2027 to 2056 under the Congressional Budget Office's February 2026 extended baseline (Congressional Budget Office 2026b). Proposition 9 fixes the accounting: no debt accumulates on the designated margin, so the path is determined by the residual budget—by the rate at which designated provision displaces general-fund outlays, from zero, where certified-sector provision is additional and the residual budget borrows the revenue it no longer collects, to one, where it replaces general-fund provision dollar for dollar—and by whether the residual primary balance responds to debt with the corrective coefficient Bohn (1998) estimates. The displacement rate is exactly the behavioural response the single-cycle model does not estimate; the projection treats it as the explicit parameter it is and reports paths across its range rather than at a point. Lapsing is set at the rates the percentage panel records, so that the levy rules (12) and (13) run against observed rather than assumed friction, and lapsed funds enter the residual budget under the statutes' $\psi = 0$. The output is a set of scenarios rather than structural estimates: a debt-to-GDP path at each α, indexed by the displacement rate and banded over the lapsing rate. Its qualitative content is known in advance from Proposition 9—the mechanism fixes what a deficit cannot reach, not the deficit—and the figure's purpose is to put a scale on the displacement rate at which the path departs from the baseline, which is the quantity on which the institution's effect on the trajectory of debt turns.

Two limits are set now. The conference version will not select a value of α or a clause, and it will not read the percentage panel as a calibration of the American case: the

statutes operate at fractions that Table 1 has already placed below the resolution of the federal deficit, and what they measure is the behaviour of citizens holding expiring claims. The first two components are scheduled for completion by November 2026, the third by the conference.

14. Conclusion

A fiscal constitution should be judged by what it guarantees against what its framer cannot know. We have studied Personal Tax Holding—a constitutionalized, generalized descendant of the percentage-philanthropy statutes—for the within-bundle problem under that standard, replacing the framer's prior with an uncertainty set and evaluating each rule by its worst-case performance. The findings are a short list of features that survive the framer's ignorance: a designation deadline that is interior, detail-free, and set to protect the most patient citizen—or, when a mean can be asserted, to balance her against the most cost-burdened one; a lapse rule with a constant marginal rate, for the reason Leviathan-constraining tax rules are flat; an enforcement coalition of passive non-designators whose blocking power rests on fiscal privacy and degrades only continuously as the government learns, with a computable breaking point; and a certification instrument that caps margins and delegates quality to competition; a designated margin that is pay-as-you-go by construction, pricing the trade between deficit discipline and tax smoothing in the single parameter α; a levy that responds to revealed demand only through the friction floor, and so never contracts a margin citizens value; and a ceiling whose reach over debt is a drafting fact—revenue-only clauses leak into borrowing, while the combined clause inherits the coalition's full blocking power—priced by the same stake and the same opacity threshold. Stripped to this detail-free core, the mechanism meets the constitutional standard; its finely tuned counterpart does not.

Four extensions seem most valuable. An overlapping-generations constitution, in which successive populations are drawn from distributions the framer cannot forecast, would let responsive deadline rules be compared honestly with fixed ones, and would let the borrowing half of the combined clause of Section 10.3 earn—or lose—the intergenerational stake the single-generation analysis cannot supply. Richer moment information—variances, tails—would trace how the guarantees sharpen as the framer's credible knowledge grows, extending Section 6.2. Deriving the enforcement ceiling of Section 7 from the coalition of Section 8, and the designators' zero stake from the lapse rule itself, would close the mechanism's internal loop. And reconciling the supply side with the search model, in the manner of Wolinsky (1986) and Anderson and Renault (1999), would discharge the paper's largest acknowledged debt; Section 9.1 records why the demand-side guarantees do not wait on it, and the conference version undertakes the formal treatment. A fifth item is empirical rather than theoretical, and it is the one the conference version takes up first: Table 1 fixes the width of the pay-as-you-go margin at each α, but a width is not a debt path, and Section 13 specifies the percentage-statute panel, the state estimates, and the debt-path scenarios that turn widths into measured behaviour and paths. Each extension tests whether the detail-free core survives; the constitutional standard says that is the only test that matters.

Appendix A. Proofs

A.1 Proofs for Section 6

Proof of Lemma 1. Fix T. The map G ↦ Λ(T; G) = ∫ p(c; T) dG(c) is affine in G. The uncertainty set $\mathcal{G}$ is the set of Borel probability measures on the compact interval [0, c̄], which is convex and weak-star compact (Banach–Alaoglu, or Prokhorov). When p(·; T) is continuous and bounded, G ↦ ∫ p dG is weak-star continuous; by the Bauer maximum principle an upper semicontinuous affine functional on a compact convex set attains its maximum at an extreme point, and the extreme points of the set of probability measures on [0, c̄] are exactly the point masses δc. Therefore supG ∈ $\mathcal{G}$ Λ(T; G) = supc ∈ [0, c̄] p(c; T). It remains to show p(·; T) is continuous. By (1), each Vs(·; c) is obtained from Vs−1 by a maximum of x with an affine functional of an expectation against the fixed kernel Φ; these operations preserve joint continuity in (x, c), so Vs is continuous in c for each s ≤ T. The reservation rs(c) is defined by a strictly monotone indifference condition and is therefore continuous. The lapsing probability p(c; T) is a finite sum of products of Φ-masses evaluated at the thresholds {rs(c)} and L, hence continuous in c. Continuity on the compact [0, c̄] makes the supremum a maximum, attained at a type—a degenerate distribution. ∎

Proof of Lemma 2. Fix T. Step 1 (reservations fall in cost). From (1), the value function is nonincreasing in c for every (s, x), by induction: at s = 0 it equals x and does not depend on c, and if it is nonincreasing in c with s − 1 periods left, then the continuation value −c + (1−δ)E[Vs−1(max{x, U}; c)] + δL is strictly decreasing in c, so the maximum defining Vs(x; c) is nonincreasing in c. The stopping value x does not depend on c while the continuation value strictly decreases in it, so the unique crossing point rs(c) is nonincreasing in c, strictly decreasing wherever it is interior. Step 2 (coupling). Fix a realization of the draw sequence U1, U2, … and of the attrition period τ, and let xt = max{L, U1, …, Ut} denote the running best option, a function of the draws alone. For a type c define the directing time σ(c) = min{ t ≤ T : xt ≥ rT−t+1(c) }, with σ(c) = ∞ if no such t exists. By Step 1 the reservations are nonincreasing in c while xt does not depend on c, so c ≤ c′ implies σ(c′) ≤ σ(c) on every path. A citizen lapses on a path exactly on the event {τ ≤ σ(c)} ∪ {σ(c) > T}; since σ(c′) ≤ σ(c), the lapsing event for c′ is contained in the lapsing event for c. Taking probabilities, p(c′; T) ≤ p(c; T) whenever c ≤ c′. If rs(c) > rs(c′) for some s ≤ T, then by Assumption 3 a draw falls strictly between the two reservations with positive probability, on which c′ directs and c does not; since δ < 1 this event occurs before attrition with positive probability, and the inequality is strict. Hence the maximum of p(·; T) over [0, c̄] is attained at c = 0 and equals p(0; T). ∎

Proof of Proposition 1. Throughout, Λw(T) = maxc p(c; T) by Lemma 1. Step 1 (T = 1). With one period a type-c citizen starts at x = L and, by Assumption 2, finds an evaluation worthwhile. She survives attrition with probability 1 − δ and draws U; she directs if and only if U > L, and lapses otherwise or if attrition strikes. Hence p(c; 1) = δ + (1 − δ)Φ(L) for every c, so Λw(1) = ΛS. Step 2 (T → ∞). Fix c. As the remaining horizon s grows, rs(c) increases to the stationary reservation r∞(c), the fixed point of the indifference implied by (1), characterized by c + δ(r − L) = (1 − δ) ∫ from r to umax of (u − r) dΦ(u); the attrition term δ(r − L) must not be dropped, since at c = 0 it alone pins the stationary reservation strictly below the top of the support, and r∞(c) is strictly decreasing in c. In

the stationary regime each period delivers attrition with probability δ and designation with probability (1 − δ)(1 − Φ(r∞(c))), so the probability that attrition arrives before designation is the geometric ratio q∞(c) = δ/[δ + (1 − δ)(1 − Φ(r∞(c)))]. The finitely many early periods carry higher reservations but contribute vanishing probability as T → ∞, so p(c; T) → q∞(c) pointwise. By Lemma 2 the maximum in Lemma 1 is attained at c = 0 for every T, so pointwise convergence at c = 0 suffices; q∞ is decreasing in c, whence limT→∞ Λw(T) = q∞(0) = ΛD. Step 3 (bound). Fix c and an integer n. A type-c citizen lapses only if attrition strikes within n periods—probability 1 − (1 − δ)n—or no attrition occurs and she never draws a value above L, in which case she evaluates every period and all of her at most n draws lie at or below L—probability at most Φ(L)n. By the union bound p(c; n) ≤ M̄(n), so Λw(n) ≤ M̄(n). Step 4 (characterization). By definition TR is interior precisely when the minimum of Λw over deadlines is strictly below both corner values Λw(1) = ΛS and limT Λw(T) = ΛD, which is IDC. If minn M̄(n) < min{ΛS, ΛD} with minimizer n∗, then Λw(n∗) ≤ M̄(n∗) < min{ΛS, ΛD}, so IDC holds; since Λw(T) → ΛD > Λw(n∗), the set {T : Λw(T) ≤ Λw(n∗)} is nonempty and bounded, so Λw attains its minimum at a finite TR ≥ 2. When IDC fails the minimum of Λw equals min{ΛS, ΛD} and is attained at a corner: at T = 1 when ΛS ≤ ΛD, and as a limit along T → ∞ when ΛD < ΛS. Step 5 (detail-freeness and the per-type optimum). Each of Λw, ΛS, ΛD, and M̄(n) is determined by (δ, L, Φ) alone—by Lemma 2 the maximization is solved at c = 0, so c̄ enters only through Assumption 2—and TR inherits this dependence. Finally, because r∞(c) is strictly decreasing in c, a higher-cost type holds out for less and directs sooner, spends fewer periods exposed to attrition, and the marginal attrition cost of lengthening the deadline falls in c; the minimizer T∗(c) of p(c; ·) is therefore increasing in c over the regime in which the interior optimum obtains, and no single T∗(c) is implementable without knowledge of the realized composition of types. ∎

Proof of Proposition 2. Write cav p for the concave envelope of p(·; T) on [0, c̄]. Since cav p ≥ p and cav p is concave, Jensen's inequality gives ∫ p dG ≤ ∫ cav p dG ≤ cav p(∫ c dG) = cav p(m) for every G ∈ 𝒢(m). Conversely, p(·; T) is continuous on the compact interval [0, c̄] by Lemma 1, so its concave envelope at m is attained as a convex combination λ′·p(c1) + (1 − λ′)·p(c2) with λ′·c1 + (1 − λ′)·c2 = m; the two-point law with these weights belongs to 𝒢(m) and attains the bound. When p(·; T) is convex its concave envelope on [0, c̄] is the chord joining (0, p(0; T)) and (c̄, p(c̄; T)), so c1 = 0, c2 = c̄, and λ′ = 1 − m/c̄. ∎

A.2 Proof for Section 7

Proof of Proposition 3. A rule is a function ψ on the realizable range of lapsed revenue, with R = ατY·Λ and Λ ∈ [0, 1]. Feasibility requires 0 ≤ ψ(R) ≤ R and enforcement-compatibility ψ(R) ≤ B(R). As G varies over 𝒢, the aggregate Λ(T; G) = ∫ p(c; T) dG(c) takes every value between minc p(c; T) and maxc p(c; T)—a nondegenerate interval for T ≥ 2, since p(·; T) is continuous and non-constant and intermediate values are attained by mixing the point masses at the minimizing and maximizing types; at T = 1, p(·; 1) is constant by Step 1 of the proof of Proposition 1, R is degenerate and the statement is vacuous. Realized R thus ranges over a nondegenerate interval I. Welfare decomposes as W(ψ; G) = W0(G) + γgR + (γd − γg)ψ(R), where W0(G) and R are determined by G; since γd > γg, W is strictly increasing in ψ(R) at each realized R. Hence if ψ(R) ≥ ψ′(R)

for all R ∈ I, then W(ψ; G) ≥ W(ψ′; G) for every G, so V(ψ) ≥ V(ψ′). The pointwise-largest admissible rule is ψ∗(R) = min{R, B(R)}, the lower envelope of the two ceilings; therefore V(ψ∗) ≥ V(ψ) for every admissible ψ. The function ψ∗ is the minimum of an increasing affine map and a decreasing affine map, hence concave and piecewise linear with a single kink at R0 = αtY|L|/(αtY + |L|); for R > R0, ψ∗(R) = B(R) with constant marginal rate −|L|/(αtY). For domination, any admissible ψ̃ differing from ψ∗ on a subset of I of positive length satisfies ψ̃ < ψ∗ there—it cannot exceed ψ∗ without violating a constraint—so some G gives W(ψ̃; G) < W(ψ∗; G), and taking that G in the minimum gives V(ψ̃) < V(ψ∗); in particular any rule with a non-constant marginal rate on I differs from ψ∗ and is dominated, and ψ∗ is the unique worst-case-undominated admissible rule. On the binding region ψ∗ has constant marginal rate while its average rate ψ∗(R)/R falls with R; linearity is a property of the marginal, not the average, rate. ∎

A.3 Proofs for Section 8

Proof of Proposition 4. With μ ∈ (1 − Λ, 1] the consenting mass absent payments is 1 − Λ < μ, so the ceiling holds and the government must secure ζ = μ − (1 − Λ) > 0 additional consents from non-designators. Consider any side-payment scheme conditioned only on variables independent of designation status di. To obtain a non-designator's consent the scheme must pay her at least w, so any scheme securing ζ non-designator consents pays at least w to a set S containing at least ζ non-designators. Because the conditioning variables are uninformative about di, the government cannot raise the non-designator share of S above the population rate Λ; hence the mass of S is at least ζ/Λ, and since every paid member of S collects the transfer—designators accept it as well—the cost is at least w·ζ/Λ. By the calibration w·ζ/Λ ≥ Δπ, no budget-feasible scheme yields positive net surplus, and the coalition blocks. ∎

Proof of Proposition 5. Let the government observe σi ∈ {ND, D} with Pr(σi = di) = ρ ∈ [1/2, 1], symmetric across states. By Bayes' rule the share of true non-designators among citizens flagged ND is π(ρ) = ρΛ/[ρΛ + (1 − ρ)(1 − Λ)]. The cheapest breaching scheme offers w to citizens flagged ND; to net ζ true non-designators it must cover a flagged set of mass ζ/π(ρ), at cost K(ρ) = w·ζ/π(ρ). At ρ = 1/2, π = Λ and K = w·ζ/Λ, recovering Proposition 4; at ρ = 1, π = 1 and K = w·ζ. Since π is strictly increasing in ρ, K is continuous and strictly decreasing. The government breaches if and only if Δπ > K(ρ). When w·ζ < Δπ < w·ζ/Λ, the value w·ζ/Δπ lies strictly between π(1/2) = Λ and π(1) = 1, so ρ∗ solving π(ρ∗) = w·ζ/Δπ exists in (1/2, 1) and is unique; for ρ < ρ∗ the coalition blocks, for ρ > ρ∗ it fails, and the blocking power K(ρ) is continuous and strictly decreasing in ρ. ∎

A.4 Supply side: derivation and proofs for Section 9

Derivation of (8)–(9). Take n organizations evenly spaced on the circle, all at quality q, with one deviating to q′. For q′ near q the indifferent-citizen condition v(q′) − t·d = v(q) − t·(1/n − d) gives arc-distance d∗(q′) = 1/(2n) + [v(q′) − v(q)]/(2t) and assigned mass

$$D(q') = 1/n + [\, v(q') - v(q) \,] / t \,. \quad (A.1)$$

The deviator earns Π(q′) = β·D(q′) − κ(q′) − F with β = (1 − Λ)b. Then Π″(q′) = βv″(q′)/t − κ″(q′) < 0, since v is concave and κ strictly convex, so Π is strictly concave and its

stationary point is the unique maximizer. Imposing symmetry $q' = q = q*$ gives the first-order condition $\beta v'(q*)/t = \kappa'(q*)$; the map $q \mapsto \kappa'(q) - \beta v'(q)/t$ is continuous, negative at $q = 0$ ($\kappa'(0) = 0$, $v' > 0$) and positive for large q ($\kappa' \to \infty$), and strictly increasing, so an interior root exists and is unique. Free entry drives profit to zero at the symmetric configuration, where each organization's assigned mass is $1/n$: $\beta/n - \kappa(q*) - F = 0$ gives $n* = \beta/[\kappa(q*) + F]$, the usual large-market approximation treating $n*$ as continuous. For the induced Φ: because quality is an experience attribute, a single evaluation draws a uniformly random organization; by rotational symmetry and Assumption 5 the arc-distance from a citizen's ideal point to a uniformly random organization is uniform on $[0, 1/2]$; with linear mismatch the sampled designation value $U = v(q*) - t \cdot d$ is an affine decreasing image of a uniform variate, hence uniform on $[v(q*) - t/2, v(q*)]$ with density $2/t > 0$. For a general strictly increasing mismatch Δ, the arc-distance is still uniform and Φ inherits a continuous positive density, with Δ' replacing t in the first-order condition; linearity adds the closed form, nothing else.

Proof of Proposition 6. Existence, uniqueness, and (9) are established above. Under the fraction rule (10), L lies strictly inside $[u_{min}, u_{max}]$ for every realized $(q*, t)$, so Assumption 1 holds; since Φ is uniform, $\Phi(L) = \lambda$, and Assumption 3 holds with density $2/t$. For Assumption 2, the expected acceptable surplus of one draw is $E[(U - L)+] = (u_{max} - L)^2/t = (1 - \lambda)^2 t/4$, so Assumption 2 reads $\bar{c} \le (1 - \delta)(1 - \lambda)^2 t/4$. With Assumptions 1–3 in force and Φ continuous with bounded support, the hypotheses of Lemmas 1–2 and Propositions 1–5 are met for the endogenous Φ. Monotonicity of the location and spread is immediate from (8)–(9). Recursive determination under (10) follows from Step 1 of the proof of Proposition 7 below: every $p(c; T)$ is a function of (c, t, δ, λ) alone, so Λ is independent of the supply side, and the configuration follows recursively—Λ, then $\beta = (1 - \Lambda)b$, then $(q*, n*)$. Under a fixed nominal penalty the determination is a fixed point in (β, Λ), treated in Lemma A.1. ∎

Lemma A.1 (Determination of the configuration). *Maintain the hypotheses of Proposition 6. (i) Under the fraction rule (10), $\Lambda_w(T)$ and $\Lambda(T; \Phi)$ under any fixed deadline depend on $(T, \bar{c}, \delta, t, \lambda)$ alone and are independent of the effective margin β; the equilibrium is recursive. (ii) Under a fixed nominal penalty $L \in (u_{min}, u_{max})$, the lapsing rate and the configuration are jointly determined by a fixed point of the map $\Lambda \mapsto \Lambda(T; \Phi(\beta(\Lambda)))$ on $[0,1]$. A fixed point always exists, and it is unique whenever $(2b/t^2) \cdot \sup_q[v'(q)]^2 / \inf_q \kappa''(q) < 1$—in particular whenever the effective margin b is small enough.*

Proof. (i) By the computation in the proof of Proposition 7, the stationary reservation for the uniform family satisfies $u_{max} - r = \sqrt{(tc/(1 - \delta))}$, so the acceptance probability $1 - \Phi(r) = \sqrt{(4c/((1 - \delta)t))}$ and the search-failure probability $\Phi(L) = \lambda$ are independent of the location u_{max}; hence every $p(c; T)$, and with it $\Lambda(T; \Phi)$ and $\Lambda_w(T)$, is independent of u_{max} and so of $q*$ and β. (ii) Under a fixed nominal L, $\Phi(L) = (L - u_{max})(2/t) + 1$ varies with u_{max}. The composite map $T(\Lambda) := \Lambda(T; \Phi(\beta(\Lambda)))$ is continuous on $[0,1]$, so a fixed point exists by Brouwer. For uniqueness, bound the slope by the chain rule: $d\beta/d\Lambda = -b$; $\partial\Phi(L)/\partial u_{max} = -2/t$; differentiating the quality condition, $du_{max}/d\beta = [v'(q*)]^2/[t\kappa''(q*) - \beta v''(q*)] \le \sup_q[v'(q)]^2/(t \cdot \inf_q \kappa''(q))$; and a fixed nominal L affects lapsing only through the final period, in which a still-searching citizen fails with probability $\Phi(L)$, so |

$\partial\Lambda/\partial\Phi(L)| \le 1 - \delta \le 1$. Multiplying, $|T'(\Lambda)| \le (2b/t^2)\cdot \sup_q[v'(q)]^2/\inf_q \kappa''(q)$; when this is below one the map is a contraction and the fixed point is unique by Banach. ∎

Proof of Proposition 7. Step 1 (the lapsing objects depend on $(\bar{c}, \delta, t, \lambda)$ only). By Proposition 6 the endogenous Φ is uniform with spread $t/2$ and, under (10), $\Phi(L) = \lambda$. The one-shot rate is $\Lambda S = \delta + (1 - \delta)\lambda$, a function of (δ, λ) alone. For the uniform family the stationary reservation of a type-c citizen solves $c = (1 - \delta)(u_{max} - r)^2/t$, so $u_{max} - r = \sqrt{(tc/(1 - \delta))}$ and the acceptance probability $1 - \Phi(r) = \sqrt{(4c/((1 - \delta)t))}$, which depends on (c, t, δ) but not on the location u_{max}; likewise $\Phi(L) = \lambda$ and every $\Phi(r_s(c))$ depend on relative positions within the support only. Thus every $p(c; T)$, and hence $\Lambda_w(T)$, depends on $(\bar{c}, \delta, t, \lambda)$ only. For the long-horizon value, the stationary indifference at $c = 0$ is $\delta(r - L) = (1 - \delta) \int$ from r to u_{max} of $(u - r)\, d\Phi(u)$; its left side is strictly positive at $r = u_{max}$ while its right side vanishes, so $r_\infty(0) < u_{max}$ and $\Lambda_D < 1$ for every configuration. With the support normalized to $[0,1]$ (so $L = \lambda$) and $a = 1 - \delta$, the closed form is $r_\infty(0) = 1/a - \sqrt{(1/a^2 - 1 - 2\delta\lambda/a)}$, whence $\Lambda_D = \delta/[\delta + a(1 - r_\infty(0))]$ depends on (δ, λ) alone; at $(\delta, \lambda) = (0.15, 0.35)$ it equals 0.3457. Step 2 (redistribution and enforcement). The ceiling $B(R) = |L|(1 - R/(\alpha\tau Y))$ depends on Φ only through $|L|$, which under (10) is fixed once $(\lambda, t, v(q^*))$ are, and the binding marginal rate $-|L|/(\alpha\tau Y)$ likewise; the threshold ρ^* of Proposition 5 depends on Φ only through the lapsing rate Λ. None of these depends on the location $u_{max} = v(q^*)$, which is the sole channel through which (κ, F, v) enter, via (8). Hence every robust characterization is invariant to (κ, F, v) given the coarse summaries (t, λ). ∎

Lemma A.2 (A sufficient condition for Condition T). *In the linear-Salop benchmark, parametrize the admissible class E by a scalar $\eta \in [\eta_0, \eta_1]$ such that the efficient quality $q_e(\eta)$ and the competitive quality $q^*(m; \eta)$ are continuous and strictly increasing in η, with $q^*(m; \eta)$ strictly increasing in m and, for each η, ranging over an interval that contains $q_e(\eta)$ as m varies. Let $\bar{m}$ match at the midpoint, $q^*(\bar{m}; \eta_m) = q_e(\eta_m)$ with $\eta_m = (\eta_0 + \eta_1)/2$, and suppose the gap $g(\eta) = q_e(\eta) - q^*(\bar{m}; \eta)$ is nondecreasing in η. Then Condition T holds—so the margin cap strictly dominates every quality target—whenever the competitive quality is at least half as responsive as the efficient quality on each side of the midpoint: $q^*(\bar{m}; \eta_1) - q^*(\bar{m}; \eta_m) \ge \tfrac{1}{2}[q_e(\eta_1) - q_e(\eta_m)]$ and $q^*(\bar{m}; \eta_m) - q^*(\bar{m}; \eta_0) \ge \tfrac{1}{2}[q_e(\eta_m) - q_e(\eta_0)]$.*

Proof of Proposition 8 and Lemma A.2. For the target: write $Q = \{q_e : e \in E\}$, an interval of width $\Delta q > 0$. A binding target induces the constant quality q_{min} in every environment, so its worst-case misallocation is $\max\{q_{min} - \inf Q, \sup Q - q_{min}\} \ge \Delta q/2$, with equality only at the midpoint; no target does better. For the cap, under Condition T the best margin cap has worst-case misallocation at most $A < \Delta q/2$, which is the dominance claim. For Lemma A.2: since g is nondecreasing with $g(\eta_m) = 0$, the cap's worst-case misallocation is $A = \max\{-g(\eta_0), g(\eta_1)\}$. By the midpoint match, $g(\eta_1) = [q_e(\eta_1) - q_e(\eta_m)] - [q^*(\bar{m}; \eta_1) - q^*(\bar{m}; \eta_m)]$, and the first responsiveness inequality makes the subtracted term at least half the first, so $g(\eta_1) \le \tfrac{1}{2}[q_e(\eta_1) - q_e(\eta_m)] < \tfrac{1}{2}\Delta q$; symmetrically $-g(\eta_0) \le \tfrac{1}{2}[q_e(\eta_m) - q_e(\eta_0)] < \tfrac{1}{2}\Delta q$. Hence $A < \Delta q/2$, which is Condition T with this A. The inequalities are satisfiable in the benchmark: $q^*(m; \eta)$ solves $m\alpha\tau Y \cdot v'(q)/t = \kappa'(q)$ and is strictly increasing in m and in any index η that raises the value of quality or lowers its cost, while $q_e(\eta)$ is nondecreasing in that index by Topkis;

calibrating $\bar{m}$ at the midpoint and choosing the parametrization so that the competitive quality is not too flat delivers the two inequalities. ∎

A.5 Proofs for Section 10

Proof of Proposition 9. Feasibility of ψ gives $\psi(R_t) \le R_t = P_t\Lambda_t$, so designated-margin outlays satisfy $(1-\Lambda_t)P_t + \psi(R_t) \le (1-\Lambda_t)P_t + \Lambda_t P_t = P_t$, which is (11), with strict inequality when $\Lambda_t > 0$ and $\psi(R_t) < R_t$. Claims expire at the year's close and cannot be pledged against future years (Section 3), so no liability on the designated margin survives any year and the margin's contribution to the debt stock is zero along every path. The government's flow budget constraint then attributes the whole primary deficit to the residual budget: $D_t = E_t - [(1-\alpha)\tau Y + (R_t - \psi(R_t))]$. No property of G, of behaviour, or of the government's objective was used. ∎

Proof of Proposition 10. (i) For any type c and any deadline, lapsing occurs if attrition strikes in the first period, an event of probability δ that no strategy avoids: a citizen holding only $x = L$ has nothing above L to direct to (Assumption 1), so she cannot direct before her first draw resolves, and the hazard applies. Hence $p(c; TR) \ge \delta$ for every c, and with a continuum of citizens $\Lambda_t = \int p(c; TR)\, dG_t(c) \ge \delta$ for every G_t. Under (12) the pool is multiplied by $1 - \Lambda_t \le 1 - \delta < 1$ each year, so $P_t \le (1-\delta)^{(t-1)} P_1 \to 0$. (ii) By Lemmas 1 and 2, $\Lambda(TR; G) \le \max_{c \in [0, \bar{c}]} p(c; TR) = \Lambda_w(TR)$ for every $G \in \mathcal{G}$, so the excess $\Lambda_t - \Lambda_w(TR)$ in (13) is never positive: the multiplier equals one in every year and the pool is constant. (iii) A type for which Assumption 2 fails strictly prefers never to evaluate; holding $x = L$ forever, she never directs (Assumption 1) and lapses with probability one, as recorded in the remark on Assumption 2 in Section 5.4. For the participating types, $p(\cdot; TR)$ is nonincreasing by Lemma 2, so each lapses with probability at least $p_{min} = p(\bar{c}; TR)$. Hence $\Lambda_t \ge \nu \cdot 1 + (1-\nu)p_{min}$ for every admissible configuration of the remaining types. The adjusted rule contracts exactly when $\Lambda_t > \Lambda_w(TR)$; the displayed lower bound exceeds $\Lambda_w(TR)$ if and only if $\nu(1 - p_{min}) > \Lambda_w(TR) - p_{min}$, that is, $\nu > \bar{\nu} = (\Lambda_w(TR) - p_{min})/(1 - p_{min})$, which lies in $[0, 1)$ since $p_{min} \le \Lambda_w(TR) = p(0; TR) < 1$. Conversely, for $\nu \le \bar{\nu}$ the configuration placing the whole participating mass at $\bar{c}$ yields $\Lambda_t = \nu + (1-\nu)p_{min} \le \Lambda_w(TR)$, and the rule does not contract. ∎

References

Alesina, Alberto, and Guido Tabellini. 1990. "A Positive Theory of Fiscal Deficits and Government Debt." *Review of Economic Studies* 57 (3): 403–414.

Amador, Manuel, and Kyle Bagwell. 2013. "The Theory of Optimal Delegation with an Application to Tariff Caps." *Econometrica* 81 (4): 1541–1599.

Anderson, Simon P., and Régis Renault. 1999. "Pricing, Product Diversity, and Search Costs: A Bertrand–Chamberlin–Diamond Model." *RAND Journal of Economics* 30 (4): 719–735.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy* 87 (5): 940–971.

Bergemann, Dirk, and Stephen Morris. 2005. “Robust Mechanism Design.” *Econometrica* 73 (6): 1771–1813.

Bohn, Henning. 1998. “The Behavior of U.S. Public Debt and Deficits.” *Quarterly Journal of Economics* 113 (3): 949–963.

Brennan, Geoffrey, and James M. Buchanan. 1980. *The Power to Tax: Analytical Foundations of a Fiscal Constitution.* Cambridge: Cambridge University Press.

Brennan, Geoffrey, and James M. Buchanan. 1985. *The Reason of Rules: Constitutional Political Economy.* Cambridge: Cambridge University Press.

Buchanan, James M. 1963. “The Economics of Earmarked Taxes.” *Journal of Political Economy* 71 (5): 457–469.

Buchanan, James M., and Roger D. Congleton. 1998. *Politics by Principle, Not Interest: Towards Nondiscriminatory Democracy.* Cambridge: Cambridge University Press.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy.* Ann Arbor: University of Michigan Press.

Buchanan, James M., and Richard E. Wagner. 1977. *Democracy in Deficit: The Political Legacy of Lord Keynes.* New York: Academic Press.

Bullain, Nilda. 2004. “Explaining Percentage Philanthropy: Legal Nature, Rationales, Impacts.” *International Journal of Not-for-Profit Law* 6 (4).

Bunch, Beverly S. 1991. “The Effect of Constitutional Debt Limits on State Governments’ Use of Public Authorities.” *Public Choice* 68 (1–3): 57–69.

Cabannes, Yves. 2004. “Participatory Budgeting: A Significant Contribution to Participatory Democracy.” *Environment and Urbanization* 16 (1): 27–46.

Carroll, Gabriel. 2015. “Robustness and Linear Contracts.” *American Economic Review* 105 (2): 536–563.

Carroll, Gabriel. 2019. “Robustness in Mechanism Design and Contracting.” *Annual Review of Economics* 11: 139–166.

Carroll, Gabriel, and Delong Meng. 2016. “Robust Contracting with Additive Noise.” *Journal of Economic Theory* 166: 586–604.

Chassang, Sylvain. 2013. “Calibrated Incentive Contracts.” *Econometrica* 81 (5): 1935–1971.

Congressional Budget Office. 2026a. *The Budget and Economic Outlook: 2026 to 2036.* Washington, DC: Congressional Budget Office.

Congressional Budget Office. 2026b. *The Long-Term Budget Outlook Data: 2026 to 2056.* Washington, DC: Congressional Budget Office.

Deller, Steven, Craig Maher, Lindsay Amiel, and Judith I. Stallmann. 2013. “The Restrictiveness of State Tax and Expenditure Limitations and State Debt.” *Public Finance and Management* 13 (1): 1–25.

Dworczak, Piotr, and Alessandro Pavan. 2022. “Preparing for the Worst but Hoping for the Best: Robust (Bayesian) Persuasion.” *Econometrica* 90 (5): 2017–2047.

Frankel, Alexander. 2014. “Aligned Delegation.” *American Economic Review* 104 (1): 66–83.

Frey, Bruno S., and Reiner Eichenberger. 1996. “FOCJ: Competitive Governments for Europe.” *International Review of Law and Economics* 16 (3): 315–327.

Frey, Bruno S., and Reiner Eichenberger. 1999. *The New Democratic Federalism for Europe: Functional, Overlapping and Competing Jurisdictions.* Cheltenham: Edward Elgar.

Friedman, Milton, and Rose D. Friedman. 1980. *Free to Choose: A Personal Statement.* New York: Harcourt Brace Jovanovich.

Garrett, Daniel F., and Alessandro Pavan. 2015. “Dynamic Managerial Compensation: A Variational Approach.” *Journal of Economic Theory* 159: 775–818.

Halac, Marina, and Pierre Yared. 2014. “Fiscal Rules and Discretion under Persistent Shocks.” *Econometrica* 82 (5): 1557–1614.

Harsanyi, John C. 1955. “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility.” *Journal of Political Economy* 63 (4): 309–321.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States.* Cambridge, MA: Harvard University Press.

Hurwicz, Leonid. 1972. “On Informationally Decentralized Systems.” In *Decision and Organization*, edited by C. B. McGuire and Roy Radner, 297–336. Amsterdam: North-Holland.

Kiewiet, D. Roderick, and Kristin Szakaly. 1996. “Constitutional Limitations on Borrowing: An Analysis of State Bonded Indebtedness.” *Journal of Law, Economics, and Organization* 12 (1): 62–97.

Knight, Brian G. 2000. “Supermajority Voting Requirements for Tax Increases: Evidence from the States.” *Journal of Public Economics* 76 (1): 41–67.

Liebman, Jeffrey B., and Neale Mahoney. 2017. “Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement.” *American Economic Review* 107 (11): 3510–3549.

Lippman, Steven A., and John J. McCall. 1976. “The Economics of Job Search: A Survey.” *Economic Inquiry* 14 (2): 155–189.

Oates, Wallace E. 1972. *Fiscal Federalism.* New York: Harcourt Brace Jovanovich.

Ostrom, Vincent, Charles M. Tiebout, and Robert Warren. 1961. “The Organization of Government in Metropolitan Areas: A Theoretical Inquiry.” *American Political Science Review* 55 (4): 831–842.

Persson, Torsten, and Lars E. O. Svensson. 1989. “Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences.” *Quarterly Journal of Economics* 104 (2): 325–345.

Poterba, James M. 1994. “State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics.” *Journal of Political Economy* 102 (4): 799–821.

Poterba, James M. 1996. “Budget Institutions and Fiscal Policy in the U.S. States.” *American Economic Review* 86 (2): 395–400.

Rawls, John. 1971. *A Theory of Justice.* Cambridge, MA: Belknap Press of Harvard University Press.

Salop, Steven C. 1979. “Monopolistic Competition with Outside Goods.” *Bell Journal of Economics* 10 (1): 141–156.

Shadbegian, Ronald J. 1999. “The Effect of Tax and Expenditure Limitations on the Revenue Structure of Local Government, 1962–87.” *National Tax Journal* 52 (2): 221–238.

Tiebout, Charles M. 1956. “A Pure Theory of Local Expenditures.” *Journal of Political Economy* 64 (5): 416–424.

Török, Marianna, and David Moss, eds. 2004. *Percentage Philanthropy.* Budapest: Nonprofit Information and Training Centre.

von Hagen, Jürgen. 1991. “A Note on the Empirical Effectiveness of Formal Fiscal Restraints.” *Journal of Public Economics* 44 (2): 199–210.

Wicksell, Knut. 1896. “Ein neues Prinzip der gerechten Besteuerung.” In *Finanztheoretische Untersuchungen.* Jena: Gustav Fischer. English translation: “A New Principle of Just Taxation,” in *Classics in the Theory of Public Finance*, edited by Richard A. Musgrave and Alan T. Peacock. London: Macmillan, 1958.

Wilson, Robert. 1987. “Game-Theoretic Analyses of Trading Processes.” In *Advances in Economic Theory: Fifth World Congress*, edited by Truman Bewley, 33–70. Cambridge: Cambridge University Press.

Wolinsky, Asher. 1986. “True Monopolistic Competition as a Result of Imperfect Information.” *Quarterly Journal of Economics* 101 (3): 493–511.

Yang, Kai Hao. 2022. “Selling Consumer Data for Profit: Optimal Market-Segmentation Design and Its Consequences.” *American Economic Review* 112 (4): 1364–1393.

Yared, Pierre. 2019. “Rising Government Debt: Causes and Solutions.” *Journal of Economic Perspectives* 33 (2): 115–140.