

October 2025

“Government Reputation in Ramsey Taxation”

Georgy Lukyanov and Emin Ablyatifov

Government Reputation in Ramsey Taxation*

Georgy Lukyanov[†]

Emin Ablyatifov[‡]

Abstract

We embed honesty-based reputation into a Ramsey taxation framework with competitive firms and households. In a static benchmark with exogenous trust, there is a sharp cutoff below which the optimal policy sets no taxes and above which the optimal tax take rises with trust. In the dynamic model, beliefs evolve through noisy public monitoring of delivered public goods; the planner's problem is well posed, the value is increasing and convex in beliefs, and optimal revenue is monotone in reputation with a trust threshold that is weakly below the static cutoff. With multiple broad instruments and symmetric monitoring, the dynamic force acts through the total revenue scale; the tax mix is indeterminate along an equivalence frontier. Blackwell-improving monitoring and greater type persistence expand the optimal scale and shift the trust threshold inward. The model delivers clear policy prescriptions for building fiscal capacity in low-trust environments and testable links between measured trust, verifiability, and revenue.

Keywords: Optimal taxation; Government reputation; Ramsey problem; Credibility; Fiscal capacity.

JEL: H21; H30; E62; D82; C73.

1 Introduction

Recent experience in Russia during the pandemic provides a salient backdrop. Large fiscal packages were announced in an environment of fragile institutional trust, alongside highly publicized episodes of both successful and failed delivery. In such settings, public spending has both an allocative and an informational role: it provides public goods and simultaneously

*We thank Arthur Izgarshev, David Li, Anastasia Makhmudova, Konstantin Shamruk, Yuliia Tukmakova, Anna Vlasova and Stanislav Vyushkov for valuable comments; all remaining errors are our own.

[†]Toulouse School of Economics. Email: georgy.lukyanov@tse-fr.eu

[‡]Independent researcher, Brentwood, United Kingdom. Email: eminablyatifov@gmail.com

reveals information about the government’s honesty. Credible implementation raises beliefs; shortfalls depress them. This paper asks how optimal taxation should be designed when reputation is an endogenous state variable determined in equilibrium by observed delivery and noisy public monitoring.

We ask three questions. First, how should the *scale* of taxation respond to reputation? Second, what is the *composition* of taxes when several broad instruments are available? Third, how do monitoring, partial enforcement, and economic structure shift these prescriptions? We study a static benchmark—where trust is an exogenous index—and a dynamic environment in which reputation follows a Markov process and evolves through Bayesian updating from a public signal whose informativeness increases with delivered public goods.

This paper places a Ramsey taxation environment inside a reputational equilibrium with noisy monitoring and delivers three core advances. First, it shows that *reputation disciplines fiscal capacity primarily through the tax scale*. In the static benchmark there is a sharp trust threshold for taxing (Proposition 1). In the dynamic model, existence holds and the value is increasing and convex in beliefs (Theorem 6; Proposition 7); the informational term tightens the incentive to raise revenue so the no-tax region moves inward (Proposition 4). This yields a simple, testable policy rule: the optimal revenue scale is (weakly) increasing in reputation and adjusts in a step-like manner after news (Corollary 1). Second, with multiple broad instruments and symmetric monitoring, the problem *separates*: the dynamic force acts through revenue scale, while the instrument mix is indeterminate along a static equivalence frontier (Propositions 2 and 8). Third, the framework isolates when composition matters again—precisely when informativeness is instrument-specific or when incidence concerns break equivalence—providing a clear bridge to empirics and policy (Section 8).

Relative to classical public finance, the paper endogenizes the social return to public funds via Bayesian reputation while retaining the production-efficiency discipline of Diamond and Mirrlees (1971a,b) and the tax-mix logic of Atkinson and Stiglitz (1976). Unlike standard Ramsey analyses (e.g., Ramsey, 1927; Chari et al., 1994), the planner’s objective now embeds an informational continuation value that depends on delivered revenue; this is the lever behind the dynamic trust threshold (Proposition 4) and the scale monotonicity. Relative to the credibility and time-inconsistency tradition (Kydland and Prescott, 1977; Chari and Kehoe, 1990), the government’s payoff-relevant attribute is *honesty in delivery*, and noisy public monitoring ties current fiscal effort to future beliefs. Closest are reputation models in policy settings such as Phelan (2006) and the honesty results of Fudenberg et al. (2022), as well as Lu (2013). Our contribution is to embed honesty-based reputation into a fully specified static and dynamic *Ramsey* model with competitive firms and households, to characterize the dynamic trust cutoff and convexity in beliefs (Theorem 6; Proposition 7), and

to establish scale–mix separation with clean failure conditions (Proposition 8; Remark 2).

2 Related Literature

This paper connects three literatures: optimal taxation in public finance, credibility and time inconsistency in policy design, and formal models of reputation—with a particular focus on honesty—applied to fiscal policy.

First, we relate to the classical Ramsey program and production efficiency results in second-best taxation. The benchmark for efficient tax design dates back to Ramsey (1927) and the production-efficiency theorems of Diamond and Mirrlees (1971a,b), with the tax-mix insights of Atkinson and Stiglitz (1976). Our static and dynamic Ramsey environments adopt the standard general equilibrium logic while departing along a single dimension: the government’s concern for maintaining a reputation for honesty (modeled as a belief about type) that feeds back into equilibrium behavior and thus into the planner’s implementability set. We also speak to the long-run capital taxation debate inaugurated by Chamley (1986) and Judd (1985), and to environments where market incompleteness overturns zero-capital-tax conclusions (Aiyagari, 1995). Our dynamic results on the scale and mix of taxes under reputation constraints are therefore naturally compared to the unconstrained Ramsey allocations in Chari et al. (1994) and the capital-tax benchmarks above.

Second, our analysis builds on credibility and time inconsistency in policy. The seminal insight that optimal discretionary policies are dynamically inconsistent (Kydland and Prescott, 1977) motivated a large literature on rules, discretion, and reputational enforcement (Barro and Gordon, 1983; Chari and Kehoe, 1990). We adapt that logic to fiscal policy by treating “honesty” as the payoff-relevant attribute that sustains credible tax choices and delivery of public goods. In contrast to inflation games, our government’s announcements and realized tax schedules interact with equilibrium labor supply and firm behavior in general equilibrium, so that reputational concerns shift not only expectations but also resource allocations and the government budget constraint.

Third, we connect directly to reputation models in policy environments. In particular, Phelan (2006) models public trust and government betrayal when types are unobservable and compliance is costly; Lu (2013) develops optimal policy with credibility concerns in a reputation framework; and Fudenberg et al. (2022) formalize conditions under which a long-run player can build a *reputation for honesty* and thereby approximate commitment outcomes. We leverage these insights in a Ramsey context: our government’s honesty is defined through the verifiable delivery of public spending and tax schedules, and beliefs update via Bayes’ rule from observed fiscal outcomes. Relative to this literature, our contribution is

to embed honesty-based reputation into a fully specified static and dynamic optimal taxation model with firms and households, derive comparative statics of the optimal *tax scale* and *tax mix* with respect to reputation, and characterize when reputational constraints implement production efficiency versus force tax distortions.

A complementary strand studies fiscal rules as devices to restore credibility. Halac and Yared (2014) characterize optimal discretion in the presence of persistent shocks; Halac and Yared (2018) analyze coordinated rules in a world economy; and Halac and Yared (2022) study rules under limited enforcement. We share their emphasis on institutional mechanisms disciplining policy, but differ in primitive enforcement: in our setting, reputational concerns (shaped by Bayesian beliefs) rather than externally enforced rules constrain the planner, yielding testable predictions linking measured public trust to the optimal tax scale and tax composition.

Methodologically, our comparative statics in beliefs and signal informativeness build on classical orderings of information structures and monotone methods. The partial order over monitoring technologies uses Blackwell’s comparison of experiments (Blackwell, 1953). Monotone comparative statics and lattice methods underpin our results on the shape of the value function and the optimal policy’s dependence on beliefs (Milgrom and Shannon, 1994; Topkis, 1998; Athey, 2002). These tools allow us to deliver clean threshold and single-crossing characterizations of the optimal tax scale and mix when reputation varies.

Finally, this paper is part of a broader research agenda on credibility, public communication, and policy design. Our brief companion paper, Lukyanov and Ablyatifov (2025), focuses on a compact static benchmark that maps measured trust into the optimal tax scale/mix and offers a policy-ready calibration; the present article provides the general equilibrium foundations and a dynamic Ramsey treatment with endogenous reputation. Together, they aim to bridge theory and empirics on how imperfect trust reshapes optimal taxation.

Roadmap Section 3 describes the environment. Section 4 develops the static benchmark with exogenous trust and characterizes the trust threshold and the equivalence frontier. Section 5 introduces the dynamic model with Markov reputation and noisy monitoring, and defines equilibrium. Section 6 establishes existence, convexity in beliefs, the monotone policy rule, the dynamic threshold, and the scale–mix separation. Section 7 presents the quantitative illustration. Section 8 examines robustness and policy-relevant extensions. Section 9 distills policy implications. Section 10 concludes. The appendices collect proofs.

3 Environment

We consider a competitive economy with a representative household, a representative firm, and a government that can raise distortionary taxes. Time is discrete. This section lays out the static primitives used both for the one-period benchmark and as the per-period environment in the dynamic model.

3.1 Agents, Technology, and Markets

The representative household has period utility

$$u(C, G, L) = \tilde{u}(C, L) + G,$$

where $\tilde{u} : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is C^1 , strictly increasing in C , strictly decreasing in L , and strictly concave. The linear term in G captures that delivered public consumption (if any) enters utility one-for-one.

A competitive firm produces the private good according to $Y = f(L)$, where $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is C^1 , strictly increasing, strictly concave, and satisfies $f(0) = 0$. The private good is the numeraire. The representative household owns the firm and receives profits π .

The government chooses a vector of ad valorem tax rates $\tau = (\tau_L, \tau_B) \in [0, 1]^2$, where τ_L is a labor-income tax and τ_B is a broad-base tax levied at source on the firm's revenue (e.g., an output or value-added tax).¹ In the baseline we assume no instrument-specific costs: the planner's objective and feasibility set depend on instruments only through the net-of-tax product and the revenue delivered.

The government has a per-period type $g \in \{H, O\}$. If $g = H$ (honest), delivered public consumption equals collected revenue $G = R(\tau)$; if $g = O$ (opportunistic), $G = 0$. In the static benchmark, the probability that $g = H$ is an exogenous trust parameter $\theta \in (0, 1)$; in the dynamic model, types evolve as a Markov chain and beliefs update via Bayes' rule (Section 5).

3.2 Timing and Information

In a period, the sequence is: (i) beliefs about honesty θ are public; (ii) the government sets taxes τ ; (iii) private agents choose (C, L) and the firm chooses L given τ ; (iv) nature draws g and delivers G according to type; (v) in the dynamic model a public signal is realized and beliefs update (Section 5). In the static benchmark, only the prior θ matters.

¹In Section 8.4 we interpret the broad base as capital taxation. The analysis below is written to cover either interpretation.

3.3 Competitive Equilibrium for Given Taxes

Given τ and prices, the household chooses (C, L) to maximize $u(C, G, L)$ subject to

$$C \leq (1 - \tau_L) w L + \pi,$$

taking G as parametric (determined by the government's type) and where w is the wage and π are profits rebated from the firm.

Given τ_B , the firm solves

$$\max_{L \geq 0} (1 - \tau_B)Y - wL = (1 - \tau_B)f(L) - wL.$$

First-order conditions give

$$w = (1 - \tau_B) f'(L). \quad (1)$$

The private good is the numeraire and the household owns the firm; hence feasibility is $C + G = Y$. Government revenue is

$$R(\tau) = \tau_L w L + \tau_B Y, \quad (2)$$

with w given by (1). Delivered public consumption equals $G = R(\tau)$ if $g = H$ and $G = 0$ if $g = O$.

Assumption 1. $\tilde{u}$ is C^1 , strictly increasing in C , strictly decreasing in L , and strictly concave; f is C^1 , strictly increasing, strictly concave, and $f(0) = 0$; $a(\tau) = a_L(\tau_L) + a_B(\tau_B)$ with $a_i(0) = a'_i(0) = 0$ and $a''_i(\cdot) > 0$.

Lemma 1. *Under Assumption 1, for any $\tau \in [0, 1]^2$ there exists a unique competitive allocation $(C(\tau), L(\tau), Y(\tau))$. Moreover, private choices depend on taxes only through the composite net-of-tax factor*

$$s(\tau) \equiv (1 - \tau_L)(1 - \tau_B) \in (0, 1],$$

in the sense that $L(\tau) = \mathcal{L}(s(\tau))$ for a continuous, strictly increasing function $\mathcal{L}$ characterized by the MRS condition

$$-\frac{\partial \tilde{u}(C, L)}{\partial L} \bigg/ \frac{\partial \tilde{u}(C, L)}{\partial C} = s(\tau) f'(L), \quad (3)$$

with $C = Y = f(L)$. Consequently $Y(\tau) = \mathcal{Y}(s(\tau))$ and, writing $W(s) \equiv f'(\mathcal{L}(s)) \mathcal{L}(s)$,

revenue can be expressed as

$$R(\tau) = (1 - \tau_B) \tau_L W(s(\tau)) + \tau_B \mathcal{Y}(s(\tau)). \quad (4)$$

4 Static Benchmark

This section analyzes the one-period problem when trust in honesty is an exogenous primitive $\theta \in [0, 1]$. The government chooses a labor-income tax τ_L and a broad-base ad valorem tax τ_B (commodity or output) taking as given households' and firms' competitive responses. Public spending equals collected revenue if the government is honest and zero otherwise. Expected welfare is the representative household's utility with the public component weighted by θ .

4.1 Primitives and private equilibrium

Given $\tau := (\tau_L, \tau_B)$, the firm solves $\max_L (1 - \tau_B)f(L) - wL$, yielding $w = (1 - \tau_B)f'(L)$. A small atomistic household takes w as parametric and chooses L to solve

$$\max_{L \geq 0} U(C) - V(L) \quad \text{s.t.} \quad C \leq (1 - \tau_L)wL + \pi(\tau),$$

where $\pi(\tau)$ are profits rebated lump-sum. The intratemporal first-order condition is

$$\frac{V'(L)}{U'(C)} = (1 - \tau_L)w = (1 - \tau_L)(1 - \tau_B)f'(L). \quad (5)$$

Define the *net-of-tax product*

$$S(\tau) \equiv (1 - \tau_L)(1 - \tau_B) \in (0, 1],$$

so that (5) becomes $V'(L)/U'(C) = S(\tau)f'(L)$. The competitive allocation $(C(\tau), L(\tau), Y(\tau))$ depends on τ only through $S(\tau)$; the induced output is $Y(\tau) = f(L(S))$. Government revenue is

$$R(\tau) = \tau_L w L + \tau_B Y = \tau_L(1 - \tau_B)f'(L)L + \tau_B f(L), \quad (6)$$

and the goods resource constraint is $C + G = Y$.

Lemma 2. *For any (τ_L, τ_B) and (τ'_L, τ'_B) with $S(\tau) = S(\tau')$, the private allocation coincides: $L(\tau) = L(\tau')$ and $Y(\tau) = Y(\tau')$. Consequently, distortions at the labor-leisure margin are summarized by the scalar $S \in (0, 1]$.*

The aggregator in Lemma 2 reduces the static problem to two scalars: the private wedge S and delivered revenue R . With this reduction in hand, the next step is to pin down when the planner uses the revenue instrument at all. The following proposition shows that optimal taxation exhibits a trust cutoff: below it, the planner sets $R = 0$ and relies exclusively on private allocation efficiency; above it, revenue is strictly positive and increases with trust.

4.2 Planner's problem and welfare representation

An honest government converts revenue one-for-one into public consumption while an opportunistic government delivers none. Expected welfare at τ is

$$W(\tau; \theta) = U(C(\tau)) - V(L(\tau)) + \theta G(\tau), \quad (7)$$

with $G(\tau) = R(\tau)$ and $C(\tau) = Y(\tau) - G(\tau)$. Using $C + G = Y$, (7) can be written as

$$W(\tau; \theta) = U(Y(\tau) - R(\tau)) - V(L(\tau)) + \theta R(\tau). \quad (8)$$

The Ramsey planner chooses τ (equivalently, a pair (S, τ_B) or (S, τ_L)) to maximize $W(\tau; \theta)$ subject to (5) and (6).

Proposition 1. *Let (C^0, L^0, Y^0) denote the zero-tax allocation, characterized by $V'(L^0) = U'(Y^0)f'(L^0)$. Define the trust threshold $\bar{\theta} \equiv U'(Y^0)$. Then: (i) if $\theta \leq \bar{\theta}$, the optimal policy sets $\tau_L = \tau_B = 0$; (ii) if $\theta > \bar{\theta}$, any optimal policy raises positive revenue.*

Having characterized the scale decision, we turn to the composition of taxation. Because competitive allocations depend on the net-of-tax product S and the revenue requirement R , multiple instrument pairs can implement the same allocation. The next result makes this indifference precise by describing the static equivalence frontier for the mix.

A small distortionary tax changes welfare at the zero-tax allocation by

$$\left. \frac{dW}{d\tau} \right|_{\tau=0} = -U'(Y^0) \left. \frac{dR}{d\tau} \right|_{\tau=0} + \theta \left. \frac{dR}{d\tau} \right|_{\tau=0},$$

because the envelope condition removes private-substitution terms at (C^0, L^0) . The sign is determined by the wedge $\theta - U'(Y^0)$, delivering a clean cutoff.

Proposition 2. *Suppose $\theta > \bar{\theta}$ and no instrument-specific costs ($a \equiv 0$). Then any interior optimum satisfies the scale rule*

$$U'(C^*) = \theta, \quad C^* = Y^* - G^*, \quad (9)$$

so the optimal private consumption C^* depends only on θ . Let S^* be the net-of-tax product that supports $Y^* = f(L(S^*))$ consistent with C^* . The set of optimal tax mixes is a non-degenerate equivalence frontier:

$$\mathcal{F}(\theta) \equiv \left\{ (\tau_L, \tau_B) \in [0, 1]^2 : S(\tau) = S^*, R(\tau) = G^* \right\}.$$

All policies in $\mathcal{F}(\theta)$ implement the same allocation (C^*, L^*, Y^*) and achieve the same welfare $W^*(\theta)$.

The equivalence frontier formalizes a simple point: absent instrument-specific frictions or incidence concerns, the planner is indifferent across mixes that deliver the same (S, R) . The remark below translates this into practical implications and clarifies how the frontier collapses once we depart from symmetric instruments.

Remark 1. In the baseline with no instrument-specific costs, the planner is indifferent over the entire equivalence frontier: any mix that attains S^* and R^* implements (C^*, L^*, Y^*) and achieves the same value. Selection of a unique mix requires considerations beyond the model (administration, salience, or enforcement), which we omit here.

The scale rule (9) is the static sufficient-statistics analogue of a Samuelson condition under imperfect trust: marginal utility of private consumption is equated to the marginal expected value of public funds. Given C^* , feasibility pins down $G^* = Y^* - C^*$. Distortions at the labor–leisure margin are governed by S^* (Lemma 2); the remaining degree of freedom selects the mix (τ_L, τ_B) that delivers exactly G^* along that private allocation. With two broad instruments, this indifference yields a continuum of optimal mixes.

If $U(C) = \ln C$ and $V(L) = \frac{1}{2}L^2$, the threshold becomes $\bar{\theta} = 1/Y^0$, the scale rule implies $C^*(\theta) = 1/\theta$, delivered public spending is $G^*(\theta) = Y^* - 1/\theta$, and $W^*(\theta) = -\ln \theta + Y^*\theta - \text{const}$. These closed forms are convenient for calibration-ready plots; the equivalence frontier is easily visualized as the locus of (τ_L, τ_B) pairs holding $S(\tau)$ fixed while meeting $R(\tau) = G^*(\theta)$.

All proofs are provided in A. The next section embeds this static benchmark into an infinite-horizon environment with endogenous reputation, where belief dynamics discipline the intertemporal scale of taxation while instrument costs select the mix period by period.

5 Dynamic Model with Markov Reputation

We study an infinite-horizon version of the environment where the government’s honesty is uncertain and evolves stochastically. Time is discrete, $t = 0, 1, 2, \dots$

The government has a type $g_t \in \{H, O\}$ each period. Type H (honest) transforms revenue one-for-one into public consumption; type O (opportunistic) diverts revenue so that delivered public consumption is zero. Types evolve according to a two-state Markov chain with transition matrix

$$\Pi = \begin{bmatrix} \pi_{HH} & 1 - \pi_{HH} \\ 1 - \pi_{OO} & \pi_{OO} \end{bmatrix}, \quad \pi_{HH}, \pi_{OO} \in (0, 1).$$

Let $\theta_t \equiv \Pr(g_t = H \mid h_t)$ denote the public belief (reputation) at the start of period t , conditioned on public history h_t .

Preferences and technology are as in the static benchmark: the representative household has $u(C_t, G_t, L_t) = \tilde{u}(C_t, L_t) + G_t$, with $\tilde{u}$ strictly concave, and the competitive firm produces $Y_t = f(L_t)$ with $f' > 0 > f''$. The government chooses a tax vector $\tau_t = (\tau_{L,t}, \tau_{B,t})$ on labor income and a broad base B (e.g., output or capital). Given τ_t , private choices pin down a unique competitive allocation $(C_t(\tau_t), L_t(\tau_t), Y_t(\tau_t))$ and revenue $R(\tau_t)$; we allow small convex instrument costs $a(\tau_t) = a_L(\tau_{L,t}) + a_B(\tau_{B,t})$ with $a_i(0) = a'_i(0) = 0$.²

If $g_t = H$, delivered public consumption equals $G_t = R(\tau_t)$; if $g_t = O$, $G_t = 0$. Society does not observe g_t but does observe a public signal $s_t \in \{0, 1\}$ whose likelihood depends on the delivered G_t through primitives $q_H(\cdot)$ and $q_O(\cdot)$:

$$\Pr(s_t = 1 \mid g_t = H, \tau_t) = q_H(R(\tau_t)), \quad \Pr(s_t = 1 \mid g_t = O, \tau_t) = q_O(R(\tau_t)),$$

with $q_i : [0, \bar{R}] \rightarrow (0, 1)$ continuous. We impose an MLRP-type condition: the likelihood ratio $q_H(R)/(1 - q_H(R))$ divided by $q_O(R)/(1 - q_O(R))$ is strictly increasing in R . Intuitively, higher delivered public consumption makes a favorable signal more indicative of honesty. The perfect-monitoring case is the limit $q_H \equiv 1, q_O \equiv 0$.

At the start of t , reputation θ_t is public. The government chooses τ_t . Households and firms take τ_t as given and implement (C_t, L_t, Y_t) . Nature draws g_t using θ_t as the prior and delivers G_t according to type. The public signal s_t is realized and publicly observed. Beliefs are updated to $\hat{\theta}_t \equiv \Pr(g_t = H \mid h_t, s_t, \tau_t)$, and the Markov transition maps $\hat{\theta}_t$ into next period's prior:

$$\theta_{t+1} = \Phi(\hat{\theta}_t) \equiv \pi_{HH} \hat{\theta}_t + (1 - \pi_{OO})(1 - \hat{\theta}_t). \quad (10)$$

Let $\ell_H(R) \equiv q_H(R)$ and $\ell_O(R) \equiv q_O(R)$. Given (θ_t, τ_t) , the probability of a favorable

²As in the static note, this nests the product form for private distortions and a linear public-value aggregator. Nothing in the dynamic statements below requires the isoelastic specialization, which we reserve for the quantitative section.

signal is

$$p_1(\theta_t, \tau_t) = \theta_t \ell_H(R(\tau_t)) + (1 - \theta_t) \ell_O(R(\tau_t)).$$

Posteriors after observing s_t are

$$\hat{\theta}_t(1) = \frac{\theta_t \ell_H(R(\tau_t))}{p_1(\theta_t, \tau_t)}, \quad \hat{\theta}_t(0) = \frac{\theta_t [1 - \ell_H(R(\tau_t))]}{1 - p_1(\theta_t, \tau_t)}.$$

Combining with (10) yields $\theta_{t+1} = \Theta(\theta_t, \tau_t, s_t)$ with two possible next-period values $\theta_{t+1}^{(1)}$ and $\theta_{t+1}^{(0)}$.

A benevolent government maximizes the representative household's expected discounted welfare:

$$\max_{\{\tau_t\}} \mathbb{E}_0 \sum_{t \geq 0} \beta^t \left\{ \tilde{u}(C_t(\tau_t), L_t(\tau_t)) + \mathbf{1}\{g_t = H\} R(\tau_t) - a(\tau_t) \right\}, \quad \beta \in (0, 1).$$

Reputation matters because current revenue $R(\tau_t)$ changes the informativeness of s_t via $q_i(\cdot)$, which shifts θ_{t+1} through Bayes and the Markov transition, altering future welfare.

5.1 Government's Recursive Problem and Equilibrium Concept

The payoff-relevant public state is the reputation $\theta \in [0, 1]$. Let $w(\tau)$ denote current expected welfare given τ and belief θ :

$$w(\tau; \theta) \equiv \tilde{u}(C(\tau), L(\tau)) + \theta R(\tau) - a(\tau),$$

where $C(\tau), L(\tau)$ are the competitive responses. Given (θ, τ) , the probability of $s = 1$ is $p_1(\theta, \tau)$, and the two possible next beliefs are $\theta^{(1)}(\theta, \tau)$ and $\theta^{(0)}(\theta, \tau)$ as above. The Bellman equation is

$$V(\theta) = \max_{\tau \in \mathcal{T}} \left\{ w(\tau; \theta) + \beta [p_1(\theta, \tau) V(\theta^{(1)}(\theta, \tau)) + (1 - p_1(\theta, \tau)) V(\theta^{(0)}(\theta, \tau))] \right\}. \quad (11)$$

Here $\mathcal{T} \subset [0, 1]^2$ is a compact set of tax rates (or instruments), and $a(\cdot)$ is continuous and strictly convex componentwise.

Definition 1. A Markov Perfect Equilibrium consists of (i) a value function $V : [0, 1] \rightarrow \mathbb{R}$, (ii) a measurable policy $\tau^* : [0, 1] \rightarrow \mathcal{T}$, and (iii) belief updates given by Bayes' rule and (10), such that (11) holds for all θ and private allocations $(C(\tau), L(\tau), Y(\tau))$ are competitive equilibrium outcomes at $\tau^*(\theta)$.

Assumption 2. $\mathcal{T}$ is compact; a is C^1 and strictly convex; $\tilde{u}$ is continuous, strictly concave, and bounded above; $R(\cdot)$ and $C(\cdot), L(\cdot)$ are continuous; q_H, q_O are continuous and satisfy MLRP.

Theorem 1. *Under Assumption 2, the Bellman operator*

$$(TV)(\theta) = \max_{(S,R) \in \mathcal{F}_0} \left\{ U(Y(S) - R) - V(L(S)) + \theta R + \Gamma_V(\theta; R) \right\}$$

with

$$\Gamma_V(\theta; R) = \beta \mathbb{E}[V(\theta^+) \mid \theta, R]$$

is a contraction on the space of bounded continuous functions $(\mathcal{C}(\Theta), \|\cdot\|_\infty)$ with modulus β . It has a unique fixed point $V^* \in \mathcal{C}(\Theta)$. The argmax correspondence is nonempty, compact-valued, and upper hemicontinuous in θ ; a measurable selector $\theta \mapsto (S^*(\theta), R^*(\theta))$ exists.

The operator is monotone and a β -contraction under the sup norm. Continuity follows from the maximum theorem because the objective is continuous and the instrument set is compact. A measurable selection then delivers a Markov policy. *Proof:* B, Lemma 3 and Proposition 18.

With existence and well-posedness established, the key question is the shape of the value in beliefs. The next proposition shows that the value is increasing and convex in θ , a property that underlies all subsequent monotone policy results.

6 Characterization of Optimal Policy

This section records three target results we will prove under Assumption 2. They deliver existence, monotone comparative statics in reputation, and the separation between *scale* (revenue) and *mix* (instrument composition).

Proposition 3. *Let Assumption 2 hold. Then the operator T preserves monotonicity and convexity in θ . Consequently, the unique fixed point V^* is increasing and convex on $[0, 1]$.*

Convexity and monotonicity generate a single-crossing structure in (θ, R) . Evaluating the zero-revenue margin at low beliefs then delivers a dynamic analogue of the static cutoff, which is weakly lower because delivered revenue carries informational value.

Higher reputation raises the expected current payoff one-for-one with revenue and also improves the distribution of next period's beliefs in the convex order, which benefits a convex value function. *Proof:* B, Lemmas 4–5 and Proposition 19.

Corollary 1. *Let $\tau^*(\theta)$ be an equilibrium policy and $R^*(\theta) \equiv R(\tau^*(\theta))$. Then $R^*(\theta)$ is (weakly) increasing in θ . Moreover, for any belief θ and realized signal $s \in \{0, 1\}$, the implied next prior $\theta^{(s)}$ satisfies*

$$R^*(\theta^{(0)}) \leq R^*(\theta) \leq R^*(\theta^{(1)}),$$

with strict inequalities whenever the signal is informative and interior solutions obtain.

Having characterized the scale, we return to the instrument mix. Under symmetric monitoring and the static feasibility set from Lemma 2, the dynamic force operates entirely through R ; composition remains a static indifference. The next proposition states this separation formally.

With MLRP and a convex value, the continuation payoff has increasing differences in beliefs and the revenue scale, so optimal revenue rises in reputation. Good news steps the scale up; bad news steps it down. *Proof:* B, Corollary 2.

Proposition 4. *There exists $\bar{\theta}^{\text{dyn}} \in [0, 1]$ such that $R^*(\theta) = 0$ for all $\theta \leq \bar{\theta}^{\text{dyn}}$ and $R^*(\theta) > 0$ for $\theta > \bar{\theta}^{\text{dyn}}$. Moreover, $R^*(\theta)$ is (weakly) increasing in θ . If the experiment is informative at $R = 0$ (the likelihood ratio is non-constant in R at 0), then $\bar{\theta}^{\text{dyn}} < \bar{\theta}^{\text{stat}}$.*

The dynamic trust threshold implies a step-like revenue schedule: small belief movements near the cutoff trigger discrete policy adjustments. The corollary below formalizes this history dependence via belief updates after favorable or unfavorable signals.

A small tax has the usual static trade-off and an informational gain: by increasing delivered public consumption it makes the public signal more informative, which is valuable when the value function is convex in beliefs. This informational term weakly lowers the no-tax region relative to the static benchmark. *Proof:* B, Assumption A1, Lemma 6, and Proposition 20.

Proposition 5. *Under Assumption 2 and instrument-symmetric monitoring (the kernel $Q(\cdot|g, R)$ depends on revenue R but not on the tax mix conditional on R), the Bellman maximization is equivalent to:*

$$\max_{(S, R) \in \mathcal{F}_0} U(Y(S) - R) - V(L(S)) + \theta R + \Gamma_{V^*}(\theta; R),$$

followed by implementation with any $(\tau_L, \tau_B) \in \mathcal{T}$ satisfying $(1 - \tau_L)(1 - \tau_B) = S$ and $R(\tau_L, \tau_B) = R$. The continuation term depends on the scale R but is independent of the mix. Hence, along the static equivalence frontier the mix is indeterminate in the baseline.

Separation is a knife-edge property: it hinges on symmetry in monitoring and incidence. The following remark delineates precisely when composition becomes first order again.

Remark 2. If Q depends on the instrument conditional on R , then Γ_{V^*} depends on the mix, and separation breaks: the optimal policy tilts toward the more informative base.

Private allocations depend on instruments only through S , while monitoring depends on delivered revenue R . Hence the dynamic problem selects the scale; without instrument costs the mix remains a free degree of freedom along the static frontier.

Proposition 6. *Under Assumption 2, the Bellman equation (11) admits a solution V that is bounded and continuous on $[0, 1]$. There exists a measurable selector $\tau^*(\theta) \in \arg \max_{\tau \in \mathcal{T}} \{\cdot\}$, hence an MPE exists.*

Proposition 7. *Suppose in addition that $R(\cdot)$ is single-crossing in each instrument and q_H, q_O satisfy MLRP. Then $V(\theta)$ is increasing and convex in θ . The optimal revenue scale $R^*(\theta) \equiv R(\tau^*(\theta))$ is (weakly) increasing in θ . There exists a threshold $\bar{\theta} \in (0, 1)$ such that for $\theta \leq \bar{\theta}$, no taxation is optimal ($\tau^*(\theta) = 0$), while for $\theta > \bar{\theta}$ the optimal policy raises positive revenue. After an unfavorable signal $s = 0$, the induced drop from θ to $\theta^{(0)}(\theta, \tau^*(\theta))$ strictly lowers next period's optimal revenue, $R^*(\theta^{(0)}(\cdot)) < R^*(\theta)$, and conversely for $s = 1$.*

Proposition 8. *Let $a(\tau) = a_L(\tau_L) + a_B(\tau_B)$ with a_i strictly convex and small. For any fixed $\theta > \bar{\theta}$, the optimization in (11) separates into (i) choosing a revenue scale R that maximizes*

$$\tilde{W}(R; \theta) \equiv \max_{\tau \in \mathcal{T}: R(\tau)=R} \left\{ \tilde{u}(C(\tau), L(\tau)) + \theta R - a(\tau) \right\} + \beta \mathbb{E}[V(\theta') \mid \theta, R],$$

and (ii) selecting the cheapest instrument mix that attains R along the static equivalence frontier. Consequently, the dynamic problem pins down a unique mix via $a(\cdot)$ and pins down the scale via the marginal value of reputation, which depends only on (θ, R) and the signal technology q_H, q_O .

Remark 3. MLRP implies that larger delivered public consumption makes good news more informative about honesty. The convexity of V captures the dynamic marginal value of reputation: an extra unit of trust at higher θ is worth more because it sustains a larger tax scale with lower future distortion. Proposition 7 delivers a simple history dependence: after bad news, scale steps down; after good news, it steps up. Proposition 8 says the dynamic channel acts through the *scale*, while the *mix* is chosen statically each period by instrument cost—consistent with the static frontier logic.

For transparency, one can adopt $f(L) = aL^\beta$, $\tilde{u}(C, L) = \ln C - \frac{1}{2}L^2$, $q_H(R) = \rho + (1 - \rho)\sigma(R)$, $q_O(R) = (1 - \kappa)\sigma(R)$ where $\sigma(R)$ is strictly increasing and $\rho, \kappa \in [0, 1]$ govern false

positives/negatives. With Π diagonal ($\pi_{HH} = \pi_{OO} = \pi$), the belief mapping collapses to a one-dimensional stochastic kernel $\theta \mapsto \{\theta^{(1)}(\theta, R), \theta^{(0)}(\theta, R)\}$ with weights $p_1(\theta, R)$, allowing closed-form comparative statics for $R^*(\theta)$ and explicit policy functions for figures.

The characterization establishes four pillars announced in Section 1. First, a well-posed dynamic Ramsey problem with a unique value exists (Theorem 6). Second, the value is increasing and convex in beliefs, which underpins the single-crossing and monotone policy shape (Proposition 7). Third, the revenue scale exhibits a dynamic trust threshold—weakly below the static cutoff when monitoring is informative at zero—and rises (weakly) with reputation (Proposition 4). Fourth, with symmetric monitoring the problem separates: reputation acts through the *scale* R , while the instrument mix is indeterminate along the static equivalence frontier (Propositions 2 and 8); Remark 2 pinpoints when composition matters again. These results are the backbone for the quantitative section and for the extensions on monitoring, partial enforcement, and heterogeneity.

7 Quantitative Illustration

This section provides a compact quantitative illustration of the static benchmark and its dynamic implications. We adopt transparent primitives and keep the exercise deliberately stylized to highlight the mechanisms. Households have utility $u(C, G, L) = \ln C - \frac{1}{2}L^2 + \theta G$, firms produce $Y = 2\sqrt{L}$ under perfect competition, and the government uses a labor tax τ_L and a broad commodity (output) tax τ_Y . Given (τ_L, τ_Y) , the competitive equilibrium satisfies $L^2 = \frac{1-\tau_L}{2-\tau_L}$, $C = (1 - \tau_Y)(2 - \tau_L)\sqrt{L}$, and delivered public consumption is $G = \sqrt{L}[\tau_L(1 - \tau_Y) + 2\tau_Y]$. Welfare is $W = \ln C - \frac{1}{2}L^2 + \theta G$.

We set θ on a grid from 0.2 to 0.95 and compute the optimal policy on a dense grid over $(\tau_L, \tau_Y) \in [0, 0.9]^2$. For the dynamic signal used in the belief map, we set $\pi_{HH} = \pi_{OO} = 0.9$ and use a binary public signal with $q_H(R) = 0.2 + 0.8 \frac{R}{R+1}$ and $q_O(R) = 0.1 \frac{R}{R+1}$, which satisfies monotone likelihood ratio and links information content to delivered revenue.

Figure 1 shows the optimal revenue scale $R^*(\theta)$. For low trust, the optimal scale is zero; beyond a cutoff, $R^*(\theta)$ rises with θ .

Figure 2 reports the optimal tax rates $\tau_Y^*(\theta)$ and $\tau_L^*(\theta)$ without instrument costs. Even in the costless case, the revenue requirement and the net-of-tax product pin down a unique mix at the optimum for each θ ; the broad base typically bears most of the burden in this parametric example.

Using the policy-induced scale, Figure 3 plots the next-period prior after good and bad news. Good news maps beliefs above the 45° line and bad news maps them below, delivering the step-up/step-down history dependence predicted by the theory.

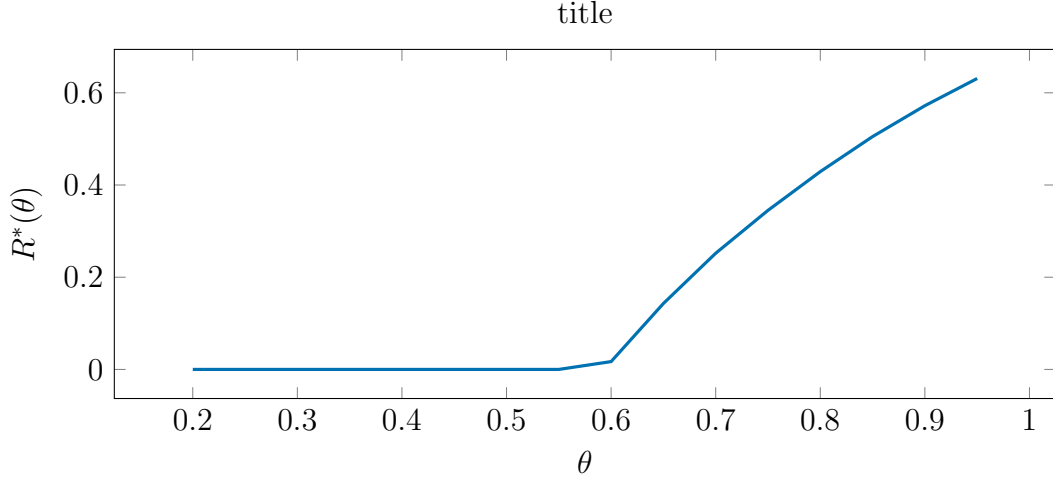


Figure 1: Optimal revenue scale $R^*(\theta)$ in the baseline (no instrument-specific costs; symmetric monitoring). The scale is (weakly) increasing in reputation and exhibits a zero-tax region at low beliefs (Proposition 4).

Figure 4 shows the equilibrium welfare; the function is increasing and convex in trust, mirroring the value convexity result.

All curves use a colorblind-friendly palette with distinct line styles and remain interpretable in grayscale print. Parameter values are listed in each caption.

8 Extensions and Robustness

This section studies four extensions that speak directly to external validity and policy design. We show that the main comparative statics—monotonicity of the optimal revenue scale in reputation and the existence of a trust threshold—survive noisy monitoring; partial commitment shifts the scale but weakens the informational motive; the scale–mix separation persists for broad instruments unless the signal technology is instrument–specific; and heterogeneity breaks the static equivalence frontier through incidence, while preserving the central belief–scale logic when signals depend only on aggregate delivery.

8.1 Robustness of the Core Results

This subsection records four robustness results: minimal primitives for our static aggregator and feasibility, signal assumptions beyond MLRP, the role of type persistence, and a worked-out heterogeneous-agent case.

Our static aggregator $(S, R) \mapsto (C, L, Y)$ does not rely on the log–square example. The following abstracts from parametric forms.

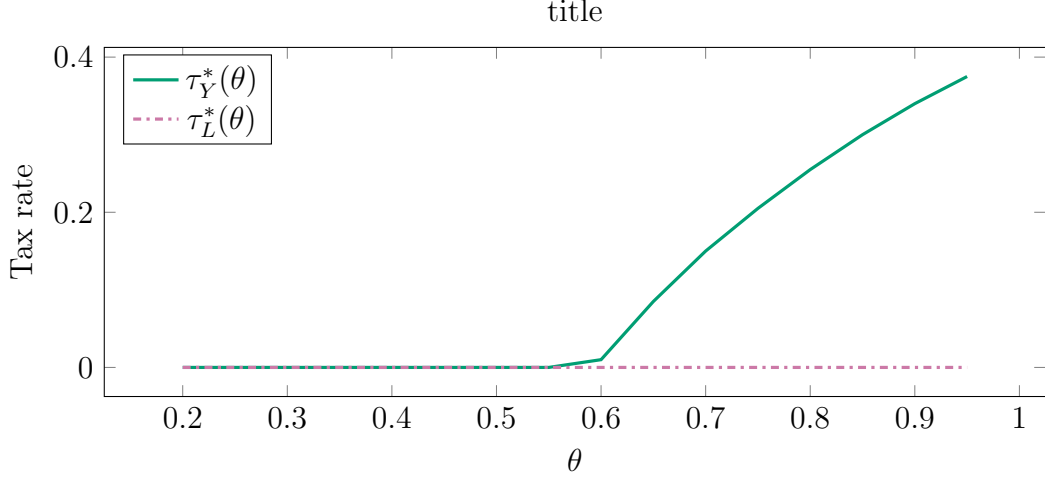


Figure 2: Equivalence frontier for the tax mix in the baseline. For each θ , any mix attaining the same $(S, R^*(\theta))$ implements the same allocation and value (Propositions 2 and 8). The plot displays a representative optimal path for $(\tau_Y^*(\theta), \tau_L^*(\theta))$.

Proposition 9. *Suppose (i) preferences are $U(C) - V(L)$ with U strictly increasing, strictly concave and V convex, continuous; (ii) firms are competitive with a strictly concave technology in labor that yields a unique competitive allocation for any constant net-of-tax product $S \in (0, 1]$; (iii) instruments map continuously from a compact set $\mathcal{T}$ into (S, R) . Then there exist continuous functions $Y(S)$ and $L(S)$ with Y strictly increasing and concave, L weakly decreasing, such that for each (S, R) the competitive allocation is $C = Y(S) - R$, $L = L(S)$. The feasibility image $\mathcal{F}_0$ is nonempty and compact.*

Signals beyond MLRP Our monotonicity in beliefs and the dynamic cutoff use MLRP in R to obtain single crossing in (θ, R) . For comparative statics in the *informativeness* of monitoring, MLRP is not required, it is enough that monitoring becomes more informative in the Blackwell sense. The next proposition records the corresponding comparative statics.

Proposition 10. *Let experiments be indexed by $\iota \in I$ and assume that for every R , $Q_{\iota'}(\cdot | g, R)$ Blackwell-dominates $Q_{\iota}(\cdot | g, R)$ whenever $\iota' \succ \iota$. Then for all θ , $R_{\iota'}^*(\theta) \geq R_{\iota}^*(\theta)$ and the dynamic trust threshold satisfies $\bar{\theta}^{\text{dyn}}(\iota') \leq \bar{\theta}^{\text{dyn}}(\iota)$, with strict inequalities on sets of positive measure when solutions are interior. This comparative static does not require MLRP.*

A simple thresholded-visibility example makes the logic concrete: even when MLRP may fail near the threshold, a Blackwell improvement still pushes the optimal scale up and the trust cutoff inward.

Example 1. Let the public signal be $s = \mathbf{1}\{G \geq \kappa\}$ observed with symmetric bit-flip noise $\varepsilon \in (0, \frac{1}{2})$; i.e., $\Pr(s \text{ observed correctly}) = 1 - \varepsilon$. Lower ε Blackwell-dominates higher ε at

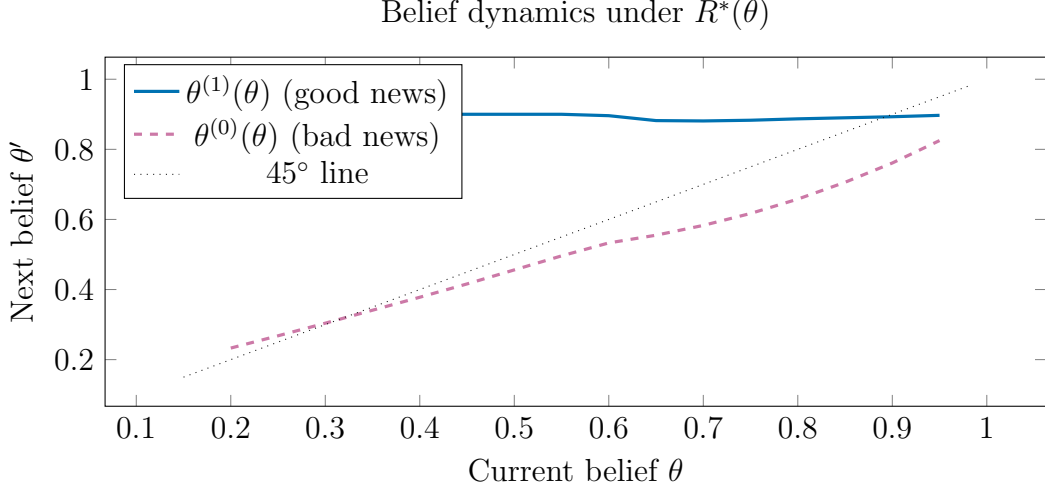


Figure 3: Belief dynamics under the baseline policy $R^*(\theta)$. Favorable (unfavorable) signals shift next-period beliefs above (below) the 45° line, inducing step-up (step-down) adjustments in the scale (Corollary 1).

each R ; hence $R^*(\theta)$ increases and $\bar{\theta}^{\text{dyn}}$ falls as $\varepsilon \downarrow 0$, even though MLRP in R may fail globally around the threshold κ .

Information is more valuable when types are persistent. Strengthening persistence amplifies the continuation value of delivered revenue and therefore expands the optimal scale, as the next result shows.

Proposition 11. *Order transition matrices by persistence: $\Pi' \succeq \Pi$ if $\pi'_{HH} \geq \pi_{HH}$ and $\pi'_{OO} \geq \pi_{OO}$. Under Assumption 2, for every belief θ we have $R^*_{\Pi'}(\theta) \geq R^*_{\Pi}(\theta)$ and $\bar{\theta}^{\text{dyn}}(\Pi') \leq \bar{\theta}^{\text{dyn}}(\Pi)$, with strict inequalities on sets of positive measure at interior solutions.*

Finally, allowing heterogeneous workers preserves the scale results but breaks mix indifference through incidence. The following proposition separates what survives from what becomes pinned by distributional weights.

Proposition 12. *Consider two worker types $i \in \{1, 2\}$ with utilities $U(C_i) - V_i(L_i)$, V_i convex, w common, and a utilitarian social aggregator with weights $(\lambda, 1 - \lambda)$. Suppose the signal depends only on aggregate delivered revenue. Then (i) $R^*(\theta)$ remains (weakly) increasing in θ and a dynamic trust threshold exists; (ii) for given (S, R) , the social marginal deadweight loss of revenue via labor vs. commodity taxation differs unless $V_1 \equiv V_2$, so the instrument mix is generically unique and determined by incidence.*

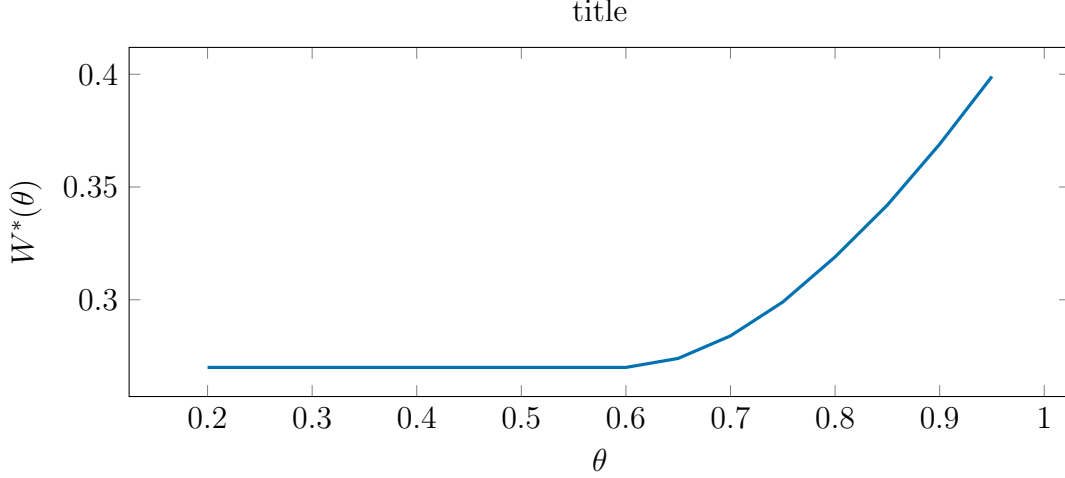


Figure 4: Welfare $W^*(\theta)$ under the baseline policy. The convexity of the value in beliefs (Proposition 7) and the monotone scale map into a smooth welfare profile, with faster gains where monitoring is more informative.

8.2 Noisy Monitoring and Verifiable Delivery

Let the public signal satisfy the MLRP assumptions in the baseline and introduce an *informativeness index* $\iota \in [0, 1]$ that scales the likelihood ratio. Formally, consider a one-parameter family of experiments $\{(q_H^\iota, q_O^\iota)\}_{\iota \in [0, 1]}$ with

$$\log \frac{q_H^\iota(R)}{1 - q_H^\iota(R)} - \log \frac{q_O^\iota(R)}{1 - q_O^\iota(R)} = \iota \cdot \left(\log \frac{q_H(R)}{1 - q_H(R)} - \log \frac{q_O(R)}{1 - q_O(R)} \right),$$

so that $\iota' = 1$ Blackwell-dominates ι whenever $\iota' > \iota$. The perfect-monitoring limit is $\iota \downarrow 0$ if informativeness collapses (signals independent of type) or $\iota \uparrow 1$ if one starts from a minimally informative baseline.

Proposition 13. *Fix primitives and the instrument set. If $\iota' > \iota$ in the Blackwell sense, then for every belief θ the optimal revenue scale satisfies $R_{\iota'}^*(\theta) \geq R_\iota^*(\theta)$, with strict inequality on a set of beliefs of positive measure whenever interior solutions obtain. Moreover, the dynamic no-tax region shrinks: $\bar{\theta}^{\text{dyn}}(\iota') \leq \bar{\theta}^{\text{dyn}}(\iota)$.*

More informative monitoring raises the continuation value of an additional unit of revenue because beliefs move in the convex order in a way that favors a convex value function. This pushes up the optimal scale and lowers the reputation cutoff at which it becomes optimal to tax.

A salient special case is *verifiable delivery*. Suppose an observable fraction $\lambda \in [0, 1]$ of delivered public consumption is publicly verified (e.g., by third-party audit), while the

remainder is subject to the baseline noisy signal. Then the posterior is a mixture of a perfectly revealing component and the original experiment, which Blackwell–dominates the baseline. Proposition 13 implies $R^*(\theta)$ weakly increases in λ and $\bar{\theta}^{\text{dyn}}$ weakly falls.

Proposition 14. *Under perfect monitoring (signals reveal the current type independently of R), the informational term in the dynamic Euler equation is zero. The dynamic cutoff coincides with the static one, $\bar{\theta}^{\text{dyn}} = \bar{\theta}^{\text{stat}}$, and the policy function $R^*(\theta)$ reduces to the static benchmark.*

When current type is publicly observed, changing revenue does not change the informativeness of the signal; reputation evolves only through the exogenous Markov transition. The dynamic problem collapses to a sequence of static ones.

8.3 Partial Commitment and Enforcement Devices

When instruments target different bases, composition matters only if the informational or incidence content differs. This subsection clarifies those cases and the empirical footprints they leave.

Partial commitment and enforcement add a parallel discipline to reputation. We model them as either verifiable fractions of delivery or penalties for shortfalls; both operate like outward shifts of the feasibility set or the monitoring kernel.

Suppose legal or institutional devices force an opportunistic government to deliver a fraction $\varphi \in [0, 1]$ of collected revenue, while an honest government still delivers all of it. Delivered public consumption is

$$G_t = \mathbf{1}\{g_t = H\} R_t + \mathbf{1}\{g_t = O\} \varphi R_t,$$

so expected contemporaneous benefits from revenue equal $(\theta_t + (1 - \theta_t)\varphi)R_t$. When the enforced fraction φ is publicly verifiable, the signal about G_t becomes *less* informative about type because delivery partially ceases to be type–contingent.

Proposition 15. *For any belief θ , partial commitment raises the contemporaneous marginal value of revenue from θ to $\theta + (1 - \theta)\varphi$, so the static scale rule becomes $U'(C^*) = \theta + (1 - \theta)\varphi$. At the same time, if enforcement is publicly verifiable, the informational marginal value of revenue weakens relative to the baseline. Consequently, the effect of φ on the dynamic trust threshold $\bar{\theta}^{\text{dyn}}$ is the sum of (i) a direct reduction via the higher contemporaneous return and (ii) an indirect increase via the loss of informational leverage. In the limit $\varphi \rightarrow 1$, types are observationally equivalent and the problem collapses to the static Ramsey benchmark with effective weight one on G .*

Earmarking, escrow, or outside verification acts like insurance against opportunism: it lifts the expected payoff from revenue today but erodes the reputational benefits of using revenue to generate news. The net effect on the no-tax region depends on which force dominates, a testable implication across institutional environments.

8.4 Capital vs. Commodity Taxation

Extend the environment to two periods (or infinite horizon) with savings. Households choose (C_t, L_t, S_t) with $K_{t+1} = S_t$ and firms produce $Y_t = f(L_t, K_t)$; the government can levy a consumption/commodity tax $\tau_{C,t}$ and/or a capital income tax $\tau_{K,t}$, with total revenue $R_t = R(\tau_{C,t}, \tau_{K,t})$. Assume the public signal depends on delivered revenue, not on the instrument per se.

Proposition 16. *Under standard separability and representative-agent assumptions, there is an intertemporal equivalence frontier: sequences $\{\tau_{C,t}, \tau_{K,t}\}$ that implement the same intertemporal wedges deliver the same private allocation for a given revenue path $\{R_t\}$. If the signal technology depends on R_t only, the dynamic reputation problem selects the revenue scale path $\{R_t^*\}$, while instrument choice along the frontier remains indeterminate in the baseline. If, however, one instrument yields a strictly more informative experiment about honesty per unit of revenue in the Blackwell sense, the frontier tilts toward that instrument and the scale–mix separation breaks.*

The reputation channel operates through the *scale*: more revenue today produces news about honesty tomorrow. If monitoring treats dollars from different bases symmetrically, the dynamic problem cannot distinguish among revenue sources and the instrument mix is left to second-order criteria. If, in contrast, some bases are more *auditable* (e.g., consumption taxes with public receipts) and thus more informative, the planner has a reason to tilt the mix toward those bases even absent instrument-specific frictions. A corollary is that transitory capital taxation can be optimal as a reputational investment when near-term revenue is especially informative.

8.5 Heterogeneity and Distributional Considerations

Heterogeneity restores a unique mix by weighting deadweight losses across types. The revenue scale remains monotone in belief, but the planner now chooses composition to minimize distribution-weighted distortions.

Now let households be heterogeneous in productivity or tastes and let the government evaluate welfare with a strictly concave social aggregator. With linear instruments, the static

equivalence frontier generally breaks because different tax mixes imply different incidence, even for the same net-of-tax product S and revenue R .

Proposition 17. *Suppose the signal depends only on aggregate delivered revenue. Then, for any fixed social aggregator, the optimal revenue scale $R^*(\theta)$ remains (weakly) increasing in reputation and a dynamic trust threshold exists. The instrument mix is, however, determined by distributional concerns even in the baseline and need not be indeterminate along a frontier.*

The belief dynamics are aggregate and continue to load only on R : the convexity and monotonicity arguments in beliefs go through, so the scale inherits the same shape in θ . But heterogeneity restores a first-order role for the mix because incidence differs across instruments. If monitoring is group specific (e.g., some projects generate verifiable benefits concentrated on identifiable groups), then even the reputational channel can become mix sensitive: instrument choice and targeting can be used to shape the informativeness of signals, potentially overturning the separation. This yields a sharp empirical prediction: in low-trust environments with group-specific verifiability, revenue may be shifted toward instruments or uses that *prove* honesty most visibly, even when those instruments are not incidence efficient.

9 Policy Implications

This section interprets the model’s prescriptions in environments where trust in government varies across time and jurisdictions. Two features are central. First, optimal tax *scale* rises with reputation (Proposition 7), with a clear no-tax region at low trust (Proposition 4). Second, in the baseline with broad instruments and symmetric monitoring, the *mix* of instruments is indeterminate along an equivalence frontier (Proposition 8); dynamic forces act through revenue scale, not composition. Extensions show how monitoring and enforcement move these prescriptions.

The static scale rule equates the marginal utility of private consumption to the expected marginal payoff of public funds; in the dynamic model the same intuition survives with an informational term that decreases the no-tax region. Practically, measured trust (surveys, audit outcomes, delivery records) can be mapped to a target revenue scale: low trust calls for caution (small or zero tax take), while higher trust justifies a larger tax capacity. Because $R^*(\theta)$ is increasing in θ (Proposition 7), calibrating a trust index to θ yields a monotone schedule for the budget envelope.

Belief dynamics imply step-downs after bad news and step-ups after good news (Corollary 1). Following a negative signal, the planner should temporarily scale down taxes to

limit distortion, then rebuild reputation with policy choices that generate informative good news (Section 8). This creates a rationale for *reputation stabilizers*: automatic rules that relax the tax take after credibility shocks and gradually restore it as signals improve.

The informational motive suggests that governments can expand fiscal capacity not only by changing tax rates but also by making delivery more *verifiable*. Stronger monitoring (audits, third-party verification, public dashboards) Blackwell-dominates weaker schemes and therefore increases the optimal revenue scale while lowering the no-tax cutoff (Proposition 13). In the limit of perfect monitoring, the informational motive vanishes and the dynamic cutoff collapses to the static one (Proposition 14). This points to “credibility capital” as an input into fiscal capacity.

In the baseline, instrument choice is second order: if monitoring treats revenue symmetrically, any mix that attains the optimal $R^*(\theta)$ implements the same allocation (Proposition 8). Two forces break this indifference. First, if some bases are *more informative per unit revenue* (e.g., taxes coupled to publicly observable receipts), the frontier tilts toward those bases (Proposition 16). Second, with heterogeneity, incidence makes the mix a first-order choice even when informativeness is symmetric (Proposition 17). Both forces yield a practical rule: in low-trust environments, favor revenue sources and spending uses that are easiest to verify, provided distributional costs are acceptable.

When a fraction of delivery is verifiable (earmarks with audits, escrowed spending, milestone certifications), expected contemporaneous returns to revenue rise, but the reputational leverage of revenue weakens (Proposition 15). The net effect on the tax scale depends on these opposing forces. A practical design is to front-load *high-visibility* projects early in a mandate to raise θ , then exploit the higher trust to finance less verifiable but high-value projects later.

Three predictions invite testing. (i) Jurisdictions with higher measured trust should exhibit larger tax takes and smaller no-tax regions; policy should step up (down) after good (bad) signals about delivery. (ii) Exogenous improvements in monitoring (digitized audits, public reporting mandates) should expand fiscal capacity at given fundamentals. (iii) Where monitoring is group-specific, governments should tilt revenue and spending toward verifiable bases/uses, especially after negative signals. Each prediction links observable institutional changes to shifts in $R^*(\theta)$ and belief dynamics.

The baseline abstracts from instrument-specific features and enforcement heterogeneity; both matter for incidence and compliance. Our extensions clarify how they re-enter without overturning the core scale logic: reputation shapes the *scale*, while the *mix* is disciplined by informativeness and distribution once those features are modeled explicitly.

10 Conclusion

This paper places a Ramsey taxation problem inside a reputational equilibrium. Citizens hold beliefs about whether the government is honest, monitoring is noisy but informative about delivered public goods, and beliefs evolve through Bayesian updating. The core message is simple and policy relevant: credibility disciplines fiscal capacity primarily through the *scale* of taxation, while the instrument mix is secondary in the baseline. In the static benchmark, optimal taxation features a sharp trust cutoff (Proposition 1). In the dynamic environment, we establish existence of a well-posed problem and show that the value is increasing and convex in beliefs (Theorem 6; Proposition 7). The informational continuation value tightens the incentive to raise revenue so that the no-tax region shrinks relative to the static threshold (Proposition 4), and policy adjusts in a step-like manner after news (Corollary 1). With multiple broad instruments and symmetric monitoring, the problem separates: reputation acts through *revenue scale* R , and the instrument mix is indeterminate along an equivalence frontier (Propositions 2 and 8).

The model generates clear comparative statics that map into implementable reforms. Blackwell-improving monitoring—audits, milestone certification, public dashboards—raises the optimal scale and moves the trust cutoff inward (Proposition 10); greater persistence of government type amplifies these effects (Proposition 11). In environments with heterogeneity and incidence concerns, composition re-enters and the equivalence frontier collapses in favor of mixes that minimize distribution-weighted deadweight loss (Proposition 12). Together these results parse policy into two levers: invest in credibility (more verifiable delivery) to enlarge the optimal scale, and, where incidence matters, allocate the burden across bases with an eye to information content and distributional costs.

Several predictions are testable. First, holding fundamentals fixed, measured trust should be positively related to the aggregate tax take—a cross-sectional and time-series monotonicity. Second, large, visible delivery shocks (e.g., project completion or verified transfer roll-outs) should tilt beliefs above the 45° line and induce discrete upward adjustments in the revenue scale, while scandals or failed delivery should induce symmetric downward adjustments (Corollary 1). Third, exogenous improvements in verifiability (for example, random audit expansions or the rollout of third-party certification) should raise the tax scale and lower the effective trust cutoff (Proposition 10). Fourth, when incidence differences are first order, instrument choices should covary with heterogeneity (Proposition 12). These implications suggest straightforward empirical strategies: difference-in-differences around monitoring reforms, event studies on high-salience delivery, and cross-jurisdictional designs exploiting staggered audit expansions or transparency mandates.

The quantitative illustration is deliberately spare; its role is to visualize the monotone policy rule in beliefs, the belief dynamics after signals, and the welfare gains from verifiability (Section 7). Nothing in the mechanism hinges on parametric choices: the static aggregator requires only monotone/concave technology and standard preferences; the dynamic results rely on Blackwell improvements and single crossing rather than a particular signal parametric form (Subsection 8.1). The payoff is a small set of robust prescriptions that travel well across institutional contexts.

The framework is intentionally minimal, and that is also its limitation. We use binary government types, a representative household, and symmetric monitoring across instruments in the baseline. Commitment is absent beyond reputational incentives; monitoring technology is taken as given; and political frictions are outside the model. Each simplification sharpens a mechanism but leaves scope for work that is both feasible and important. Extending the analysis to *instrument-specific* monitoring would endogenize composition even in the baseline and refine the scale–mix trade-off; incorporating *partial commitment* or enforcement would put institutional design (rules, escape clauses) on the same footing as reputation; and allowing *heterogeneous* agents with incidence objectives would connect more directly to distributional policy. A dynamic mechanism-design treatment with audit choice as an explicit instrument would be a natural next step, as would a political-economy extension with electoral accountability and multi-level government.

In sum, the paper delivers a tractable Ramsey model in which honesty-based reputation is a state variable. It provides existence and convexity (Theorem 6; Proposition 7), a dynamic trust threshold and monotone revenue scale (Proposition 4), a clean separation between scale and mix under symmetric monitoring (Propositions 2 and 8), and comparative statics for monitoring and persistence (Propositions 10–11). These elements yield a compact policy message: when trust is scarce, do less but verify more; when trust is ample, expand the scale; and when incidence bites, allocate across bases with an eye to both information and distribution. We view this as a useful template for thinking about fiscal capacity in low-trust environments and a foundation for empirical work linking measured trust, verifiability, and the size and shape of taxation.

A Proofs for Section 4

A.1 Proof of Lemma 2

Fix primitives U, V, f as in the text. For any tax pair $\tau = (\tau_L, \tau_B)$, the competitive firm's FOC is

$$w = (1 - \tau_B)f'(L).$$

The representative household takes (w, τ_L) as given and chooses L to solve

$$\max_{L \geq 0} U(C) - V(L) \quad \text{s.t.} \quad C \leq (1 - \tau_L)wL + \pi(\tau),$$

with firm profits rebated lump-sum. The intratemporal FOC and the envelope condition imply

$$\frac{V'(L)}{U'(C)} = (1 - \tau_L)w = (1 - \tau_L)(1 - \tau_B)f'(L). \quad (12)$$

Define $S(\tau) := (1 - \tau_L)(1 - \tau_B) \in (0, 1]$. Equation (12) can be written as

$$\frac{V'(L)}{U'(C)} = S(\tau)f'(L).$$

Because U is strictly concave and V strictly convex, and f is strictly concave, the competitive allocation $(C(\tau), L(\tau))$ is unique whenever a solution exists; moreover, (12) together with feasibility $C + G = Y = f(L)$ and the government budget identity pins down (C, L, Y) as a function of the single scalar $S(\tau)$. Therefore, for any τ, τ' with $S(\tau) = S(\tau')$ we have the same unique solution (C, L, Y) , and in particular $L(\tau) = L(\tau')$ and $Y(\tau) = Y(\tau')$. This proves the claim. $\square$

A.2 Proof of Proposition 1

Let (C^0, L^0, Y^0) denote the zero-tax allocation, characterized by $V'(L^0) = U'(Y^0)f'(L^0)$ and $Y^0 = f(L^0)$. Consider a small revenue-raising perturbation of taxes from zero. It suffices to vary the broad-base tax τ_B while keeping $\tau_L = 0$. For τ_B small and positive, firms' FOC gives $w = (1 - \tau_B)f'(L)$; at $\tau = 0$ this reduces to $w = f'(L^0)$. Revenue at $(\tau_L, \tau_B) = (0, \tau_B)$ is

$$R(0, \tau_B) = \tau_B f(L(\tau_B)).$$

Hence $\left. \frac{dR}{d\tau_B} \right|_{\tau=0} = f(L^0) = Y^0 > 0$.

Expected welfare is $W(\tau; \theta) = U(Y(\tau) - R(\tau)) - V(L(\tau)) + \theta R(\tau)$. By the standard envelope logic at the competitive allocation (Diamond–Mirrlees production efficiency), changes

in C and L induced by a small tax enter only through the revenue channel at the zero-tax point; formally,

$$\left. \frac{dW}{d\tau_B} \right|_{\tau=0} = \left(-U'(Y^0) + \theta \right) \left. \frac{dR}{d\tau_B} \right|_{\tau=0} = (\theta - U'(Y^0)) Y^0.$$

Define $\bar{\theta} := U'(Y^0)$. If $\theta \leq \bar{\theta}$, every sufficiently small revenue-raising perturbation weakly lowers welfare, so the origin is optimal by continuity and global concavity of W in a neighborhood of zero. If $\theta > \bar{\theta}$, a small increase in τ_B strictly increases welfare, so zero taxes cannot be optimal; hence any optimum raises positive revenue (not necessarily through τ_B alone). $\square$

A.3 Proof of Proposition 2

Fix $\theta > \bar{\theta}$ and assume $a \equiv 0$. By Lemma 2, the private allocation is indexed by $S \in (0, 1]$. For a given S , denote the associated unique private quantities by $L(S)$ and $Y(S) = f(L(S))$. Along the iso- S locus, distortions at the labor-leisure margin are held fixed, and the only remaining policy choice is how much revenue R to collect by varying the *mix* (τ_L, τ_B) subject to $S(\tau) = S$.

Holding S fixed (hence holding $Y = Y(S)$ fixed), the planner chooses $R \in [0, Y(S)]$ to maximize

$$\widetilde{W}(R; S, \theta) = U(Y(S) - R) - V(L(S)) + \theta R.$$

The function is strictly concave in R because U is strictly concave. Therefore any interior optimum satisfies the first-order condition

$$-U'(Y(S) - R^*) + \theta = 0 \iff U'(C^*) = \theta, \quad C^* := Y(S) - R^*,$$

which is exactly the scale rule stated in (9). The FOC pins down C^* as a function only of θ ; given $Y(S)$, it pins down $G^* = R^* = Y(S) - C^*$.

For a fixed $S \in (0, 1]$, parametrize the iso- S set by $\tau_B \in [0, 1]$ and define $\tau_L(\tau_B) := 1 - \frac{S}{1-\tau_B}$, which ensures $(1 - \tau_L(\tau_B))(1 - \tau_B) = S$. Because L and Y depend only on S , the revenue mapping along the iso- S locus is

$$\begin{aligned} R(\tau_B; S) &= \tau_L(\tau_B) (1 - \tau_B) f'(L(S)) L(S) + \tau_B f(L(S)) \\ &= ((1 - \tau_B) - S) f'(L(S)) L(S) + \tau_B f(L(S)) \\ &= -\tau_B f'(L(S)) L(S) + \tau_B f(L(S)) + \underbrace{(1 - S) f'(L(S)) L(S)}_{\text{constant in } \tau_B}. \end{aligned}$$

Thus $R(\tau_B; S)$ is affine in τ_B , with range

$$\mathcal{R}(S) = \left[(1-S)f'(L(S))L(S), (1-S)f'(L(S))L(S) + f(L(S)) - f'(L(S))L(S) \right].$$

Because $f(L) - f'(L)L > 0$ for strictly concave f and $L > 0$, $\mathcal{R}(S)$ is a non-degenerate interval. Consequently, for any target $R^* \in \mathcal{R}(S)$ there exists a continuum of mixes $\tau_B \in [0, 1)$ (with the implied $\tau_L(\tau_B)$) such that $R(\tau_B; S) = R^*$.

At the optimum, S is chosen so that (i) C^* satisfies $U'(C^*) = \theta$ and (ii) $R^* = Y(S) - C^*$ lies in $\mathcal{R}(S)$; the latter is satisfied for any S sufficiently close to one, and by continuity for the maximizing S^* . All mixes (τ_L, τ_B) that (a) keep $S(\tau) = S^*$ and (b) deliver $R(\tau) = R^*$ implement the same allocation (C^*, L^*, Y^*) and achieve the same value. The set of such mixes is a non-degenerate one-dimensional locus, the *equivalence frontier* $\mathcal{F}(\theta)$ defined in the proposition. $\square$

B Proofs for Section 5

Throughout this subsection, primitives satisfy Assumption 2 in the main text: $\mathcal{T}$ is compact; $a(\cdot)$ is C^1 and strictly convex componentwise with $a_i(0) = a'_i(0) = 0$; $\tilde{u}$ is continuous and strictly concave in (C, L) and bounded above; R, C, L are continuous in τ ; and q_H, q_O are continuous and satisfy an MLRP condition in the sense stated in the paper. We recall the Bellman operator

$$(TV)(\theta) \equiv \sup_{\tau \in \mathcal{T}} \left\{ w(\tau; \theta) + \beta \mathbb{E}[V(\theta') \mid \theta, \tau] \right\}, \quad w(\tau; \theta) = \tilde{u}(C(\tau), L(\tau)) + \theta R(\tau) - a(\tau),$$

where $\theta' \in \{\theta^{(1)}(\theta, \tau), \theta^{(0)}(\theta, \tau)\}$ with

$$\Pr(\theta' = \theta^{(1)}(\theta, \tau) \mid \theta, \tau) = p_1(\theta, \tau) = \theta q_H(R(\tau)) + (1 - \theta) q_O(R(\tau)),$$

and $\theta^{(s)}(\theta, \tau) = \Phi(\hat{\theta}(s; \theta, \tau))$, $\Phi(\hat{\theta}) = \pi_{HH}\hat{\theta} + (1 - \pi_{OO})(1 - \hat{\theta})$.

B.1 Existence, uniqueness and measurable selector

Lemma 3. *For the sup-norm on bounded functions, T is monotone and a contraction with modulus $\beta \in (0, 1)$: $\|TV - TW\|_\infty \leq \beta\|V - W\|_\infty$. Moreover, T maps bounded continuous functions into bounded continuous functions.*

Proof. Monotonicity. If $V \leq W$ pointwise, then for any (θ, τ) , $w(\tau; \theta) + \beta \mathbb{E}[V(\theta') \mid \theta, \tau] \leq w(\tau; \theta) + \beta \mathbb{E}[W(\theta') \mid \theta, \tau]$, and taking supremum over τ preserves the inequality.

Contraction. For any θ and τ ,

$$|\beta \mathbb{E}[V(\theta') - W(\theta') \mid \theta, \tau]| \leq \beta \|V - W\|_\infty,$$

hence by taking the supremum over τ and then over θ , $\|TV - TW\|_\infty \leq \beta \|V - W\|_\infty$.

Continuity. Fix bounded continuous V . By Assumption 2, $(C(\tau), L(\tau), R(\tau))$ are continuous in τ ; $w(\tau; \theta)$ is continuous in (θ, τ) (affine in θ , continuous in τ). The signal kernel is continuous in (θ, τ) because q_H, q_O are continuous and R is continuous, so $(\theta, \tau) \mapsto \mathbb{E}[V(\theta') \mid \theta, \tau]$ is continuous. With $\mathcal{T}$ compact, the objective is continuous on $[0, 1] \times \mathcal{T}$; by the maximum theorem, $\theta \mapsto (TV)(\theta)$ is continuous. $\square$

Proposition 18. *There is a unique bounded continuous fixed point V^* of T . The argmax correspondence $\Gamma(\theta) \equiv \arg \max_{\tau \in \mathcal{T}} \{w(\tau; \theta) + \beta \mathbb{E}[V^*(\theta') \mid \theta, \tau]\}$ is nonempty, compact-valued and upper hemicontinuous; hence there exists a Borel-measurable selector $\tau^*(\theta) \in \Gamma(\theta)$.*

Proof. By Lemma 3, T is a sup-norm contraction on the complete metric space of bounded functions. Banach's fixed-point theorem yields a unique bounded fixed point V^* ; continuity follows from the continuity part of Lemma 3. For the selector, continuity of the objective in (θ, τ) and compactness of $\mathcal{T}$ imply $\Gamma(\theta)$ is nonempty and compact-valued with a closed graph; upper hemicontinuity then holds by the maximum theorem. A Borel-measurable selector exists by a measurable maximum theorem. $\square$

B.2 Convexity and monotonicity in reputation

We record two simple facts about the belief kernel.

Lemma 4. *For any prior θ and revenue R , the posterior $\hat{\theta}$ after observing $s \in \{0, 1\}$ satisfies $\mathbb{E}[\hat{\theta} \mid \theta, R] = \theta$. The next-period prior is $\theta' = \Phi(\hat{\theta})$ with Φ affine, hence $\mathbb{E}[\theta' \mid \theta, R] = \Phi(\theta)$.*

Proof. Bayes' rule gives

$$\hat{\theta}(1) = \frac{\theta q_H(R)}{\theta q_H(R) + (1 - \theta) q_O(R)}$$

and

$$\hat{\theta}(0) = \frac{\theta(1 - q_H(R))}{\theta(1 - q_H(R)) + (1 - \theta)(1 - q_O(R))}.$$

Weighting by $p_1(\theta, R)$ and $1 - p_1(\theta, R)$ yields $\mathbb{E}[\hat{\theta} \mid \theta, R] = \theta$ (law of iterated expectations). Since Φ is affine, $\mathbb{E}[\theta' \mid \theta, R] = \Phi(\mathbb{E}[\hat{\theta}]) = \Phi(\theta)$. $\square$

Lemma 5. *Let V be convex and bounded. Under the MLRP condition on (q_H, q_O) , the mapping $M_V(\theta; R) \equiv \mathbb{E}[V(\theta') \mid \theta, R]$ is increasing and convex in θ for each R .*

Proof. Fix R . Under MLRP, the likelihood ratio for $s = 1$ relative to $s = 0$ is increasing in θ , which implies the posterior kernel $\theta \mapsto \hat{\theta}(\cdot; \theta, R)$ is a mean-preserving spread that is increasing in the convex order. Since $V \circ \Phi$ is convex (composition of convex with affine), Jensen's inequality and the convex-order monotonicity imply that $\theta \mapsto \mathbb{E}[V(\theta') \mid \theta, R]$ is convex and increasing. $\square$

Proposition 19. *If V is bounded, increasing, and convex on $[0, 1]$, then TV is bounded, increasing, and convex. In particular, the fixed point V^* is increasing and convex.*

Proof. Boundedness is immediate. For any τ , the current term $w(\tau; \theta)$ is affine and increasing in θ (slope $R(\tau) \geq 0$). By Lemma 5, the continuation term $M_V(\theta; R(\tau))$ is increasing and convex in θ . The supremum over τ of functions that are each increasing and convex preserves these properties. Iterating T from $V_0 \equiv 0$ (which is increasing and convex) yields a pointwise limit V^* with the same properties. $\square$

Corollary 2. *Let $\tau^*(\theta) \in \Gamma(\theta)$ be an equilibrium selector. Then $R^*(\theta) \equiv R(\tau^*(\theta))$ is (weakly) increasing in θ .*

Proof. By Proposition 19, V^* is increasing and convex. The objective $\Psi(\theta, \tau) = w(\tau; \theta) + \beta M_{V^*}(\theta; R(\tau))$ has increasing differences in (θ, R) : the derivative in θ equals $R(\tau) + \beta \partial_\theta M_{V^*}(\theta; R(\tau))$, which is increasing in R because M_{V^*} is supermodular in (θ, R) under MLRP. Since R is a (continuous) function of τ , standard monotone comparative statics yield that any measurable maximizer has nondecreasing $R(\tau^*(\theta))$ in θ . $\square$

B.3 Dynamic threshold via local expansion

For the local statement we add a mild differentiability assumption at the origin.

Assumption A1. Along a revenue-raising one-parameter policy path $\varepsilon \mapsto \tau(\varepsilon)$ with $\tau(0) = 0$, we have: (i) $R(\tau(\varepsilon))$ is differentiable at 0 with $\dot{R}(0) > 0$; (ii) q_H and q_O are differentiable at $R = 0$.

Lemma 6. *Let V be bounded and continuous. Under Assumption A1, the Gateaux derivative of $(TV)(\theta)$ in the direction $\tau(\varepsilon)$ at $\varepsilon = 0$ exists and equals*

$$\left. \frac{d}{d\varepsilon} (TV)(\theta) \right|_{\varepsilon=0} = \left(\theta - U'(Y^0) \right) \dot{R}(0) + \beta \Delta_V(\theta) \dot{R}(0),$$

where

$$\Delta_V(\theta) := \left. \frac{\partial}{\partial R} \mathbb{E}[V(\theta') \mid \theta, R] \right|_{R=0}.$$

Moreover, $\Delta_V(\theta) \geq 0$ if V is convex and (q_H, q_O) satisfy MLRP.

Proof. At $\tau = 0$, the competitive allocation (C^0, L^0, Y^0) solves $V'(L^0) = U'(Y^0)f'(L^0)$. An infinitesimal tax changes welfare through the revenue channel only (envelope logic at the competitive allocation), so the derivative of the current term equals $(\theta - U'(Y^0))\dot{R}(0)$. For the continuation term, the belief kernel depends on ε only via $R(\tau(\varepsilon))$. Differentiability of q_H, q_O at 0 and boundedness of V justify differentiating under the expectation:

$$\left. \frac{d}{d\varepsilon} \beta \mathbb{E}[V(\theta') \mid \theta, \tau(\varepsilon)] \right|_{\varepsilon=0} = \beta \Delta_V(\theta) \dot{R}(0).$$

To sign $\Delta_V(\theta)$ when V is convex, observe that an increase in R (weakly) increases informativeness in the Blackwell sense under MLRP. Expected values of convex functions are monotone in the Blackwell order, hence $\partial_R \mathbb{E}[V(\theta') \mid \theta, R] \big|_{R=0} \geq 0$. $\square$

Proposition 20. *Let $\bar{\theta}^{\text{stat}} = U'(Y^0)$ be the static cutoff. Under Assumption A1 and MLRP, the dynamic cutoff*

$$\bar{\theta}^{\text{dyn}} := \inf \left\{ \theta : \exists \text{ revenue-raising path } \tau(\varepsilon) \text{ with } \left. \frac{d}{d\varepsilon} (TV^*)(\theta) \right|_{\varepsilon=0} \geq 0 \right\}$$

obeys $\bar{\theta}^{\text{dyn}} \leq \bar{\theta}^{\text{stat}}$, with equality iff the signal is locally uninformative at $R = 0$ (i.e., $q'_H(0) = q'_O(0)$ so that $\Delta_{V^}(\theta) = 0$ for all θ).*

Proof. Apply Lemma 6 with $V = V^*$ (bounded, convex by Proposition 19). The derivative equals $(\theta - \bar{\theta}^{\text{stat}})\dot{R}(0) + \beta \Delta_{V^*}(\theta)\dot{R}(0)$. Since $\dot{R}(0) > 0$ and $\Delta_{V^*}(\theta) \geq 0$, the smallest θ that makes the derivative nonnegative is weakly below $\bar{\theta}^{\text{stat}}$, with equality iff $\Delta_{V^*}(\theta) \equiv 0$ near the origin (local noninformativeness). $\square$

B.4 Scale–mix separation in the recursive problem

Lemma 7. *For any θ , the maximization over $\tau \in \mathcal{T}$ is equivalent to*

$$\max_{(S,R) \in \mathcal{F}_0} \left\{ U(Y(S) - R) - V(L(S)) + \theta R - \tilde{a}(S, R) + \Gamma_{V^*}(\theta; R) \right\},$$

where $\mathcal{F}_0$ is the static feasibility set of implementable (S, R) and

$$\tilde{a}(S, R) = \min_{\tau \in \mathcal{T}} \{a(\tau) : (1 - \tau_L)(1 - \tau_B) = S, R(\tau) = R\}$$

is the minimal instrument cost of implementing (S, R) .

Proof. By Lemma 2 (static benchmark), private allocations (C, L, Y) depend on τ only through $S = (1 - \tau_L)(1 - \tau_B)$. The signal kernel depends on τ only through $R(\tau)$. Therefore

the Bellman objective at (θ, τ) depends on τ only via (S, R) and the direct instrument cost $a(\tau)$. For any implementable (S, R) , the set of mixes that deliver (S, R) is nonempty and compact; minimizing $a(\tau)$ over that set yields $\tilde{a}(S, R)$ by Weierstrass. Taking the supremum over τ is thus equivalent to taking the supremum over implementable pairs (S, R) of the reduced objective. $\square$

Proposition 21. *Suppose a_i are strictly convex in their arguments. For each θ :*

1. *The choice of (S, R) solves*

$$(S^*(\theta), R^*(\theta)) \in \arg \max_{(S, R) \in \mathcal{F}_0} U(Y(S) - R) - V(L(S)) + \theta R - \tilde{a}(S, R) + \Gamma_{V^*}(\theta; R).$$

2. *Conditional on (S^*, R^*) , the set of mixes $\{\tau \in \mathcal{T} : (1 - \tau_L)(1 - \tau_B) = S^*, R(\tau) = R^*\}$ is nonempty and compact, and the strictly convex program $\min_{\tau} a(\tau)$ over that set has a unique minimizer $(\tau_L^\dagger, \tau_B^\dagger)$.*

Thus the scale $R^*(\theta)$ is chosen by trading off current utility, expected public value, and the continuation value Γ_{V^*} , while the mix $(\tau_L^\dagger, \tau_B^\dagger)$ is selected statically by instrument costs alone.

Proof. (1) follows from Lemma 7. For (2), nonemptiness and compactness of the iso-feasible set are inherited from continuity of the constraints and compactness of $\mathcal{T}$. Strict convexity of a on a convex combination of mixes (the feasible set is a compact curve and admits local convexification via Lagrange multipliers) yields a unique minimizer by standard convex analysis. The continuation term $\Gamma_{V^*}(\theta; R)$ depends only on R , not on the mix, hence it plays no role in step (2). $\square$

C Proofs for Section 6

Lemma 8. *Under Assumption 2(iii), $\mathcal{F}_0$ is nonempty and compact. The objective in the Bellman operator is continuous in (θ, S, R) and measurable. Hence, by Berge's Maximum Theorem, the maximizer correspondence is nonempty, compact-valued, and upper hemicontinuous; a measurable selector exists.*

Lemma 9. *For bounded V, W ,*

$$|(TV)(\theta) - (TW)(\theta)| \leq \beta \sup_{\theta' \in [0, 1]} |V(\theta') - W(\theta')|.$$

Thus T is a contraction on $(\mathcal{C}(\Theta), \|\cdot\|_\infty)$.

Proof of Theorem 6. By Lemma 9, T has a unique fixed point $V^* \in \mathcal{C}(\Theta)$. Lemma 8 gives existence and regularity of maximizers and a measurable selector. $\square$

Lemma 10. *Fix θ . If two experiments Q', Q satisfy Blackwell dominance (Q' is more informative than Q), then for every convex v , $\mathbb{E}[v(\theta')|\theta, R; Q'] \geq \mathbb{E}[v(\theta')|\theta, R; Q]$. If Q satisfies MLRP in R , then for convex v , the mapping $R \mapsto \mathbb{E}[v(\theta')|\theta, R; Q]$ is increasing.*

Lemma 11. *If V is increasing (resp. convex), then so is $\Gamma_V(\theta; R) = \beta \mathbb{E}[V(\theta')|\theta, R]$ in θ . Since the current payoff is affine in θ , the pointwise maximum over (S, R) preserves monotonicity and convexity in θ . Hence T maps increasing (resp. convex) functions to increasing (resp. convex) functions.*

Proof of Proposition 7. Iterate T from any bounded V_0 that is increasing and convex in θ (e.g., a constant). By Lemma 11, all iterates are increasing and convex; by contraction, $T^n V_0 \rightarrow V^*$, which inherits these properties. $\square$

Lemma 12. *Assume the objective is continuously differentiable in R at $R = 0$ and the argmax is single-valued (or admits a maximizer with $R = 0$). Then*

$$\left. \frac{\partial}{\partial R} \left[U(Y(S) - R) + \theta R + \Gamma_{V^*}(\theta; R) \right] \right|_{R=0} = -U'(Y(S)) + \theta + \partial_R \Gamma_{V^*}(\theta; R)|_{R=0}.$$

Moreover, under MLRP and dominance assumptions, $\partial_R \Gamma_{V^*}(\theta; 0) \geq 0$, with strict inequality if the experiment is informative at zero.

Proof of Proposition 4. Define the marginal value at zero as

$$M(\theta) := \max_{S \in \mathcal{S}} \{-U'(Y(S)) + \theta + \partial_R \Gamma_{V^*}(\theta; 0)\}.$$

By Lemma 12, if $M(\theta) < 0$ then any positive R strictly lowers the objective, so $R^*(\theta) = 0$. If $M(\theta) > 0$ the optimum has $R^*(\theta) > 0$. Continuity implies a threshold $\bar{\theta}^{\text{dyn}} := \inf\{\theta : M(\theta) \geq 0\}$. Under informativeness at zero, $\partial_R \Gamma_{V^*}(\theta; 0) > 0$, so the dynamic threshold is strictly below the static cutoff (where $\partial_R \Gamma_{V^*} \equiv 0$). Monotonicity of $R^*(\theta)$ follows from increasing differences in (θ, R) of the objective (Milgrom–Shannon single crossing). $\square$

Lemma 13. *Let $\mathcal{F}_0$ be the image of $\mathcal{T}$ under $(\tau_L, \tau_B) \mapsto (S(\tau), R(\tau))$. Then*

$$\begin{aligned} & \max_{\tau \in \mathcal{T}} \left\{ U(Y(S(\tau)) - R(\tau)) - V(L(S(\tau))) + \theta R(\tau) + \Gamma_{V^*}(\theta; R(\tau)) \right\} \\ &= \max_{(S, R) \in \mathcal{F}_0} \left\{ U(Y(S) - R) - V(L(S)) + \theta R + \Gamma_{V^*}(\theta; R) \right\}. \end{aligned} \quad (13)$$

Proof of Proposition 8. Under instrument-symmetric monitoring, $\Gamma_{V^*}(\theta; R)$ depends only on R . By Lemma 13, the problem separates into choosing (S, R) and then any mix implementing it. Indeterminacy of the mix along the frontier follows since all such mixes deliver the same (S, R) and hence the same value. $\square$

Remark 4. Let two instruments $\iota \in \{A, B\}$ implement the same (S, R) . Suppose $Q(\cdot|g, R, \iota = A)$ Blackwell-dominates $Q(\cdot|g, R, \iota = B)$. Then for convex V^* , $\mathbb{E}[V^*(\theta')|\theta, R, \iota = A] > \mathbb{E}[V^*(\theta')|\theta, R, \iota = B]$, so Γ_{V^*} depends on ι and the planner strictly prefers the more informative instrument. Hence separation fails.

D Proofs for Section 8

Throughout, primitives satisfy Assumption 2. We use the dynamic objects and notation from B. In particular, V^* denotes the unique bounded continuous fixed point, which is increasing and convex by Proposition 19. We rely on Lemmas 4–5 (posterior martingale and convex order) and on the two-stage reduction in Lemma 7.

D.1 Proof of Proposition 9

Strict concavity and continuity imply that for any fixed S the competitive equilibrium in the private sector is unique and varies continuously with S (standard arguments via firm and household FOCs and the implicit function theorem). Define $Y(S)$ as equilibrium output and $L(S)$ as equilibrium labor; then Y is increasing and concave in S , L is weakly decreasing. Continuity of the instrument map $\tau \mapsto (S(\tau), R(\tau))$ from a compact $\mathcal{T}$ implies $\mathcal{F}_0$ is nonempty and compact by continuity and the extreme-value theorem.

D.2 Proof of Proposition 10

Fix θ and (S, R) . Let V be convex and bounded. If $Q_{\iota'}(\cdot|g, R)$ Blackwell-dominates $Q_{\iota}(\cdot|g, R)$ for each R , then by Blackwell's theorem the posterior distribution of θ^+ under ι' is a mean-preserving spread of that under ι (conditional on R), so $\mathbb{E}[V(\theta^+)|\theta, R; \iota'] \geq \mathbb{E}[V(\theta^+)|\theta, R; \iota]$. Hence the continuation term satisfies $\Gamma_V(\theta; R; \iota') \geq \Gamma_V(\theta; R; \iota)$ pointwise in R . The Bellman objective thus has increasing differences in (R, ι) (indeed, a nonnegative parallel shift in R), and Milgrom–Shannon monotone comparative statics yields $R_{\iota'}^*(\theta) \geq R_{\iota}^*(\theta)$. The derivative-at-zero argument from Proposition 4 then implies $\bar{\theta}^{\text{dyn}}(\iota') \leq \bar{\theta}^{\text{dyn}}(\iota)$. Strict inequalities follow when the solution is interior and the Blackwell improvement is strict on a set of signals of positive measure.

D.3 Proof of Example 1

For fixed R , decreasing the bit-flip noise ε yields an experiment that is a mean-preserving spread (via the Blackwell ordering of binary symmetric channels). Therefore $\mathbb{E}[V^*(\theta^+) | \theta, R; \varepsilon]$ is increasing in R and (pointwise) nonincreasing in ε for convex V^* . The same single-crossing argument as in the proof of Proposition 10 establishes the claims.

D.4 Proof of Proposition 11

Under $\Pi' \succeq \Pi$, the type process is more persistent in the sense of first-order stochastic dominance of the next type conditional on the current one. Because posteriors are martingales and V^* is increasing and convex (Proposition 7), more persistent types increase the value of informative signals in the convex order: for any R and θ , $\mathbb{E}[V^*(\theta^+) | \theta, R; \Pi'] \geq \mathbb{E}[V^*(\theta^+) | \theta, R; \Pi]$. Hence the Bellman objective has increasing differences in (R, Π) and the same single-crossing argument implies $R_{\Pi'}^*(\theta) \geq R_{\Pi}^*(\theta)$ and $\bar{\theta}^{\text{dyn}}(\Pi') \leq \bar{\theta}^{\text{dyn}}(\Pi)$. Strict inequalities obtain at interior solutions when persistence strictly rises.

D.5 Proof of Proposition 12

(i) With an aggregate signal that depends only on delivered revenue, belief dynamics and the continuation term are functions of (θ, R) alone, independent of the incidence of taxes across types. The convexity and monotonicity arguments in θ carry over verbatim (Proposition 7), yielding an increasing $R^*(\theta)$ and a dynamic threshold via the same derivative-at-zero argument. (ii) For a fixed (S, R) , the set of mixes (τ_L, τ_B) that implement (S, R) yield different individual allocations (C_i, L_i) across types because the tax wedge on labor changes type-specific margins when $V_1 \neq V_2$. A utilitarian aggregator is strictly concave in (C_1, C_2) and convex in (L_1, L_2) , so the social marginal cost of revenue via labor differs from that via commodity unless $V_1 \equiv V_2$. Therefore the equivalence frontier collapses generically: the planner's FOCs pick a unique mix that minimizes incidence-weighted deadweight loss for the given (S, R) .

D.6 Proof of Proposition 13

Fix θ and primitives. Let $\iota \mapsto (q_H', q_O')$ be the family of experiments with ι' Blackwell-dominating ι if $\iota' > \iota$. For any bounded convex V , Blackwell dominance implies

$$\mathbb{E}[V(\theta') | \theta, R; \iota'] \geq \mathbb{E}[V(\theta') | \theta, R; \iota] \quad \text{for all } (\theta, R),$$

with strict inequality on a set of positive measure when the solution is interior (standard convex-order/Blackwell argument; cf. Lemma 5). Hence the Bellman objective

$$\Psi(\theta, R; \iota) = U(Y(S) - R) - V(L(S)) + \theta R + \Gamma_{V^*}^\iota(\theta; R)$$

has *increasing differences* in (R, ι) (the current term does not depend on ι). Milgrom–Shannon monotone comparative statics then imply $R_{\iota'}^*(\theta) \geq R_\iota^*(\theta)$. For the cutoff, the local derivative at zero (Appendix Lemma 6) is

$$\left. \frac{d}{d\varepsilon} (TV^*)(\theta) \right|_{\varepsilon=0} = (\theta - \bar{\theta}^{\text{stat}}) \dot{R}(0) + \beta \Delta_{V^*}^\iota(\theta) \dot{R}(0),$$

where $\Delta_{V^*}^\iota(\theta) = \partial_R \mathbb{E}[V^*(\theta') \mid \theta, R; \iota]_{R=0}$. Blackwell dominance yields $\Delta_{V^*}^{\iota'}(\theta) \geq \Delta_{V^*}^\iota(\theta)$, so the smallest θ making the derivative nonnegative weakly decreases as ι rises; thus $\bar{\theta}^{\text{dyn}}(\iota') \leq \bar{\theta}^{\text{dyn}}(\iota)$. $\square$

D.7 Proof of Proposition 14

Under perfect monitoring, the public signal reveals g_t independently of R . Therefore the posterior $\hat{\theta}$ and the next prior $\theta' = \Phi(\hat{\theta})$ are independent of R , so $\Gamma_{V^*}(\theta; R)$ is constant in R . The Bellman objective reduces to

$$\max_{(S, R) \in \mathcal{F}_0} U(Y(S) - R) - V(L(S)) + \theta R + \text{constant},$$

and the optimal $R^*(\theta)$ and the zero-tax cutoff coincide with the static benchmark (Proposition 1). $\square$

D.8 Proof of Proposition 15

With enforcement fraction $\varphi \in [0, 1]$, delivered G equals R for H and φR for O , so the contemporaneous marginal value of R becomes $\theta + (1 - \theta)\varphi$. The static scale rule therefore reads $U'(C^*) = \theta + (1 - \theta)\varphi$. If φ is publicly verifiable, the signal becomes less type-discriminating for a given R (types become more observationally equivalent), so for any convex V^* ,

$$\mathbb{E}[V^*(\theta') \mid \theta, R; \varphi] \text{ is (weakly) less increasing in } R \text{ as } \varphi \uparrow,$$

i.e., the informational term weakens in the Blackwell order. The dynamic derivative at zero therefore changes to

$$(\theta - \bar{\theta}^{\text{stat}}(\varphi)) \dot{R}(0) + \beta \Delta_{V^*}^\varphi(\theta) \dot{R}(0),$$

with $\bar{\theta}^{\text{stat}}(\varphi) = U'(Y^0)$ evaluated at the same Y^0 but with contemporaneous weight $\theta + (1 - \theta)\varphi$ entering the scale rule. The first term lowers the cutoff in φ (direct scale effect), while the second raises it (information dilution); the net effect is the sum stated in the proposition. In the limit $\varphi \rightarrow 1$, types are observationally equivalent, the informational term vanishes, and the problem collapses to static Ramsey with weight one on G . $\square$

D.9 Proof of Proposition 16

Under representative-agent separability, the private allocation is characterized by intertemporal wedges. For any implementable path of wedges there exists an *intertemporal equivalence frontier*: sequences $\{\tau_{C,t}, \tau_{K,t}\}$ that replicate the same wedges (hence the same allocation) for a given revenue path $\{R_t\}$. If the public signal depends only on delivered revenue, the continuation value depends only on $\{R_t\}$ and not on the composition of instruments. Applying the two-stage reduction period by period (Lemma 7) shows that the dynamic problem selects $\{R_t^*\}$ and leaves the mix indeterminate along the frontier in the baseline. If, however, one instrument yields a strictly more informative experiment per unit revenue (Blackwell dominance conditional on the instrument), the objective exhibits increasing differences between “use of that instrument” and the value function, tilting the frontier toward that instrument and breaking separation. $\square$

D.10 Proof of Proposition 17

Let the government aggregate individual utilities with a strictly concave social aggregator. The belief process and the signal depend only on aggregate delivered revenue R , not on incidence. Then the Bellman continuation term retains the same convex-order monotonicity in θ as in Lemma 5, so V^* remains increasing and convex. As a result, the optimal revenue scale $R^*(\theta)$ is (weakly) increasing in θ and a dynamic trust threshold exists by the same derivative-at-zero argument (Lemma 6). However, because incidence differs across instruments, the static equivalence frontier breaks under heterogeneity: for a given (S, R) , different mixes (τ_L, τ_Y) generally yield different social marginal costs. Hence the mix is determined by distributional concerns even in the baseline, while the belief-scale logic remains unchanged as long as the signal is aggregate in R . If monitoring is group-specific, composition can also affect informativeness, and the separation result can fail by the same argument as in Proposition 16. $\square$

Table 1: Notation used in the paper

$\beta \in (0, 1)$	Discount factor.
$\theta \in [0, 1]$	Public belief (prior) that the government is honest.
$g \in \{H, O\}$	Government type: honest (H); opportunistic (O).
Π	Type transition matrix.
$\Phi(\cdot)$	Next-period prior given current posterior.
$s \in \mathcal{S}, Q(\cdot \mid g, R)$	Public signal and signal kernel; Q dominated and measurable.
ι	Index of experiment informativeness.
$S \in (0, 1]$	Net-of-tax product (private wedge).
$R \in [0, \bar{R}]$	Delivered revenue (scale).
τ_L, τ_Y	Labor and commodity/output tax instruments.
$\mathcal{F}_0$	Feasible frontier in (S, R) induced by $\mathcal{T}$.
$Y(S), L(S)$	Competitive output and labor induced by S .
C	Private consumption: $C = Y(S) - R$.
$U(\cdot), V(\cdot)$	Preferences: period utility $U(C) - V(L) + \theta R$.
$V^*(\theta)$	Government value function.
$\Gamma_V(\theta; R)$	Continuation term $\beta \mathbb{E}[V(\theta^+) \mid \theta, R]$.
$\bar{\theta}^{\text{stat}}, \bar{\theta}^{\text{dyn}}$	Static and dynamic trust cutoffs.
λ, φ	Fractions for verifiable delivery and enforcement.

E Notation and Objects

References

- Aiyagari, S. R. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy*, 103(6):1158–1175.
- Athey, S. (2002). Monotone comparative statics under uncertainty. *Quarterly Journal of Economics*, 117(1):187–223.
- Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.
- Barro, R. J. and Gordon, D. B. (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1):101–121.
- Blackwell, D. (1953). Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24(2):265–272.
- Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54(3):607–622.
- Chari, V. V., Christiano, L. J., and Kehoe, P. J. (1994). Optimal fiscal policy in a business cycle model. *Journal of Political Economy*, 102(4):617–652.

- Chari, V. V. and Kehoe, P. J. (1990). Sustainable plans. *Journal of Political Economy*, 98(4):783–802.
- Diamond, P. A. and Mirrlees, J. A. (1971a). Optimal taxation and public production i: Production efficiency. *American Economic Review*, 61(1):8–27.
- Diamond, P. A. and Mirrlees, J. A. (1971b). Optimal taxation and public production ii: Tax rules. *American Economic Review*, 61(3):261–278.
- Fudenberg, D., Gao, Y., and Pei, H. (2022). A reputation for honesty. *Journal of Economic Theory*, 204. Article no. 105450.
- Halac, M. and Yared, P. (2014). Fiscal rules and discretion under persistent shocks. *Econometrica*, 82(5):1557–1614.
- Halac, M. and Yared, P. (2018). Fiscal rules and discretion in a world economy. *American Economic Review*, 108(8):2305–2334.
- Halac, M. and Yared, P. (2022). Fiscal rules and discretion under limited enforcement. *Econometrica*, 90(5):2093–2127.
- Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59–83.
- Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491.
- Lu, Y. K. (2013). Optimal policy with credibility concerns. *Journal of Economic Theory*, 148(5):2007–2032.
- Lukyanov, G. and Ablyatifov, E. (2025). Optimal taxation under imperfect trust. Submitted to *Economic Theory Bulletin*.
- Milgrom, P. and Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1):157–180.
- Phelan, C. (2006). Public trust and government betrayal. *Journal of Economic Theory*, 130(1):27–43.
- Ramsey, F. P. (1927). A contribution to the theory of taxation. *The Economic Journal*, 37(145):47–61.
- Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press, Princeton, NJ.