The Geometry of Ramsey Alignment: Measuring the Efficiency of the Postwar Tax System

Jackson Mejia*

January 2026

Was the postwar shift from corporate to personal taxation welfare-improving? I develop a sufficient statistics framework to evaluate the alignment of the tax system with the Ramsey optimum in dynamic general equilibrium. Alignment is captured by the angle between two vectors: one measuring each tax's first-order welfare cost, the other measuring its effect on total revenue. The sine squared of the angle measures the share of welfare gains available through tax reform. Using narrative tax shocks, I measure the full matrix of present-value own- and cross-base elasticities in general equilibrium. Applying this framework to the United States, I find that the system moved from severely misaligned in the 1950s to nearly optimal by the 2010s.

JEL-Classification: H21, H25, E62, C32

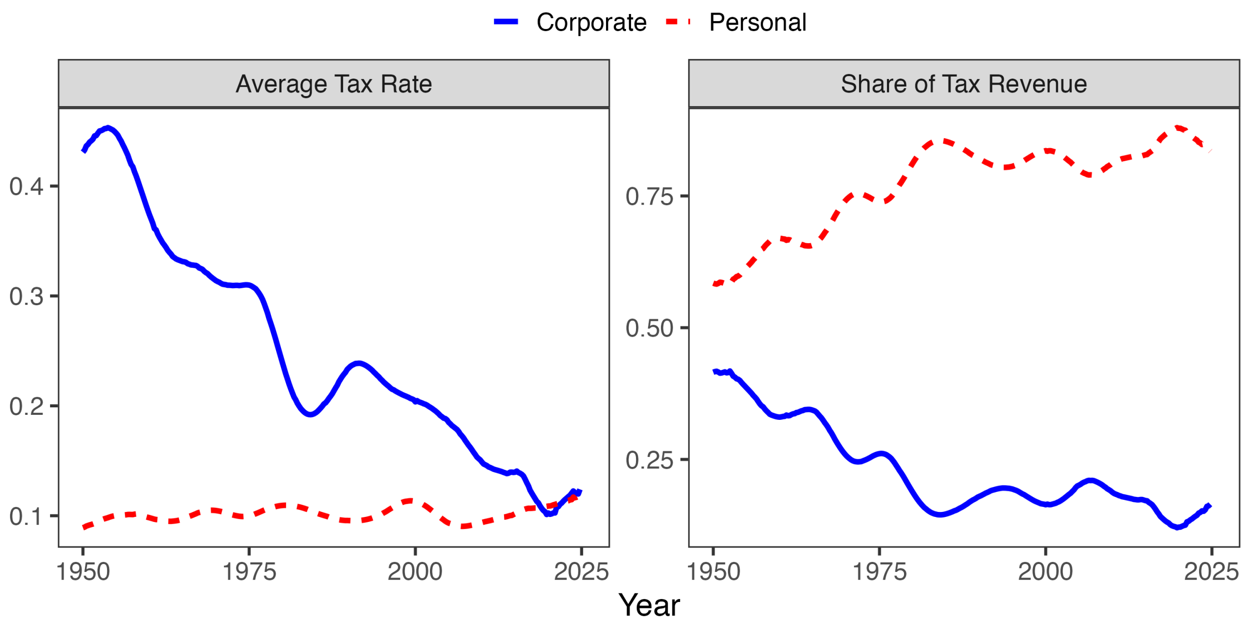
Keywords: Optimal taxation, tax reform, sufficient statistics, marginal cost of public funds, corporate taxation, narrative identification, fiscal policy

*Massachusetts Institute of Technology, jpmejia@mit.edu. First Draft: January 2026. I am grateful to Andrew Johnston, Jim Poterba, and Chris Wolf for comments and suggestions.

1 Introduction

At the end of World War II, corporate income taxes raised nearly half of federal income tax revenue. Since then, personal taxes have risen to over 85 percent of receipts, while the corporate share has fallen below 15 percent. Was this shift welfare-improving? I find that it was.

Figure 1: Changes in Postwar Sources of Revenue and Tax Rates



Note: The right panel assumes that personal and corporate taxes make up 100% of tax revenue. In reality, they sum to just over 90%. All series are trend components from the two-sided Christiano and Fitzgerald (2003). I strip out all components with a frequency less than forty quarters. See the appendix for details on data construction.

How should we evaluate this shift? Any answer requires knowing how taxes affect behavior across bases and over time. Existing methods struggle on both dimensions. When corporate taxes fall, the effects extend beyond the corporate base: wages, investment, and organizational form adjust, shifting revenue between bases. Indeed, perhaps the primary supply-side motivation for corporate tax cuts is precisely these cross-base effects, namely that lower corporate taxes will raise wages and expand the personal income tax base (Lucas 1990). Moreover, behavioral responses evolve over time. Two existing approaches handle these forces differently. Structural general equilibrium models capture both, but require parametric commitment and fundamentally disagree. Neoclassical models conclude the postwar shift was optimal (Chari, Nicolini, and Teles 2020), while overlapping generations models reach the opposite conclusion (Conesa, Kitao, and Krueger 2009).[1] Sufficient statistics approaches avoid this model dependence by working directly with elasticities (Chetty 2009), but existing implementations typically rely on rigorously estimated but static own-base responses (Saez 2001; Hendren and Sprung-Keyser 2020), missing the cross-base spillovers and dynamic forces.[2] I develop a scalar metric that summarizes how far a tax system is from optimal, and estimate the cross-base elasticities needed to compute it. The corporate base turns out to be highly elastic to both taxes; the personal base is not. But fiscal externalities scale with revenue, not the base, and as corporate tax rates fell, so did the distortion.

The framework compares two objects: a welfare gradient and a revenue gradient. The welfare gradient captures how marginal changes in each tax affect social welfare; under utilitarian weights, this is simply the vector of tax bases. The revenue gradient captures how marginal changes in each tax affect present-value government revenue, incorporating both own-base and cross-base responses as they unfold over time. At the Ramsey optimum, these gradients are proportional and no revenue-neutral reform can improve welfare. Away from optimum, they diverge. The angle θ between them characterizes efficiency: $\cos\theta$ measures alignment with the optimum, while $\sin^2\theta$ measures the share of potential welfare gains achievable through revenue-neutral reform. When $\cos\theta < 0$, the system is on the wrong side of the Laffer curve. The focus on revenue-neutral reform isolates optimal tax composition from optimal government size; I evaluate the local direction of reform rather than computing globally optimal taxes.

The projection approach builds on the classical marginal reform literature. Feldstein (1976) argued that computing globally optimal taxes is impractical; public finance should instead ask whether reforms improve welfare. This paper takes that advice: rather than computing optimal tax rates, I measure whether the postwar shift moved toward or away from optimum. The theoretical tools come from Dixit (1975), who derived conditions under which marginal reforms improve welfare, and Tirole and Guesnerie (1981), who showed that the optimal direction is the orthogonal projection of the welfare gradient onto the revenue-neutral hyperplane. Ahmad and Stern (1984) applied these ideas to Indian indirect taxes, but their approach—like others that followed—was static and required complete elasticity matrices from cross-sectional micro data. I introduce scalar alignment metrics ($\cos\theta$, $\sin^2\theta$) that summarize proximity to optimum, and estimate the required general equilibrium elasticities from time-series impulse responses.

Recent work on the marginal value of public funds (MVPF) clarifies what the first-order

1. Even within the neoclassical model, there is significant room for disagreement. Straub and Werning (2020) show that the optimal tax on capital—and hence the optimal direction of reform—is sensitive to the structure of preferences and heterogeneity.

2. There are other reasons the approach may fail even within the context of the Diamond and Mirrlees (1971) model from which it stems. For example, Bhandari, Borovička, and Yao (2025) point to robustness concerns, echoing an earlier point from Mankiw, Weinzierl, and Yagan (2009) that the Diamond and Saez (2011) result depends critically on the unknown distribution of types. Moreover, Saez, Slemrod, and Giertz (2012) argue, as this paper does, that it is critical to account for spillovers between instruments—in their case, between capital gains and labor income.

approach requires empirically (Hendren 2016; Finkelstein and Hendren 2020). To evaluate a marginal reform, one needs willingness-to-pay and the fiscal externality—both observable directly from the data or from causal evidence. Bergstrom, Dodds, and Rios (2024) extend this logic to multiple instruments, characterizing optimal reforms as raising revenue from the lowest-MVPF policy and spending on the highest. Neither the classical nor modern literature produces scalar metrics that summarize how close the system is to optimal or how much welfare improvement remains available. I introduce such metrics ($\cos\theta$, $\sin^2\theta$), and estimate the required inputs from dynamic impulse responses rather than static elasticities.

I estimate the revenue gradient using narrative tax shocks to the personal and corporate income tax bases. Mertens and Ravn (2013a, 2014), extending Romer and Romer (2010), construct exogenous shocks for both corporate and personal income taxes; I use the extension from Cloyne et al. (2022) through 2017. For each instrument, I cumulate discounted impulse responses to construct present-value revenue gradients. This builds on McKay and Wolf (2023) and Caravello, McKay, and Wolf (2024), who show that impulse responses to policy shocks are sufficient to construct Lucas-critique-robust counterfactuals.[3] My application is more limited: I evaluate whether actual tax changes moved toward or away from the locally optimal direction, which requires only the revenue gradients at the observed baseline.[4]

The elasticities this approach delivers differ from taxable income elasticities in public finance, which measure individual responses to marginal rates holding prices fixed. Standard estimates range from 0.1 to 0.4 (Saez, Slemrod, and Giertz 2012). Mertens and Montiel Olea (2018) estimate larger aggregate elasticities around 1.2 using narrative shocks, closer to my approach. Badel, Huggett, and Luo (2020) and Kleven et al. (2025) argue that even these understate long-run responses because they miss human capital accumulation. My estimates go further: they capture how aggregate bases respond to aggregate average rate changes, allowing wages, interest rates, and organizational form to adjust in general equilibrium over a five-year horizon. The average rate reflects the combined effect of all marginal rates, deductions, credits, and base definitions that move together in actual legislation, the bundle of incentives that constitutes a real-world tax reform. The approach bridges structural and sufficient statistics methods: it captures cross-base spillovers that partial equilibrium misses and dynamic accumulation that static approaches ignore, while avoiding the parametric disagreements that divide structural models.

The tax system moved from severely misaligned in the 1950s to nearly perfectly aligned by the 2010s. Alignment tracks the corporate share of income tax revenue almost perfectly. When corporate revenue was half of income tax collections, the high elasticity of the corporate base

3. Moon (2025) proposes empirical Bayes methods for incorporating uncertainty about policy impacts into optimal policy choice. I treat estimation uncertainty as measurement error, focusing on evaluating historical reforms rather than choosing policies.

4. In principle, the methodology could be extended by leveraging anticipated tax shocks.

pushed the revenue gradient away from the welfare gradient. As the corporate share fell, this wedge disappeared and the gradients converged. The results are robust to alternative discount rates, estimation methods, and welfare weights. The practical implication: there is little left to gain from reforming the income tax mix. Future efficiency improvements must come from other margins, such as base-broadening, rate structure, or the boundary between income and other taxes. The framework is silent on the optimal scale of government.

The framework also evaluates individual reforms. ERTA 1981 and TCJA 2017 moved almost exactly toward the Ramsey optimum. TRA 1986 moved the opposite direction: it shifted relative burden toward corporate income (Auerbach and Slemrod 1997), even as it improved efficiency on other margins through base-broadening and rate-flattening. But by cutting personal rates below corporate rates, TRA 1986 spurred pass-through growth that improved alignment over the following decades.

It is important to note the scope of these findings. The framework evaluates the composition of revenue between corporate and personal income taxes, not the structure within either base. Efficiency gains from flattening rate schedules, broadening bases, or reforming the treatment of capital gains operate on margins this analysis does not capture. The framework also abstracts from payroll taxes, consumption taxes, and state and local taxation, other margins where cross-base spillovers may create misalignment. Finally, the welfare analysis adopts a utilitarian benchmark; results may differ under alternative social welfare functions that weight distributional concerns more heavily.

Roadmap. The remainder of the paper proceeds as follows. Section 2 develops the projection framework for static revenue-neutral reform, showing that aggregate gradients are sufficient statistics and deriving the geometric solution. Section 3 extends the framework to incorporate dynamic effects. Sections 4 and 5 implement the framework empirically, estimating present-value revenue and welfare gradients from narrative shocks and evaluating the complete history of postwar federal reforms. Section 6 concludes.

2 The Geometry of Reform

This section formalizes the projection framework for evaluating multi-instrument tax reforms. The framework shows that the optimal direction of revenue-neutral reform depends on two sufficient statistics—the welfare gradient and the revenue gradient—which can be estimated from reduced-form evidence without specifying a full structural model. Despite the generality of the setup, the framework yields a simple geometric characterization: the optimal reform is the orthogonal projection of the welfare gradient onto the revenue-neutral hyperplane, and system efficiency is measured by a single scalar: the alignment between welfare and revenue objectives. This alignment metric is simply the cosine of the angle between the two gradients.

Section 2.1 introduces notation and derives the welfare and revenue gradients. Section 2.2 establishes the projection formula and shows the revenue gradient is a sufficient statistic. Section 2.3 develops scalar metrics that measure proximity to the Ramsey optimum and the scope for revenue-neutral reform. Section 2.4 demonstrates through a two-instrument example how partial equilibrium approaches can prescribe reforms in the opposite direction from the optimum. Section 3 extends the framework to dynamic economies.

2.1 Setup: Taxes, Bases, and Gradients

Tax instruments and bases. Let $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)^\top \in \mathbb{R}^n$ denote a vector of n tax rates levied on different economic activities. Each instrument i has a corresponding tax base $B_i(\boldsymbol{\tau})$ that depends on the entire tax vector. The vector of tax bases is $\mathbf{B} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{B}(\boldsymbol{\tau}) = (B_1(\boldsymbol{\tau}), \ldots, B_n(\boldsymbol{\tau}))^\top$. Total tax revenue is:

$$R(\boldsymbol{\tau}) = \sum_{i=1}^{n} \tau_i B_i(\boldsymbol{\tau}). \tag{1}$$

Tax bases change with tax rates for standard economic reasons. At low rates, behavior is relatively inelastic; as taxes rise, behavioral responses grow and revenue may decline. Importantly, there are spillovers from one base to the other. For example, a cut in the corporate tax rate may increase the labor tax base by raising the capital stock, which increases labor productivity and wages. Indeed, this type of spillover is the primary reason many policymakers and economists cite for corporate tax cuts.

Assumption 1 (Differentiability)**.** *Fix a baseline $\bar{\boldsymbol{\tau}} \in \mathbb{R}^n$ with $B_i(\bar{\boldsymbol{\tau}}) > 0$ for all i. Assume that $\ln B_i$ is continuously differentiable in $\ln(1-\tau_k)$ in a neighborhood of $\bar{\boldsymbol{\tau}}$, so there exists a matrix of retention-rate elasticities*

$$\boldsymbol{\varepsilon}(\bar{\boldsymbol{\tau}}) := \begin{pmatrix} \varepsilon_{11} & \cdots & \varepsilon_{1n} \\ \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \cdots & \varepsilon_{nn} \end{pmatrix}, \quad \varepsilon_{ik}(\bar{\boldsymbol{\tau}}) := \left.\frac{\partial \ln B_i(\boldsymbol{\tau})}{\partial \ln(1-\tau_k)}\right|_{\boldsymbol{\tau}=\bar{\boldsymbol{\tau}}}. \tag{2}$$

The elasticity ε_{ik} measures the percent change in base i from a one-percent decrease in the tax rate $(1-\tau_k)$. Off-diagonals ε_{ik} $(i \neq k)$ capture cross-base spillovers. For example, complementarity between capital and labor ensures that low capital taxes raise labor productivity, expanding the labor base. Assumption 1 simply encodes that the private sector is well-behaved enough to be differentiable. I intentionally leave micro-foundations unspecified, allowing the framework to accommodate any behavioral model that generates smooth tax base responses.

Given this structure, the two objects that matter for reform are how welfare and revenue respond to marginal tax changes. I now derive these gradients.

The revenue gradient. The revenue gradient, $r(\boldsymbol{\tau}) := \nabla_{\boldsymbol{\tau}} R(\boldsymbol{\tau})$, is the key input for the forthcoming reform problem. As Section 2.2 will show, this gradient serves as a sufficient statistic by aggregating all behavioral responses into a single vector that fully characterizes the optimal reform direction. Its i-th component can be derived from the full elasticity matrix $\boldsymbol{\varepsilon}$:

$$r_i = \frac{\partial R}{\partial \tau_i} = B_i \left[1 - \frac{\tau_i}{1-\tau_i} \varepsilon_{ii} - \Lambda_i \right], \qquad \Lambda_i \equiv \frac{1}{1-\tau_i} \sum_{k \neq i} \tau_k \frac{B_k}{B_i} \varepsilon_{ki}. \tag{3}$$

The i-th element of the revenue gradient is composed of three parts. The first is mechanical (+1); it is how much revenue would change if there was no fiscal externality from changing taxes. This externality is split into the own-base effect (using ε_{ii}) and the general equilibrium spillover effect (Λ_i), which aggregates all cross-base fiscal externalities. As a special case, I also define the partial equilibrium revenue gradient which has i-th entry

$$r_i^{\mathrm{PE}} = B_i \left[1 - \frac{\tau_i}{1-\tau_i} \varepsilon_{ii} \right]. \tag{4}$$

The partial equilibrium case assumes the elasticity matrix $\boldsymbol{\varepsilon}$ is diagonal. This is equivalent to assuming all cross-base spillovers are zero ($\varepsilon_{ki} = 0$ for $k \neq i$), which means the spillover term $\Lambda_i = 0$.

Welfare. The government has some money-metric welfare objective function $W(\boldsymbol{\tau})$ measured in dollars.[5] Higher W is worse (a social cost). We measure welfare changes by the dollar amount residents would pay to avoid a tax increase, aggregated across income groups using social welfare weights.

Assumption 2 (Welfare regularity). *W is Fréchet differentiable at $\bar{\boldsymbol{\tau}}$, so there exists a gradient*

$$g(\bar{\boldsymbol{\tau}}) := \nabla_{\boldsymbol{\tau}} W(\bar{\boldsymbol{\tau}}) \in \mathbb{R}^n \quad \textit{with} \quad W(\bar{\boldsymbol{\tau}} + h) = W(\bar{\boldsymbol{\tau}}) + g(\bar{\boldsymbol{\tau}})^\top h + o(\|h\|). \tag{5}$$

By the envelope theorem, the first-order welfare cost of raising tax i is the mechanical tax burden on domestic residents (Hendren 2016; Hendren and Sprung-Keyser 2020). Suppose a marginal increase in τ_i raises residents' tax payments by $\gamma_i B_i$ dollars (where $\gamma_i \in [0, 1]$ is the

5. If utility is in utils, let W be compensating variation at a fixed numeraire price and normalize the marginal value of public funds to one. All local results are invariant to this choice up to a positive scalar.

incidence share falling on domestic residents), and let $\tilde{\gamma}_i := \sum_g \omega_g s_{gi}$ be the welfare-weighted incidence constructed from group-specific social welfare weights $\omega_g > 0$ and marginal incidence shares s_{gi} (with $\sum_g s_{gi} = \gamma_i$). Then:

$$g_i(\bar{\boldsymbol{\tau}}) = \tilde{\gamma}_i B_i(\bar{\boldsymbol{\tau}}). \tag{6}$$

Equation (6) says the marginal welfare cost of instrument i is the incidence- and welfare-weighted tax base. For the empirical implementation of this framework, however, I adopt a standard utilitarian benchmark. Under this assumption, the marginal welfare cost of a tax is the size of the tax base. This simplifies the welfare gradient considerably: the welfare-weighted incidence is equal to one for domestic taxes, and the welfare gradient component g_i simplifies to the observable tax base B_i. Later, this simplification will also cleanly separate the empirical challenge. The welfare gradient g can be constructed directly from baseline data, while the revenue gradient must be estimated.

The marginal cost of public funds. Pairing the welfare gradient g_i with the revenue gradient yields the marginal cost of public funds:

$$\text{MCPF}_i = \frac{g_i}{r_i} = \frac{\tilde{\gamma}_i}{1 - \frac{\tau_i}{1-\tau_i}\varepsilon_{ii} - \Lambda_i}.$$

At a global optimum, MCPFs are equalized across instruments (Appendix A.1 derives this formally). When MCPFs differ, which is the typical case empirically, then there exists scope for welfare-improving revenue-neutral reform. Section 2.2 characterizes the optimal direction of such reforms.

2.2 Policy as Projection

The problem of reform arises when MCPFs differ across instruments. When this occurs, there exist directions of tax adjustment that improve welfare while holding total revenue fixed. The planner's local problem is to identify the reform direction $d\boldsymbol{\tau}$ that maximizes the first-order welfare gain subject to revenue neutrality:

$$\max_{d\boldsymbol{\tau}} \quad -g(\bar{\boldsymbol{\tau}})^\top d\boldsymbol{\tau} \quad \text{subject to} \quad r(\bar{\boldsymbol{\tau}})^\top d\boldsymbol{\tau} = 0. \tag{7}$$

The welfare gradient g measures how marginal tax changes affect social welfare; the revenue gradient r measures how they affect government revenue, accounting for all behavioral responses including cross-base spillovers. This constrained optimization problem has a simple geometric solution, established by Tirole and Guesnerie (1981): the optimal reform direction is the orthogonal projection of the welfare gradient onto the revenue-neutral hyperplane.

Lemma 1 (Gradient projection; Tirole and Guesnerie (1981)). *Let $\bar{\boldsymbol{\tau}}$ be a baseline with welfare gradient $g(\bar{\boldsymbol{\tau}})$ and revenue gradient $r(\bar{\boldsymbol{\tau}}) \neq 0$. The optimal revenue-neutral reform direction is*

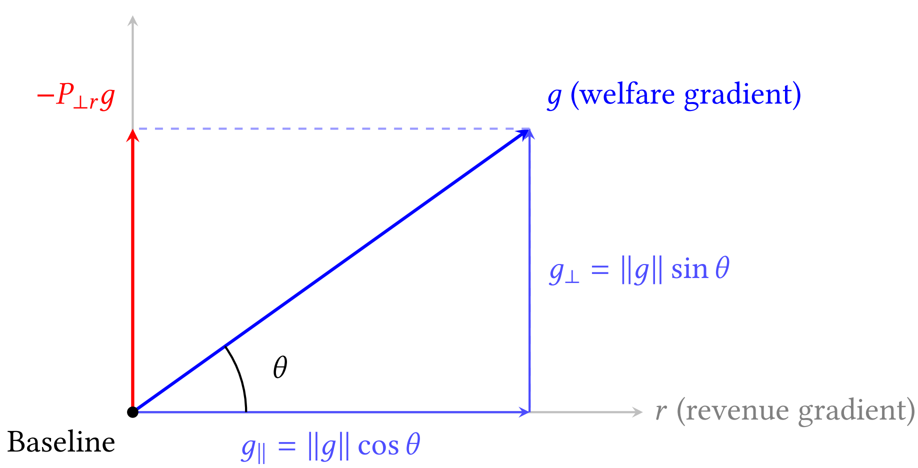
$$d\boldsymbol{\tau}^* \propto -P_{\perp r} g, \tag{8}$$

where $P_{\perp r} := I - r(r^\top r)^{-1} r^\top$ is the projection operator onto the hyperplane orthogonal to r.

The derivation is in Appendix A.2. Intuitively, the projection operator $P_{\perp r}$ decomposes the welfare gradient g into two orthogonal components: a parallel component that points along the revenue gradient r (which would change revenue) and a perpendicular component that lies in the revenue-neutral hyperplane. The optimal reform follows this perpendicular component, identifying the direction of steepest welfare improvement among all revenue-neutral reforms.

Figure 2 illustrates this decomposition. The welfare gradient g (blue) splits into a parallel component $g_{\|} = \|g\| \cos\theta$ along the revenue gradient r and an orthogonal component $g_{\perp} = \|g\| \sin\theta$ perpendicular to r. If we could move freely along g—ignoring the revenue constraint—we would achieve the steepest welfare improvement. However, moving along g would change total revenue, violating revenue-neutrality. The projection operator resolves this tension by extracting the perpendicular component, yielding $P_{\perp r} g$ (red) as the optimal revenue-neutral reform direction.

Figure 2: Geometric Decomposition of Optimal Policy



Note: The welfare gradient g (blue) decomposes into a parallel component $g_{\|} = \|g\| \cos\theta$ along the revenue gradient r and an orthogonal component $g_{\perp} = \|g\| \sin\theta$ perpendicular to r. The optimal revenue-neutral reform direction is $-P_{\perp r} g$ (red), which points opposite to the perpendicular component of welfare cost.

2.3 Alignment Metrics

While the projection formula is a classical result, this section develops scalar metrics that measure how far the tax system is from optimal and how much welfare improvement is available through revenue-neutral reform. The geometry in Figure 2 depends entirely on the angle θ between the welfare and revenue gradients. This angle determines both proximity to the Ramsey optimum and the scope for revenue-neutral reform. Define the system alignment as:

$$\cos\theta = \frac{g^\top r}{\|g\|\,\|r\|} \in [-1, 1]. \tag{9}$$

This is the standard inner product formula for the angle between vectors. Geometrically, $\cos\theta$ measures the correlation between welfare costs and revenue changes.

Proposition 1 (System alignment and revenue-neutral potential)**.** *The efficiency of the tax system is characterized by two scalar metrics:*

1. ***System alignment***: $\cos\theta$ *measures proximity to the Ramsey optimum. When* $\cos\theta = 1$*, welfare and revenue objectives are perfectly aligned and the system is optimal. When* $\cos\theta = 0$*, the gradients are orthogonal and inefficiency stems purely from tax composition. When* $\cos\theta < 0$*, raising taxes increases welfare cost but reduces revenue.*
2. ***Revenue-neutral potential***: $\sin^2\theta = 1 - \cos^2\theta$ *measures the share of potential welfare gains achievable through revenue-neutral reform. The first-order welfare gain from a reform of size* α *in the optimal direction is*

$$\Delta W \approx -\alpha\|g\|^2 \sin^2\theta. \tag{10}$$

Proof. From Lemma 1, the first-order welfare change is $\Delta W = -\alpha g^\top P_{\perp r} g = -\alpha\|P_{\perp r} g\|^2$. The squared norm decomposes as $\|P_{\perp r} g\|^2 = \|g\|^2 - (g^\top r)^2/\|r\|^2 = \|g\|^2(1 - \cos^2\theta) = \|g\|^2 \sin^2\theta$. □

These scalar metrics are new to the literature. While Tirole and Guesnerie (1981) develop methods for computing reform paths, they provide no summary statistics for measuring system efficiency or evaluating observed reforms.

The revenue-neutral potential $\sin^2\theta$ is maximized when the gradients are orthogonal ($\theta = 90°$) and declines symmetrically toward zero as θ approaches either $0°$ or $180°$. This reflects a fundamental decomposition: the welfare gradient g splits into a component parallel to the revenue gradient (which revenue-neutral reform cannot address) and a perpendicular component (which it can). Whether the system is moderately misaligned on the right side ($\theta = 60°$) or the wrong side ($\theta = 120°$) of the Laffer curve, the perpendicular component—and thus revenue-neutral potential—has the same magnitude.

The sign and magnitude of $\cos\theta$ characterize the tax system's efficiency. When $\cos\theta = 1$, the gradients are collinear: this is the Ramsey optimum, where MCPFs are equalized and no revenue-neutral reform can improve welfare. When $\cos\theta = 0$, the gradients are orthogonal: inefficiency stems purely from tax composition, and all welfare improvement is available without fiscal cost ($\sin^2\theta = 1$). Intermediate values indicate partial misalignment. For example, $\cos\theta = 0.5$ implies $\sin^2\theta = 0.75$—three-quarters of potential welfare gains are achievable revenue-neutrally. Negative values ($\cos\theta < 0$) indicate that raising taxes increases welfare cost but reduces total revenue; this occurs when cross-base spillovers dominate own-revenue effects. Appendix Table A.1 summarizes.

Beyond measuring system efficiency, the framework evaluates whether historical reforms moved in the right direction. For an observed reform $d\boldsymbol{\tau}^{\text{actual}}$, we compute its directional alignment with the optimal reform $d\boldsymbol{\tau}^*$:

$$\cos\phi = \frac{(d\boldsymbol{\tau}^{\text{actual}})^\top d\boldsymbol{\tau}^*}{\|d\boldsymbol{\tau}^{\text{actual}}\|\ \|d\boldsymbol{\tau}^*\|}. \tag{11}$$

This statistic inherits the same interpretation as $\cos\theta$. When $\phi \approx 0^\circ$ ($\cos\phi \approx 1$), the actual reform moved nearly parallel to the optimal direction. It was well-designed given the economy's position. When $\phi \approx 90^\circ$ ($\cos\phi \approx 0$), the reform was orthogonal, neither helping nor hurting. When $\phi > 90^\circ$ ($\cos\phi < 0$), the reform moved opposite to the welfare-improving direction, making the system less efficient. Together, $\cos\theta$ (system position) and $\cos\phi$ (reform direction) provide a complete characterization of tax policy.

The projection formula, while derived geometrically, enforces the classic intuition of MCPF equalization. Rewriting equation (8) component-wise:

$$d\tau_i^* \propto -g_i + \left(\frac{r^\top g}{r^\top r}\right) r_i = -r_i \left(\frac{g_i}{r_i} - \frac{r^\top g}{r^\top r}\right).$$

Define $\text{MCPF}_i = g_i/r_i$ and the revenue-weighted average $\overline{\text{MCPF}} = (r^\top g)/(r^\top r)$:

Corollary 1 (MCPF equalization)**.** *The optimal reform direction satisfies*

$$d\tau_i^* \propto -r_i(MCPF_i - \overline{MCPF}). \tag{12}$$

The projection automatically lowers taxes on instruments with above-average MCPFs and raises taxes on instruments with below-average MCPFs. MCPF equalization emerges as a consequence of the geometric projection.

2.4 A Two-Instrument Example

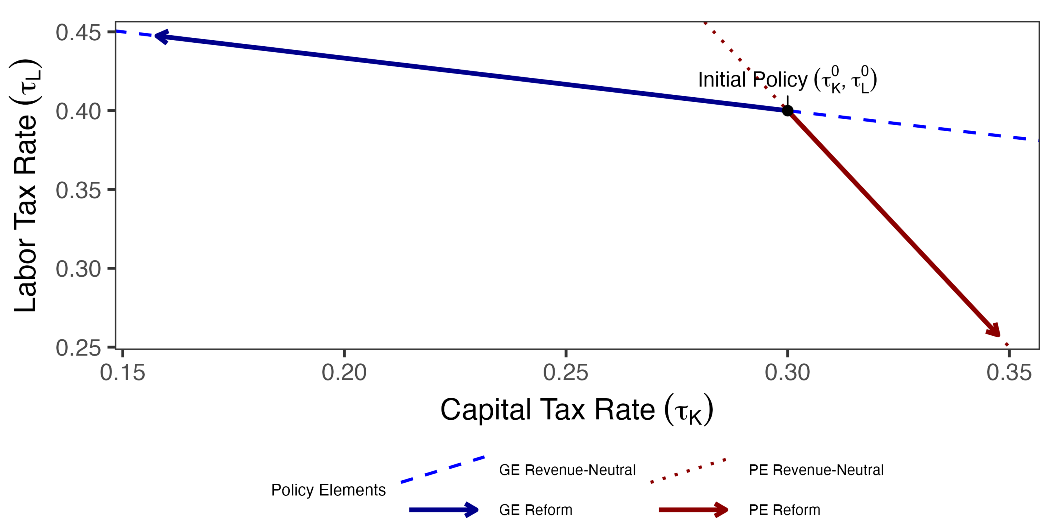
The alignment metrics depend on the revenue gradient r, which aggregates both own-base responses and cross-base spillovers. What happens if a planner ignores the spillovers and uses a partial equilibrium revenue gradient r^{PE} instead? This subsection shows that the error can be severe: partial equilibrium analysis not only mismeasures system alignment but can prescribe reforms in the opposite direction from the true optimum.

Consider an economy with two tax instruments: a tax on labor income and a tax on capital income. Output is produced with a concave production function $Y = F(K, L)$, so the tax bases are $B_K = F_K K$ and $B_L = F_L L$. If capital and labor are complements in production ($F_{KL} > 0$), a higher capital tax reduces the capital stock, which reduces wages and shrinks the labor tax base. This spillover means the true revenue gradient for capital taxes is smaller than the PE estimate:

$$r_K^{\mathrm{GE}} = B_K \left(1 - \frac{\tau_K}{1-\tau_K}\varepsilon_{KK} - \Lambda_K\right) < r_K^{\mathrm{PE}} = B_K \left(1 - \frac{\tau_K}{1-\tau_K}\varepsilon_{KK}\right),$$

where $\Lambda_K \equiv \frac{\tau_L B_L}{B_K(1-\tau_K)}\varepsilon_{LK} > 0$ captures the cross-base spillover. The PE planner thinks capital taxes raise more revenue than they actually do.

Figure 3: The Geometric Failure of Partial Equilibrium



Note: Starting from baseline policy (black dot), the partial equilibrium reform (red arrow) projects onto the PE revenue-neutral line (red dashed) and prescribes raising capital taxes. The general equilibrium reform (blue arrow) projects onto the correct revenue-neutral line (blue dashed) and prescribes lowering capital taxes—the opposite direction.

Figure 3 illustrates the consequences. The GE revenue-neutral line (blue) and PE revenue-neutral line (red) have different slopes because PE ignores spillovers. The projection formula tells the planner to project the welfare gradient onto the revenue-neutral line—but the PE planner projects onto the wrong line. The result: GE prescribes lowering capital taxes and raising labor taxes, while PE prescribes the opposite. When cross-base spillovers are large, partial equilibrium analysis can reverse the optimal reform direction entirely.

The example illustrates a general point: the alignment metric $\cos\theta$ is only as good as the revenue gradient used to compute it. A partial equilibrium gradient r^{PE} yields a different angle θ^{PE}, potentially indicating alignment when the system is actually misaligned (or vice versa).

3 An Extension to Dynamics

The static framework in Section 2.2 treats tax changes as one-time adjustments with contemporaneous effects. In reality, tax policy has persistent consequences. A corporate tax cut today raises investment, growing the capital stock and expanding future labor bases through production complementarity. A labor tax increase discourages human capital investment, shrinking future tax bases. These intertemporal spillovers are first-order determinants of the marginal cost of public funds, yet they are invisible in static analysis. This section shows that the projection framework extends naturally to dynamic settings: time is simply additional dimensions in the policy vector, and present-value gradients aggregate all intertemporal spillovers.

3.1 Dynamic Reforms

Dynamic setup. Let the government choose tax rates $\{\tau_{i,t}\}_{i=1,\ldots,n}^{t=0,\ldots,T}$ across n instruments and $T+1$ periods. Stack these into a policy vector:

$$\boldsymbol{\tau} := (\tau_{1,0}, \ldots, \tau_{n,0}, \tau_{1,1}, \ldots, \tau_{n,1}, \ldots, \tau_{1,T}, \ldots, \tau_{n,T})^\top \in \mathbb{R}^{n(T+1)}.$$

Discounting. I discount both welfare and revenue at the social discount factor $\beta = 1/(1+\rho)$, where ρ is society's rate of time preference. This assumes the government's borrowing rate equals the social discount rate, $r_{\text{debt}} = \rho$. When these diverge, the projection formula requires dual discounting as developed in Appendix G.

Tax bases at each date depend on the entire policy sequence: $B_{i,t}(\boldsymbol{\tau})$. A change in $\tau_{i,t}$ affects not only base i at date t (the contemporaneous own-base effect) but also other bases at other dates through intertemporal spillovers. Present-value revenue is:

$$R^{\mathrm{PV}}(\boldsymbol{\tau}) := \sum_{t=0}^{T} \beta^t R_t(\boldsymbol{\tau}) = \sum_{t=0}^{T} \beta^t \sum_{i=1}^{n} \tau_{i,t} B_{i,t}(\boldsymbol{\tau}). \tag{13}$$

Present-value revenue gradient. The present-value revenue gradient $r^{\text{PV}} := \nabla_{\boldsymbol{\tau}} R^{\text{PV}} \in \mathbb{R}^{n(T+1)}$ can be derived using the same decomposition as the static case. To see this, introduce stacked notation. Let $a \equiv (i,t)$ and $b \equiv (k,s)$ denote (instrument, date) pairs. Define discounted bases and tax rates

$$\tilde{B}_a := \beta^t B_{i,t}, \qquad \tau_a := \tau_{i,t},$$

and the present-value GE elasticity matrix $\varepsilon_{ba} := \frac{\partial \ln B_{k,s}}{\partial \ln(1-\tau_{i,t})}$ stacked over all (k,s) and (i,t). Then from $R^{\text{PV}} = \sum_{s=0}^{T} \beta^s \sum_k \tau_{k,s} B_{k,s}$ it follows that

$$\frac{\partial R^{\text{PV}}}{\partial \tau_{i,t}} = \beta^t B_{i,t} + \sum_{s,k} \beta^s \tau_{k,s} \frac{\partial B_{k,s}}{\partial \tau_{i,t}} = \tilde{B}_a - \frac{1}{1-\tau_a} \sum_b \tau_b \tilde{B}_b \varepsilon_{ba}.$$

Separating the own term $b = a$ from the rest gives the exact analogue of the static formula:

$$r_{i,t}^{\text{PV}} = \tilde{B}_{i,t} \left[1 - \frac{\tau_{i,t}}{1-\tau_{i,t}} \varepsilon_{(i,t),(i,t)} - \Lambda_{i,t}^{\text{PV}} \right] \tag{14}$$

The present-value spillover term $\Lambda_{i,t}^{\text{PV}}$ aggregates how tax instrument i at time t affects all other tax bases across all future periods—capturing fiscal externalities that compound over time through capital accumulation, human capital investment, and other dynamic adjustments. This is the core empirical challenge. Most existing estimates of the marginal cost of public funds are partial equilibrium: they use only the contemporaneous own-base elasticity ε_{ii} and ignore both cross-base spillovers and intertemporal persistence (Finkelstein and Hendren 2020; Hendren and Sprung-Keyser 2020). This can severely bias MCPF estimates. For instance, if capital and labor are complements, a corporate tax increase erodes the labor base over time, making the true revenue gradient much smaller than the partial equilibrium calculation suggests. Section 4 shows how to estimate $\Lambda_{i,t}^{\text{PV}}$ directly from the cross-base impulse responses in a structural VAR, capturing general equilibrium spillovers without calibrating a specific production function.

The present-value revenue gradient $r_{i,t}^{\text{PV}}$ admits two complementary decompositions. Equation (14) provides an *economic* decomposition into mechanical effect ($\tilde{B}_{i,t}$), contemporaneous own-base response ($\varepsilon_{(i,t),(i,t)}$), and all intertemporal spillovers ($\Lambda_{i,t}^{\text{PV}}$). This parallels the static formula exactly—the only difference is that spillovers now include effects across time as well as across instruments.

Alternatively, a *temporal* decomposition splits by when revenue effects occur:

$$r_{i,t}^{\text{PV}} = \underbrace{\beta^t \frac{\partial R_t}{\partial \tau_{i,t}}}_{\text{Contemporaneous}} + \underbrace{\sum_{s<t} \beta^s \frac{\partial R_s}{\partial \tau_{i,t}}}_{\text{Anticipation}} + \underbrace{\sum_{s>t} \beta^s \frac{\partial R_s}{\partial \tau_{i,t}}}_{\text{Persistence}}. \tag{15}$$

The *contemporaneous* term is the static revenue effect at date t—what partial equilibrium analysis captures. The *anticipation* and *persistence* terms capture intertemporal spillovers: how tax changes at time t affect revenue at other dates $s \neq t$. These correspond to upper-diagonal (anticipation) and lower-diagonal (persistence) entries in the Jacobian matrix $\partial R_s / \partial \tau_{i,t}$. Both anticipation and persistence effects are embedded in the spillover term $\Lambda_{i,t}^{\text{PV}}$, which aggregates all off-diagonal elasticities $\varepsilon_{(k,s),(i,t)}$ for $(k,s) \neq (i,t)$. To illustrate the economic channels underlying these spillovers, we consider three examples:

Example 1: Human capital accumulation (lower-diagonal, own-base). Consider an increase in the labor tax at $t = 0$ whose effects persist through subsequent periods. Workers respond by reducing human capital investment (education, training), as the after-tax return to skill accumulation has fallen. This shrinks the labor tax base $B_{L,t}$ in all future periods:

$$\frac{\partial B_{L,t}}{\partial \tau_{L,0}} < 0 \quad \text{for all } t > 0. \tag{16}$$

The static MCPF, which uses only the contemporaneous elasticity ε_{LL} (the short-run labor supply response), misses these persistent effects. Empirical estimates of ε_{LL} from contemporaneous variation typically yield small values around 0.2 (Chetty et al. 2011). But Badel, Huggett, and Luo (2020) and Kleven et al. (2025) show that accounting for human capital investment raises the present-value elasticity. The true $r_{L,0}^{\text{PV}}$ is much smaller than the static calculation suggests, making the labor tax more costly than it appears in static partial equilibrium:

$$r_{L,0}^{\text{PV}} = \beta^0 \frac{\partial R_0}{\partial \tau_{L,0}} + \sum_{t=1}^{T} \beta^t \frac{\partial R_t}{\partial \tau_{L,0}} \ll \frac{\partial R_0}{\partial \tau_{L,0}}.$$

Example 2: Capital accumulation (lower-diagonal, cross-base). A corporate tax increase at $\tau_{K,0}$ reduces investment, shrinking the capital stock over time. If capital and labor are complements in production ($F_{KL} > 0$), the reduced capital stock lowers the marginal product of labor (wages), eroding the labor tax base in future periods:

$$\frac{\partial B_{L,t}}{\partial \tau_{K,0}} < 0 \quad \text{for all } t > 0.$$

This cross-base, intertemporal spillover compounds the static fiscal externality from Section 2.4. The present-value revenue gradient for the capital tax includes the cumulative effect on labor revenue:

$$r_{K,0}^{\text{PV}} = \beta^0 \frac{\partial R_0}{\partial \tau_{K,0}} + \sum_{t=1}^{T} \beta^t \left(\frac{\partial R_{K,t}}{\partial \tau_{K,0}} + \frac{\partial R_{L,t}}{\partial \tau_{K,0}} \right),$$

where the second term in the sum captures the cross-base spillover. Vergara and Swonder (2025) account for the cross-partial statically, but not dynamically in their analysis of TCJA.

Example 3: Anticipation effects (upper-diagonal). Forward-looking agents adjust behavior today in response to announced future reforms. Suppose the government announces at $t = 0$ that the capital tax will be cut at $t = 5$: $\tau_{K,5}$ falls, but $\tau_{K,t<5}$ remains unchanged. Firms delay investment from periods $t < 5$ to period $t = 5$ and beyond, when the after-tax return is higher. This shrinks the capital tax base before the reform is implemented:

$$\frac{\partial B_{K,t}}{\partial \tau_{K,5}} \neq 0 \quad \text{for } t < 5.$$

These anticipation effects—upper-diagonal entries in the Jacobian—mean that the present-value revenue cost of the reform is larger than the static calculation at $t = 5$ would suggest. Mertens and Ravn (2012) exploit announcement timing to estimate these effects directly, distinguishing announcement dates from implementation dates in their narrative tax shock series. Their impulse responses trace out $\frac{\partial R_s}{\partial \tau_{i,t}}$ for all s (including $s < t$), providing reduced-form estimates of the upper-diagonal entries.

Welfare and the projection formula. The present-value welfare is defined analogously, using the social discount factor:

$$W^{\text{PV}}(\boldsymbol{\tau}) := \sum_{t=0}^{T} \beta^t W_t(\boldsymbol{\tau}), \tag{17}$$

with welfare gradient $g^{\text{PV}} := \nabla_{\boldsymbol{\tau}} W^{\text{PV}} \in \mathbb{R}^{n(T+1)}$.

Having stacked the dynamic problem into a single policy vector in $\mathbb{R}^{n(T+1)}$, the projection machinery from Section 2.2 applies immediately. The revenue-neutral hyperplane is now defined by $r^{\text{PV}\top} d\boldsymbol{\tau} = 0$, and the optimal reform projects the welfare gradient onto this hyperplane.

Corollary 2. *At an interior solution, the government's optimal revenue-neutral reform is*

$$d\boldsymbol{\tau}^* = -P_{\perp r^{PV}} g^{PV}, \tag{18}$$

where $P_{\perp r^{PV}} = I - r^{PV}(r^{PV\top} r^{PV})^{-1} r^{PV\top}$ *projects onto the hyperplane orthogonal to the present-value revenue gradient* r^{PV}*.*

The projection formula is identical to Lemma 1. The policy space is $\mathbb{R}^{n(T+1)}$ instead of $\mathbb{R}^n$, but the geometric logic is unchanged. Time is just additional dimensions. Cross-time spillovers—captured by off-diagonal entries $\varepsilon_{(k,s),(i,t)}$ for $s \neq t$—are analogous to cross-base spillovers in the static case.

3.2 Debt Finance and the Scope of the Analysis

Real-world fiscal reforms rarely occur under strict revenue neutrality. The Tax Cuts and Jobs Act of 2017 reduced taxes by approximately \$1.5 trillion over ten years without offsetting revenue increases, expanding the federal deficit (JCT 2017). How should we evaluate such deficit-financed reforms?

The alignment metric $\cos\theta$ answers a narrow question: holding total revenue fixed, is the *composition* of taxes efficient? It measures whether revenue is raised from the right mix of instruments, not whether the level of revenue is appropriate. A tax system can have $\cos\theta \approx 1$ (near-optimal composition) while running large deficits or collecting too little revenue. The metric is silent on these questions. When I find that TCJA moved in the optimal direction ($\cos\phi > 0$), this means it improved the corporate-personal mix—not that its deficit financing was justified.

Evaluating deficit financing requires a separate analysis. Issuing government debt is equivalent to choosing the timing of taxation. The transversality condition ensures all debt must eventually be repaid:

$$\lim_{T\to\infty} \beta^T D_T = 0 \quad \Rightarrow \quad D_0 = \sum_{t=0}^{\infty} \left(\frac{1}{1+r_{\text{debt}}}\right)^t [G_t - R_t], \tag{19}$$

where r_{debt} is the government's borrowing rate. Debt does not eliminate the need for taxation; it defers it. The question "should we issue more debt?" is therefore equivalent to "should we lower taxes today and raise them tomorrow?"

The dynamic projection framework developed in this section is general: it nests Chamley's (1986) zero capital tax theorem, Barro's (1979) tax smoothing result, and Ramsey's (1927) optimality condition as special cases when present-value revenue gradients take particular limiting forms. Appendix H develops these connections. But the empirical question does not require solving for the global optimum. It asks: given the tax system we observe, how far is it from optimal, and did historical reforms move in the right direction? The following section estimates the present-value gradients needed to answer this question.

Answering this requires comparing the MCPF of immediate revenue-neutral financing to the present-value MCPF of future repayment. However, the actual repayment path is unobserved for historical reforms and depends on future political choices. Evaluating the optimality of deficit timing is therefore a distinct question from evaluating the optimality of tax composition—just as optimal government spending is distinct from optimal taxation. This paper focuses on composition; Appendix G extends the framework to joint evaluation of composition and timing.

4 Estimation of General Equilibrium Dynamic Elasticities

As an application of the geometric framework, I examine whether the postwar decline in corporate tax burdens relative to personal tax burdens moved the U.S. tax system toward or away from efficiency. Figure 1 shows that the corporate share of federal revenue fell from nearly 50% in the late 1940s to around 10% by 2019, while the personal income tax share grew correspondingly. This shift raises a natural question: did the rebalancing improve welfare, or would revenue-neutral reforms in the opposite direction—higher corporate taxes, lower personal taxes—have been preferable?

The projection framework answers this question by characterizing the alignment between welfare and revenue objectives. The alignment metric $\cos\theta$ measures how close the tax system is to the Ramsey optimum, while $\sin^2\theta$ measures the share of potential welfare gains achievable through revenue-neutral rebalancing. If welfare and revenue gradients point in similar directions ($\cos\theta$ near one), the system is near-optimal and the postwar shift may have been efficient. If the gradients are misaligned ($\cos\theta$ near zero or negative), substantial welfare gains were available, and whether the observed reforms moved toward or away from the optimum depends on the sign of their directional alignment with the optimal reform. In this section, I assemble the ingredients for estimating these alignment statistics—the revenue and welfare gradients for personal and corporate income taxes—using narratively-identified VARs. Section 5 then applies these estimates to evaluate historical reforms.

4.1 Identification and Estimation

The key empirical object is the present-value revenue gradient $r_j^{PV} = \sum_{h=0}^{\infty} \beta^h \frac{\partial R_{t+h}}{\partial \tau_{j,t}}$. This captures how an exogenous change in tax rate j at date t affects total government revenue at all future dates, accounting for dynamic adjustments and general equilibrium spillovers across tax bases. Estimating this gradient requires identifying exogenous variation in tax rates and tracing out the dynamic response of tax bases to tax shocks at all horizons. Concretely, I estimate impulse responses of log tax bases to shocks in log retention rates $\ln(1-\tau_j)$, where τ_j is the average tax rate on instrument j. Cumulating and discounting these IRFs yields the present-value elasticities that enter the gradient formula.[6]

6. The empirical analysis uses average tax rates rather than statutory marginal rates for three reasons. First, the theoretical framework in Section 2.1 defines revenue as $R_i = \tau_i B_i$, where τ_i is the average rate. The revenue gradient $\partial R/\partial\tau$ therefore requires variation in average rates. Second, the U.S. tax code features graduated rate schedules with rates varying across income brackets, types of income, and categories of taxpayers, so no single statutory rate summarizes the policy stance. The average rate—total revenue divided by the base—provides the natural aggregate measure. Third, many historical reforms changed the effective tax burden without altering statutory rates at all. Base-broadening measures such as limiting deductions, accelerating depreciation schedules, or restricting credits

The welfare gradient also follows from the impulse responses. Under utilitarian welfare weights, the first-order welfare cost of a tax increase equals the mechanical burden on taxpayers—the tax base—discounted over the horizon. Unlike the revenue gradient, the welfare gradient does not depend on the persistence of the identified shock: we evaluate the welfare cost of a permanent marginal reform using elasticities identified from temporary shocks. The tax rate IRFs nonetheless play a key role: their persistence $\kappa_j = \sum_{h=0}^{H} \beta^h \cdot \mathrm{IRF}_h(\ln(1-\tau_j))$ enters the revenue gradient formula and normalizes the base responses into permanent-equivalent elasticities. The following subsection provides the precise formulas.

The narrative approach provides the exogenous variation needed to estimate revenue gradients. I follow the identification strategy developed by Romer and Romer (2010) and extended by Mertens and Ravn (2013a) to distinguish personal and corporate income taxes. Romer and Romer (2010) read the Congressional Record and Treasury documents to classify major federal tax legislation from 1945 to 2007 according to the motivations policymakers articulated at the time. They code a reform as exogenous when the legislative record indicates policymakers pursued objectives unrelated to the contemporaneous business cycle—such as long-run deficit reduction, ideological commitments, or tax simplification—rather than short-run stabilization. Mertens and Ravn (2013a) refine the series by separating personal and corporate tax changes and excluding reforms implemented more than 90 days after announcement to avoid anticipation effects. I extend their series through 2019 using the classification in Cloyne et al. (2022).

Each reform receives a quantitative measure equal to the projected revenue impact scaled by the lagged tax base, as announced at the time of enactment. This scaling delivers an approximation to the change in the average tax rate on that instrument. The resulting series are denoted m_t^{PI} for personal income taxes and m_t^{CI} for corporate income taxes.

The narrative measures serve as instruments for latent structural tax shocks rather than entering as regressors directly. As discussed in Mertens and Ravn (2013a), the two instruments are contemporaneously correlated because Congress often adjusts personal and corporate taxes in the same legislation. This correlation means the instruments identify only the two-dimensional subspace spanned by the structural shocks, not the individual shocks themselves. Additional restrictions are required to separate them.[7] Following Mertens and Ravn (2013a) and Cloyne et al. (2022), I use the narrative measures as external instruments to identify the contemporaneous impact matrix A_0. Let Z_t denote the vector of endogenous variables (tax rates, tax bases, GDP, government spending, and debt). A reduced-form VAR produces residuals u_t, which relate to

raise the average rate even when the statutory schedule remains unchanged. The revenue-based narrative measure captures these reforms, which statutory rate series would miss.

7. Cloyne et al. (2022) shows that both VAR-identified instruments are strong.

structural shocks ε_t via:

$$u_t = A_0 \varepsilon_t$$

where ε_t are mutually orthogonal structural shocks with unit variance and A_0 is the contemporaneous impact matrix. The external instruments identify the columns of A_0 corresponding to tax shocks through the moment conditions:

$$E[m_t \varepsilon'_{tax,t}] = \Phi \neq 0, \quad E[m_t \varepsilon'_{other,t}] = 0$$

where $m_t = (m_t^{PI}, m_t^{CI})'$, $\varepsilon_{tax,t}$ denotes the two structural tax shocks, and $\varepsilon_{other,t}$ denotes all other structural disturbances. The first condition requires the instruments to correlate with the tax innovations; the second requires them to be uncorrelated with non-tax shocks.

The two narrative instruments are correlated, so the proxy-SVAR identifies a two-dimensional subspace of tax shocks but does not uniquely separate personal from corporate shocks within this subspace. To isolate each shock, I apply a Cholesky decomposition to the identified subspace, ordering the shock of interest last. When estimating responses to the personal tax shock, I order the personal tax rate last, imposing that the personal shock has no contemporaneous effect on the corporate rate; when estimating responses to the corporate shock, I reverse the ordering. This approach does not require a single structural model that simultaneously satisfies both restrictions. Each column of the elasticity matrix is estimated from a separate single-shock experiment, and the alignment metric requires only the individual revenue gradients—not a joint structural interpretation of both shocks.

I estimate impulse responses using the two-stage Bayesian local projection framework developed by Cloyne et al. (2022). The first stage estimates a Bayesian VAR with hierarchical priors following Giannone, Lenza, and Primiceri (2015). Hyperparameters governing overall shrinkage are treated as random and sampled via Metropolis-Hastings. For each draw from the BVAR posterior, I apply the identification procedure described above to recover a draw of A_0. This delivers the causal impact of each tax shock on all variables in the system at horizon zero.

The second stage estimates local projections at each horizon $h \geq 1$:

$$Z_{t+h} = c^{(h)} + B_1^{(h)} Z_{t-1} + \sum_{j=2}^{p} b_j^{(h)} Z_{t-j} + u_{t+h}, \quad \text{var}(u_{t+h}) = \Omega_h \tag{20}$$

where $p = 4$ lags. At horizon zero, the residuals u_t relate to the structural shocks via $u_t = A_0 \varepsilon_t$. The impulse response at horizon h is $IRF_h = B_1^{(h-1)} A_0$. I estimate equation (20) with Bayesian methods using a Minnesota prior and Student-t errors to accommodate fat tails. Appendix C provides complete details on prior specification and the MCMC algorithm.

Bayesian estimation offers three advantages: it propagates identification uncertainty from A_0 into the impulse response posterior, it improves precision at long horizons where LP variance is high (Li, Plagborg-Møller, and Wolf 2024), and it produces posterior draws for straightforward inference on the alignment metrics. I use local projections rather than propagating shocks through a VAR because the present-value gradient requires reliable estimates over extended horizons, where VARs can suffer from misspecification bias due to lag truncation. However, I also show that the main results are robust to a variety of alternative estimation methods including standard local projections, penalized local projections, smooth local projections, Bayesian VARs, and proxy SVARs.

Data The baseline specification includes four variables:

$$Z_t = \Big[\log(1 - \text{APITR}_t), \log(1 - \text{ACITR}_t), \ln B_t^{PI}, \ln B_t^{CI} \Big],$$

where APITR_t and ACITR_t are the average personal and corporate income tax rates, and B_t^{PI} and B_t^{CI} are the corresponding tax bases. The sample covers 1947Q1–2019Q4. Following standard practice in this literature, the corporate base is BEA corporate profits, which includes both C-corporations and S-corporations; the implications of this measurement choice for interpreting the cross-elasticity are discussed in Section 4.2. Appendix B provides detailed variable definitions and data sources.

The revenue gradient requires the total derivative of tax bases with respect to tax rates: the full response a policymaker would observe, including effects operating through output, investment, and wages. Adding variables like GDP or government spending would partial out these channels, estimating a conditional elasticity rather than the total elasticity the theory requires. Narrative identification does not require additional controls; exogeneity comes from the Romer-Romer selection criterion, not from the conditioning set. Appendix F develops this argument in detail and shows that a seven-variable specification following Mertens and Ravn (2013a) yields qualitatively similar but less precise estimates.

The impulse responses are estimated using the raw (unfiltered) data, which is standard in the VAR literature. However, for the alignment calculations in Section 5, I extract the long-run component of tax rates and bases using the Christiano and Fitzgerald (2003) band-pass filter, retaining only frequencies longer than 40 quarters (10 years). This filtering serves two purposes. First, it removes business-cycle variation that is orthogonal to the permanent reform experiments the framework evaluates—the alignment metric measures efficiency of the long-run tax structure, not cyclical fluctuations in bases. Second, it smooths measurement error in quarterly tax data, which can be noisy due to timing of payments and refunds.

4.2 Impulse Response Functions and Gradients

Figure E.6 plots discounted cumulative impulse responses to personal tax shocks (left column) and corporate tax shocks (right column). Each panel shows the running sum $\sum_{s=0}^{h} \beta^s \cdot \mathrm{IRF}_s$ as the horizon h extends from 0 to 20 quarters; the terminal value is the present-value response that enters the gradient formulas. Rows display own retention rates (top), corporate base (middle), and personal base (bottom).

Three patterns emerge. First, corporate tax shocks are more persistent than personal tax shocks: the cumulative retention rate response reaches approximately 8 percentage points for corporate taxes but plateaus around 6 for personal taxes. Second, cross-base spillovers are asymmetric. Personal tax cuts substantially expand the corporate base (middle left), with cumulative responses exceeding 100 percentage points, while corporate tax cuts have slightly negative effects on the personal base (bottom right). Third, both own-base responses are positive: corporate tax cuts expand the corporate base by around 60 percentage points (middle right) and personal tax cuts expand the personal base by around 9 percentage points (bottom left), though the corporate base responds more strongly to both shocks.

These cumulative responses map into the gradient formulas. The revenue gradient is:

$$r_j^{PV} = B_{j,0} \cdot \left[1 - \frac{\tau_j}{1-\tau_j} \varepsilon_{jj} - \frac{\tau_k}{1-\tau_j} \frac{B_k}{B_j} \varepsilon_{kj} \right], \tag{21}$$

where each elasticity ε_{ij} equals the terminal value in the corresponding base panel of Figure E.6 divided by the cumulative retention rate response κ_j. Dividing by persistence converts present-value responses into permanent-equivalent elasticities, ensuring both the revenue and welfare gradients correspond to the same policy experiment: a permanent marginal reform. The welfare gradient is simply:

$$g_j^{PV} = B_{j,0}, \tag{22}$$

since the common annuity factor cancels when computing alignment metrics. The baseline uses $\beta = 0.9926$ (approximately 3% annual discount rate) and $H = 20$ quarters, following the five-year horizon in Mertens and Montiel Olea (2018). Table 1 reports the baseline elasticities in a 2×2 matrix. The key finding is asymmetric cross-base spillovers: the cross-elasticity $\varepsilon_{CL} \approx 21$ is large and positive, while $\varepsilon_{LC} \approx -0.3$ is small and negative. Personal tax cuts substantially expand the corporate base; corporate tax cuts have negligible effects on the personal base. Both own-elasticities are positive: $\varepsilon_{CC} \approx 8.5$ and $\varepsilon_{LL} \approx 1.5$.

The magnitude of ε_{CL} warrants comment. As noted in Section 4.1, the corporate base includes S-corporation profits, which are taxed at personal rates. When personal taxes fall, S-corporation activity expands, and this response appears in the measured corporate base. The cross-elasticity therefore captures both traditional general equilibrium channels (demand effects, labor-capital complementarity) and the direct sensitivity of pass-through business activity to personal tax incentives. The latter channel is quantitatively important: S-corporations grew from a negligible share of business income in the 1970s to roughly half by the 2010s (Dyrda and Pugsley 2024; Smith et al. 2022). For the alignment calculation, what matters is how measured bases respond to the policy instruments Congress actually uses. Personal tax legislation affects S-corporation incentives; this fiscal externality belongs in the revenue gradient regardless of how we classify S-corps conceptually.

These elasticities are estimated once from the full sample and held fixed across time. Time variation in the alignment metric $\cos\theta$ comes entirely from changes in tax rates (τ_L, τ_C) and the relative size of the two bases (B_C/B_L), which I measure using the long-run (Christiano-Fitzgerald filtered) components of the data. The revenue gradient formula (21) combines fixed elasticities with these time-varying bases and rates to produce a quarterly series of alignment statistics.

Table 1: Present-Value Elasticities of Tax Bases

	Personal Tax Shock	Corporate Tax Shock
$\varepsilon_{L\cdot}$ (Personal Base)	1.52	-0.34
	(0.39, 2.65)	(-0.83, 0.07)
$\varepsilon_{C\cdot}$ (Corporate Base)	20.88	8.47
	(14.53, 29.60)	(5.85, 13.54)

Note: Elasticities from Bayesian Local Projections with narrative tax shocks. 95% credible intervals in parentheses. Each ε_{ij} measures the percent change in base i per percent change in retention rate $(1-\tau_j)$.

Table E.1 in Appendix E shows robustness across estimation methods. The local projection variants—standard LP, smooth LP (Barnichon and Brownlees 2019), and bias-corrected LP (Herbst and Johannsen 2024)—yield similar point estimates with wider confidence intervals. The VAR-based methods produce noisier estimates, as expected given that local projections are better suited to long horizons.

These elasticities differ conceptually from the taxable income elasticities in the public finance literature. Standard estimates—from bunching at kinks, difference-in-differences around state tax changes, or tax return panel data—identify individual responses to *marginal* rates, holding wages and prices fixed. The elasticities here capture aggregate base responses to *average* rates, incorporating general equilibrium adjustments in wages, interest rates, and organizational form. They are also cumulative over a five-year horizon, capturing dynamic responses through investment that static estimates miss. The two objects answer different questions: the micro elasticity asks how one taxpayer responds to their own marginal rate; the macro elasticity asks how total revenue responds to an aggregate tax change.

Saez, Slemrod, and Giertz (2012) survey the taxable income elasticity literature and conclude that the best available estimates range from 0.12 to 0.40. Mertens and Montiel Olea (2018), using narrative identification similar to the present study, find larger aggregate elasticities around 1.2—somewhat below our $\varepsilon_{LL} \approx 1.5$, though well within the credible interval. The gap between micro and macro estimates reflects general equilibrium amplification and dynamic adjustment. Vergara and Swonder (2025) survey 25 estimates of the elasticity of corporate taxable profits with respect to the net-of-tax rate, finding a median of 0.58 and a range from 0.08 to 4.79; our $\varepsilon_{CC} \approx 8.5$ exceeds their maximum, consistent with cumulative responses over a five-year horizon exceeding short-run estimates. The cross-elasticity $\varepsilon_{CL} \approx 21$ has no analog in the existing literature, which focuses on own-base responses.

5 The Postwar Evolution of Tax System Efficiency

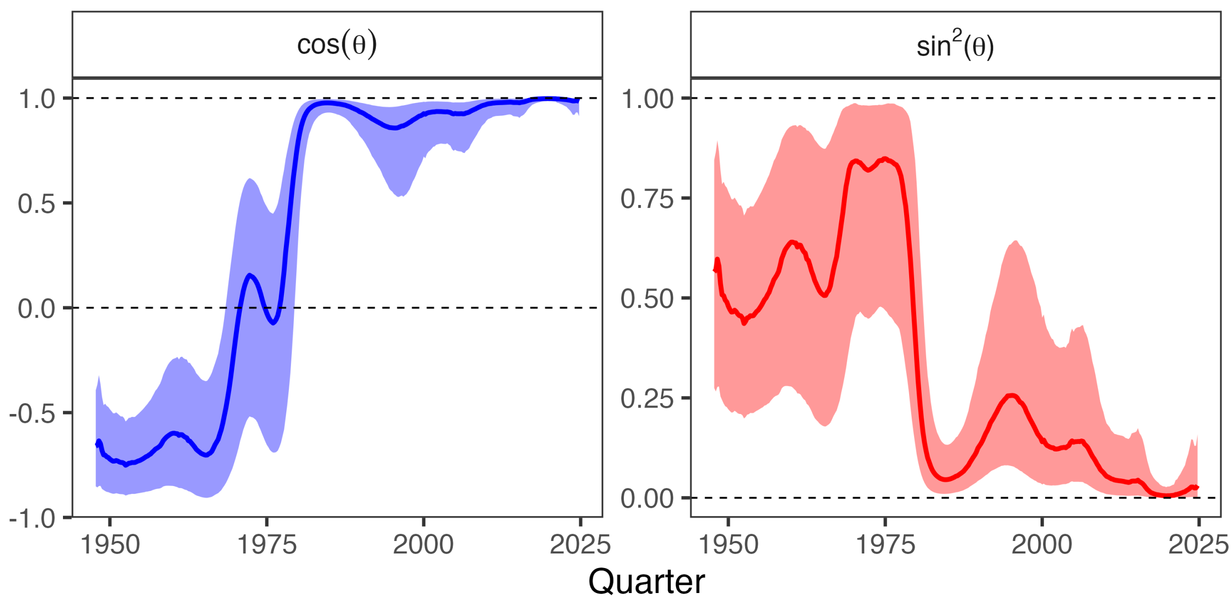
The previous section estimated the revenue gradient from narrative tax shocks; the welfare gradient is simply the vector of tax bases under utilitarian weights. With both gradients in hand, I now apply the projection framework to evaluate U.S. federal income tax policy from 1947 to 2019. For each quarter, I compute the alignment statistics developed in Section 2: $\cos\theta$ measuring proximity to the Ramsey optimum, $\sin^2\theta$ measuring the share of welfare gains available through revenue-neutral reform, and $\cos\phi$ measuring whether actual reforms moved toward or away from the optimal direction.

The alignment statistics require tax bases and rates at each date. A conceptual question arises: should we evaluate efficiency at the raw quarterly values, which include cyclical fluctuations, or at some notion of long-run values? The framework is designed to evaluate *permanent* reforms—the elasticities are cumulated over 5 years precisely to capture long-run responses. Evaluating these permanent-reform statistics at cyclical bases would conflate two distinct questions: whether the long-run tax structure is efficient (which the framework answers) and whether taxes should vary over the business cycle (which it does not). I therefore use the Christiano and Fitzgerald (2003) band-pass filter to extract the trend component of bases and rates, retaining only frequencies longer than 40 quarters. This ensures the alignment metric measures efficiency of the tax system's permanent structure rather than transitory deviations driven by recessions or booms.

5.1 Results

Figure 4 presents the central finding. The left panel plots $\cos\theta$, the alignment between welfare and revenue gradients; the right panel plots $\sin^2\theta$, the share of potential welfare gains available through revenue-neutral reform. Until 1975, $\cos\theta$ was negative—indicating that welfare and revenue gradients were pointing in opposite directions, worse than orthogonal. The gradients crossed orthogonality around 1980, when $\cos\theta$ passed through zero. At the Ramsey optimum these gradients are collinear, so near-orthogonality implies the system was far from optimal. Around 1980, alignment sharply improved as $\cos\theta$ approached 0.8. By around 2020, $\cos\theta$ exceeded 0.95 and $\sin^2\theta$ had fallen below 0.05. The tax system moved from severe misalignment to near-optimality.

Figure 4: The Evolution of Tax System Efficiency

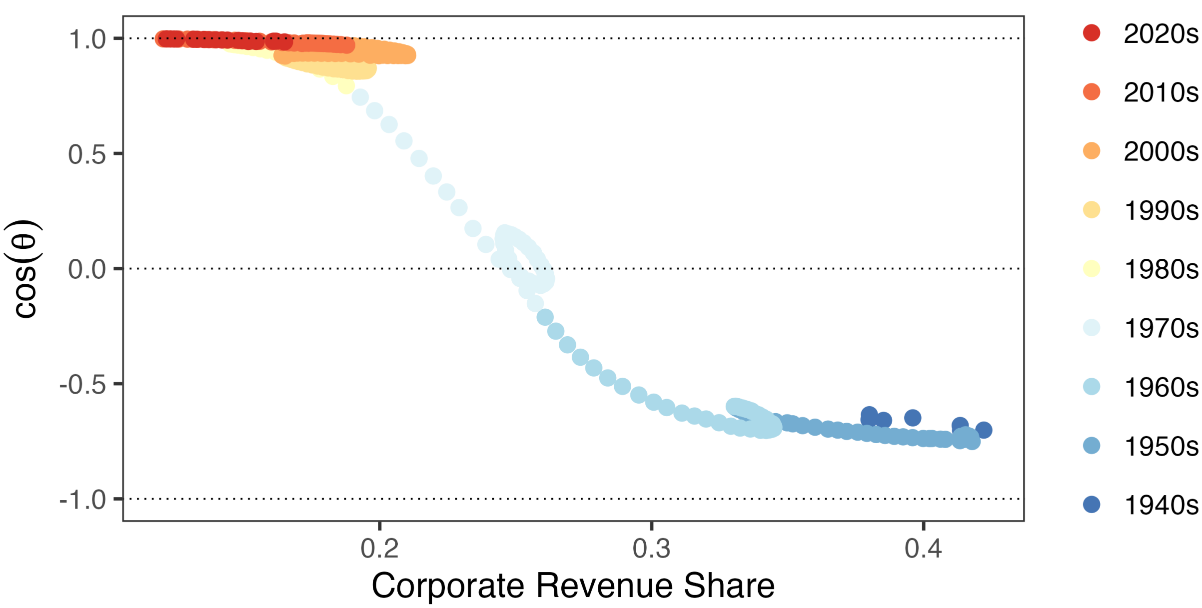


Note: Left panel plots $\cos\theta$, the alignment between welfare and revenue gradients; $\cos\theta = 1$ indicates the Ramsey optimum. Right panel plots $\sin^2\theta$, the share of potential welfare gains available through revenue-neutral reform. Shaded regions show 90% credible intervals.

What drove this improvement? Figure 5 reveals that alignment tracks the corporate share of income tax revenue almost perfectly. When corporate revenue exceeded 40% of income tax collections in the late 1940s, the system was severely misaligned; as the share fell below 20%, alignment improved sharply. The relationship is nearly monotonic. The economic logic follows from equation (21). Two features of the elasticity matrix matter: the corporate base is highly elastic to both taxes ($\varepsilon_{CC} \approx 8.5$, $\varepsilon_{CL} \approx 21$), and these elasticities enter the revenue gradient scaled by corporate revenue $\tau_C B_C$. When corporate revenue was large, both the own-base response and the cross-base spillover made corporate taxes appear expensive and personal taxes appear cheap in revenue terms relative to welfare terms, pushing the revenue gradient away from the welfare gradient. As corporate revenue shrank, these elastic responses became fiscally negligible and the gradients converged.

Why did the corporate share decline? Several forces contributed. Pass-through entities—S-corporations, partnerships, and LLCs—grew considerably between the late 1970s and the present (Dyrda and Pugsley 2024), partly because TRA 1986 cut top personal rates below the corporate rate, making pass-through structures newly attractive (Barro and Wheaton 2019; Smith et al. 2022). The economy also shifted toward human-capital-intensive services with smaller corporate profit shares. Moreover, statutory corporate rate cuts combined with base-shrinking reforms like accelerated depreciation for equipment directly reduced corporate revenue. These forces combined to shrink R_C/R_L, which in turn improved alignment.

Figure 5: System Efficiency and the Corporate Revenue Share



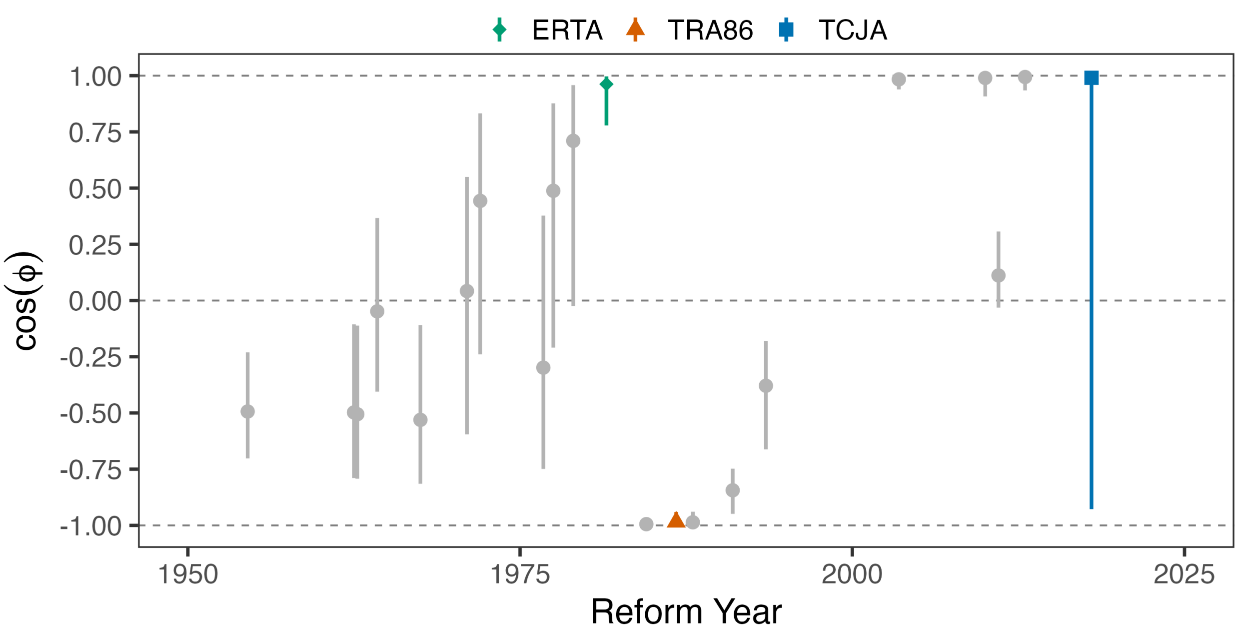
Note: Each point represents a quarter, colored by decade. The horizontal axis plots the corporate share of federal income tax revenue; the vertical axis plots $\cos\theta$, the alignment between welfare and revenue gradients. Higher values indicate greater alignment with the Ramsey optimum.

However, just because the system became more aligned does not mean that it was the result of good tax reforms. The framework offers a natural tool to investigate that via $\cos\phi$, which evaluates whether reforms actually moved in the optimal reform direction. Figure 6 plots $\cos\phi$

for each reform in the Mertens and Ravn (2013a) narrative dataset.[8] Until the Reagan era, the reforms were a mixed bag. However, after Reagan's Economic Recovery and Tax Act (ERTA) of 1981—which accelerated depreciation further—there were no major welfare-improving reforms until the 2000s, during which all exogenous reforms were welfare-improving.

It is somewhat surprising to find that the post-ERTA reforms were largely inefficient, but it follows directly from the projection framework. The Deficit Reduction Act of 1984 scaled back ERTA's investment incentives. The Tax Reform Act of 1986 (TRA86) went further, eliminating the investment tax credit, lengthening depreciation schedules, and cutting top personal rates from 50% to 28%. TRA86 shifted a large share of the tax burden from individuals to corporations—the direction the framework identifies as inefficient, despite TRA86 being widely praised for its efficiency properties (Auerbach and Slemrod 1997). The dip in $\cos\theta$ visible in Figure 4 around 1986–1987 confirms this: alignment fell temporarily before resuming its upward trend. Whatever its merits for base-broadening and equity, TRA86 moved against the efficiency margin measured here.

Figure 6: The Alignment of Reforms with their Optimal Direction



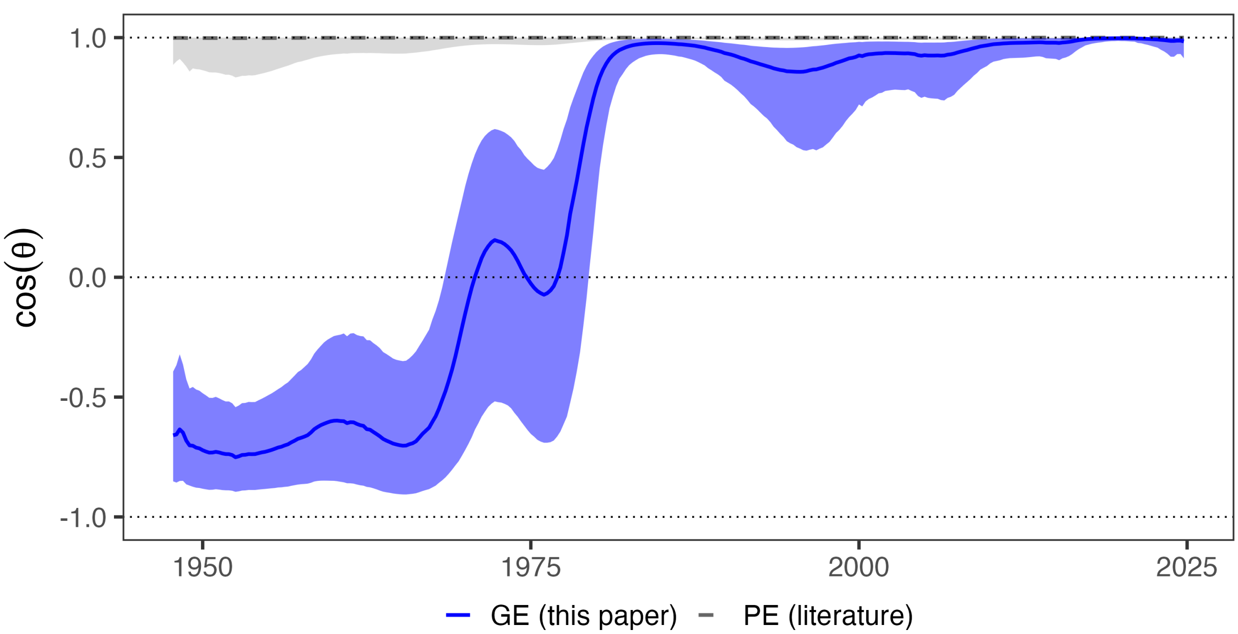
Note: Each bar plots $\cos\phi$, the alignment between the actual reform direction and the optimal direction. Values near +1 indicate reform moved toward the optimum; values near −1 indicate reform moved away. Error bars show 95% credible intervals.

The 2000s reforms, starting with accelerated depreciation policies following 9/11, were largely

8. I evaluate the reform direction using the Mertens and Ravn (2013a) direct shock.

efficient. Notably, the Tax Cuts and Jobs Act of 2017 (TCJA) was nearly ideal along the composition margin, cutting the corporate rate from 35% to 21%, increasing investment incentives, and cutting personal tax rates. Moreover, it is evident from Figure 4 that the post-TCJA system is nearly optimal, with the optimal direction still pointing toward modestly lower corporate rates—TCJA moved correctly but did not exhaust all efficiency gains. Yet the remaining gains are small ($\sin^2\theta < 0.05$). The wide credible interval on TCJA's $\cos\phi$ reflects this near-optimality: when the system is close to the Ramsey optimum, the optimal reform direction becomes nearly indeterminate, and small perturbations to the estimated gradients can swing the direction substantially. My evaluation stands somewhat in contrast to Vergara and Swonder (2025), who find TCJA unjustifiable unless implausibly large weight is placed on firm owners.[9] The divergence stems from two sources: they estimate partial equilibrium elasticities with respect to marginal rates, while I estimate general equilibrium elasticities with respect to average rates; and they evaluate distributional incidence across heterogeneous households, while I adopt a utilitarian benchmark. The frameworks answer different questions—theirs asks who benefits, mine asks whether composition is efficient.

Figure 7: Partial vs. General Equilibrium Alignment



Note: Solid blue line shows $\cos\theta$ from general equilibrium estimates with cross-base spillovers. Dashed gray line and shaded band show $\cos\theta$ using partial equilibrium own-elasticities from the literature with cross-elasticities set to zero. The PE band spans all combinations of $\varepsilon_{CC} \in [0.08, 4.79]$ and $\varepsilon_{LL} \in [0.12, 0.40]$.

9. This follows from work from Kennedy et al. (2023) showing that TCJA's benefits accrued almost entirely to the top 10% of earners.

Figure 7 illustrates the importance of the general equilibrium approach. Using own-base elasticities from the partial equilibrium literature—$\varepsilon_{CC} \in [0.08, 4.79]$ from Vergara and Swonder (2025) and $\varepsilon_{LL} \in [0.12, 0.40]$ from Saez, Slemrod, and Giertz (2012)—with cross-base spillovers set to zero, alignment appears near-perfect throughout the postwar period. The intuition is straightforward: with small own-elasticities and no spillovers, the revenue gradient $r_j \approx B_j$ is approximately proportional to the welfare gradient $g_j = B_j$, so the two are mechanically aligned regardless of the tax mix. The general equilibrium estimates reveal what PE misses: the corporate base is highly elastic to both taxes, and these elasticities enter the revenue gradient scaled by corporate revenue. When corporate revenue was large, both the own-base response and the cross-base spillover made corporate taxes appear expensive and personal taxes appear cheap in revenue terms, pushing the revenue gradient away from the welfare gradient. This created the severe early misalignment that PE cannot detect.

5.2 Robustness

The main results rely on present-value general equilibrium elasticities generated by a Bayesian local projection using utilitarian welfare and a particular discount rate. How sensitive are the findings to these choices?

Discount rates. Present-value calculations require a discount factor, and higher discount rates place more weight on near-term responses where estimation is precise. Figure E.2 varies β from 0.97 to 1. The improvement in alignment is robust across this range: all specifications with $\beta \leq 0.99$ show $\cos\theta$ rising from below 0.5 to nearly perfect alignment.

Estimation methods. The revenue gradient depends on impulse response estimates, which differ across methods. Figure E.3 compares nine approaches. The local projection variants (BLP, LP, SLP, and LP-BC) and the BVAR cluster tightly, all showing alignment rising to above 0.9 by the 2010s. The SVAR variant is an outlier and show no apparent change from the postwar period to the present, but instead are evidence of severe misalignment. Figure E.4 shows that the same is largely true when including a linear trend.

Welfare weights. The welfare gradient depends on how society values different taxpayers. Figure E.1 varies these weights. Results are almost entirely insensitive to ω_C, the weight on corporate taxpayers—all specifications in the left panel nearly overlap. Results are more sensitive to ω_L, but the pattern holds for any $\omega_L \geq 0.5$. Only extreme assumptions—placing near-zero weight on personal income taxpayers—yield meaningfully different results. The asymmetry reflects the larger personal tax base: its welfare weight matters more for the gradient calculation.

Horizon length. The elasticities depend on the length of the horizon cumulated over time. Figure E.5 shows that the upward trend in $\cos\theta$ is consistent across horizon lengths from 20 to 100 quarters under the baseline specification.

Seven-variable specification. The baseline uses four variables: tax rates and bases for both instruments. Appendix E shows robustness to a seven-variable specification that adds government spending, GDP, and debt following Mertens and Ravn (2013a). The seven-variable specification yields qualitatively similar point estimates but wider credible intervals, as expected given the additional parameters. The alignment metrics are robust to this choice: $\cos\theta$ in the 2010s exceeds 0.8 under both specifications.

In sum, the central finding—that alignment improved from severely negative to near-optimal over the postwar period—is robust to discount rates, estimation methods, welfare weights, and horizon length.

6 Conclusion

This paper develops a geometric framework for evaluating tax reform in general equilibrium. The angle between the welfare gradient and the revenue gradient measures alignment with the Ramsey optimum: when these vectors are collinear, no revenue-neutral reform can improve welfare; when they are orthogonal, the system is maximally inefficient. Computing this angle requires only two sufficient statistics, tax bases and the general equilibrium response of revenue to tax changes, without specifying preferences, production, or market structure.

Applying this framework to the U.S. federal tax system from 1947 to 2019 reveals that alignment improved dramatically over the postwar period. In the early 1950s, welfare and revenue gradients were nearly orthogonal, indicating severe misalignment: the tax system was as far from optimal as possible without actively destroying welfare, and more than 90% of potential efficiency gains were available through revenue-neutral reform. By the 2010s, $\cos\theta$ exceeded 0.95 and the share of gains available revenue-neutrally ($\sin^2\theta$) had fallen below 5%. The system moved from severe misalignment to near-optimality.

The U.S. income tax system has largely exhausted the efficiency gains available through revenue-neutral reform along the corporate-personal margin. The framework points toward margins outside the current analysis (consumption taxes, payroll taxes, wealth taxes) where cross-base spillovers may create new misalignments. More broadly, the projection approach applies wherever policymakers face multi-instrument reform problems with general equilibrium spillovers: state and local taxation, environmental policy, or social insurance. The sufficient statistics it requires, welfare costs and revenue responses, can be estimated from reduced-form evidence in any of these settings.

References

Ahmad, Ehtisham, and Nicholas Stern. 1984. "The theory of reform and indian indirect taxes." *Journal of Public Economics* 25, no. 3 (December): 259–298. ISSN: 00472727. https://doi.org/10.1016/0047-2727(84)90057-4.

Auerbach, Alan A., and Joel Slemrod. 1997. "The Economic Effects of the Tax Reform Act of 1986." *Journal of Economic Literature* 35 (2): 589–632.

Badel, Alejandro, Mark Huggett, and Wenlan Luo. 2020. "Taxing Top Earners: a Human Capital Perspective." *The Economic Journal* 130, no. 629 (July): 1200–1225. ISSN: 0013-0133. https://doi.org/10.1093/ej/ueaa021.

Barnichon, Regis, and Christian Brownlees. 2019. "Impulse Response Estimation by Smooth Local Projections." *The Review of Economics and Statistics* 101, no. 3 (July): 522–530. ISSN: 0034-6535. https://doi.org/10.1162/rest{_}a{_}00778.

Barro, Robert, and Brian Wheaton. 2019. *Taxes, Incorporation, and Productivity.* Technical report. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w25508. http://www.nber.org/papers/w25508.pdf.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy* 87, no. 5, Part 1 (October): 940–971. ISSN: 0022-3808. https://doi.org/10.1086/260807.

Bergstrom, Katy, William Dodds, and Juan Rios. 2024. "Optimal policy reforms."

Bhandari, Anmol, Jaroslav Borovička, and Yuki Yao. 2025. *Robust Bounds on Optimal Tax Progressivity.* Technical report. Cambridge, MA: National Bureau of Economic Research, September. https://doi.org/10.3386/w34218.

Caravello, Tomás, Alisdair McKay, and Christian Wolf. 2024. *Evaluating Monetary Policy Counterfactuals: (When) Do We Need Structural Models?* Technical report. Cambridge, MA: National Bureau of Economic Research, September. https://doi.org/10.3386/w32988.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica* 54 (3): 607–622.

Chari, V.V., Juan Pablo Nicolini, and Pedro Teles. 2020. "Optimal capital taxation revisited." *Journal of Monetary Economics* 116 (December): 147–165. ISSN: 03043932. https://doi.org/10.1016/j.jmoneco.2019.09.015.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1, no. 1 (September): 451–488. ISSN: 1941-1383. https://doi.org/10.1146/annurev.economics.050708.142910.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2011. "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review* 101, no. 3 (May): 471–475. ISSN: 0002-8282. https://doi.org/10.1257/aer.101.3.471.

Christiano, Lawrence J., and Terry J. Fitzgerald. 2003. "The Band Pass Filter." *International Economic Review* 44, no. 2 (May): 435–465. ISSN: 0020-6598. https://doi.org/10.1111/1468-2354.t01-1-00076.

Cloyne, James, Joseba Martinez, Haroon Mumtaz, and Paolo Surico. 2022. *Short-Term Tax Cuts, Long-Term Stimulus.* Technical report. Cambridge, MA: National Bureau of Economic Research, July. https://doi.org/10.3386/w30246.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review* 99, no. 1 (February): 25–48. ISSN: 0002-8282. https://doi.org/10.1257/aer.99.1.25.

Diamond, Peter, and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendation." *Journal of Economic Perspectives* 25, no. 4 (November): 165–190. ISSN: 0895-3309. https://doi.org/10.1257/jep.25.4.165.

Diamond, Peter A., and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1): 8–27.

Dixit, Avinash. 1975. "Welfare effects of tax and price changes." *Journal of Public Economics* 4, no. 2 (February): 103–123. ISSN: 00472727. https://doi.org/10.1016/0047-2727(75)90014-6.

Doan, Thomas, Robert Litterman, and Christopher Sims. 1984. "Forecasting and conditional projection using realistic prior distributions." *Econometric Reviews* 3, no. 1 (January): 1–100. ISSN: 0747-4938. https://doi.org/10.1080/07474938408800053.

Dyrda, Sebastian, and Benjamin Pugsley. 2024. "The Rise of Pass-Throughs: an Empirical Investigation." *The Economic Journal* 135, no. 665 (December): 387–403. ISSN: 0013-0133. https://doi.org/10.1093/ej/ueae062.

Favero, Carlo, and Francesco Giavazzi. 2012. "Measuring Tax Multipliers: The Narrative Method in Fiscal VARs." *American Economic Journal: Economic Policy* 4, no. 2 (May): 69–94. ISSN: 1945-7731. https://doi.org/10.1257/pol.4.2.69. https://pubs.aeaweb.org/doi/10.1257/pol.4.2.69.

Feldstein, Martin. 1976. "On the theory of tax reform." *Journal of Public Economics* 6, nos. 1-2 (July): 77–104. ISSN: 00472727. https://doi.org/10.1016/0047-2727(76)90042-6.

Finkelstein, Amy, and Nathaniel Hendren. 2020. "Welfare Analysis Meets Causal Inference." *Journal of Economic Perspectives* 34, no. 4 (November): 146–167. ISSN: 0895-3309. https://doi.org/10.1257/jep.34.4.146.

Geweke, J. 1993. "Bayesian treatment of the independent student-t linear model." *Journal of Applied Econometrics* 8, no. S1 (December): S19–S40. ISSN: 08837252. https://doi.org/10.1002/jae.3950080504.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri. 2015. "Prior Selection for Vector Autoregressions." *Review of Economics and Statistics* 97, no. 2 (May): 436–451. ISSN: 0034-6535. https://doi.org/10.1162/REST{_}a{_}00483.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30, no. 1 (January): 51–89. ISSN: 0892-8649. https://doi.org/10.1086/685593.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *The Quarterly Journal of Economics* 135, no. 3 (August): 1209–1318. ISSN: 0033-5533. https://doi.org/10.1093/qje/qjaa006.

Herbst, Edward P., and Benjamin K. Johannsen. 2024. "Bias in local projections." *Journal of Econometrics* 240, no. 1 (March): 105655. ISSN: 03044076. https://doi.org/10.1016/j.jeconom.2024.105655.

JCT. 2017. *Macroeconomic Analysis Of The Conference Agreement For H.R. 1, The Tax Cuts And Jobs Act.* Technical report. https://www.jct.gov/publications/2017/jcx-69-17/.

Jorda, Oscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95, no. 1 (February): 161–182. ISSN: 0002-8282. https://doi.org/10.1257/0002828053828518. https://pubs.aeaweb.org/doi/10.1257/0002828053828518.

Kennedy, Patrick J, Christine Dobridge, Paul Landefeld, and Jacob Mortenson. 2023. *The Efficiency-Equity Tradeoff of the Corporate Income Tax: Evidence from the Tax Cuts and Jobs Act.* Technical report.

Kleven, Henrik, Claus Kreiner, Kristian Larsen, and Jakob Søgaard. 2025. "Micro versus Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort." *American Economic Review* 115, no. 9 (September): 2849–2890. ISSN: 0002-8282. https://doi.org/10.1257/aer.20240554.

Li, Dake, Mikkel Plagborg-Møller, and Christian K. Wolf. 2024. "Local Projections vs. VARs: Lessons From Thousands of DGPs." *Journal of Econometrics* 244 (2). https://doi.org/https://doi.org/10.1016/j.jeconom.2024.105722.

Lucas, Robert E. 1990. "Supply-Side Economics: An Analytical Review." *Oxford Economic Papers* 42 (2): 293–316.

Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan. 2009. "Optimal Taxation in Theory and Practice." *Journal of Economic Perspectives* 23, no. 4 (November): 147–174. ISSN: 0895-3309. https://doi.org/10.1257/jep.23.4.147.

McKay, Alisdair, and Christian K. Wolf. 2023. "What Can Time-Series Regressions Tell Us About Policy Counterfactuals?" *Econometrica* 91 (5): 1695–1725. ISSN: 0012-9682. https://doi.org/10.3982/ECTA21045.

Mertens, Karel, and José Luis Montiel Olea. 2018. "Marginal Tax Rates and Income: New Time Series Evidence*." *The Quarterly Journal of Economics* 133, no. 4 (November): 1803–1884. ISSN: 0033-5533. https://doi.org/10.1093/qje/qjy008.

Mertens, Karel, and Morten O. Ravn. 2012. "Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks." *American Economic Journal: Economic Policy* 103 (4): 145–181.

———. 2013a. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review* 103, no. 4 (June): 1212–1247. ISSN: 0002-8282. https://doi.org/10.1257/aer.103.4.1212. https://pubs.aeaweb.org/doi/10.1257/aer.103.4.1212.

———. 2013b. "The Dynamic Effects of Personal and Corporate Income Taxes in the United States." *American Economic Review* 103:1212–1247.

———. 2014. "A reconciliation of SVAR and narrative estimates of tax multipliers." *Journal of Monetary Economics* 68 (December): S1–S19. ISSN: 03043932. https://doi.org/10.1016/j.jmoneco.2013.04.004. https://linkinghub.elsevier.com/retrieve/pii/S0304393213000536.

Moon, Sarah. 2025. "Optimal Policy Choices Under Uncertainty."

Ramsey, F. P. 1927. "A Contribution to the Theory of Taxation." *The Economic Journal* 37, no. 145 (March): 47. ISSN: 00130133. https://doi.org/10.2307/2222721.

Romer, Christina D, and David H Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3): 763–801.

Saez, E. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *The Review of Economic Studies* 68, no. 1 (January): 205–229. ISSN: 0034-6527. https://doi.org/10.1111/1467-937X.00166.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50, no. 1 (March): 3–50. ISSN: 0022-0515. https://doi.org/10.1257/jel.50.1.3.

Sims, Christopher A. 1993. "A Nine-Variable Probabilistic Macroeconomic Forecasting Model." In *Business Cycles, Indicators and Forecasting,* edited by James H. Stock and Mark W. Watson. National Bureau of Economic Research.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick. 2022. "The Rise of Pass-Throughs and the Decline of the Labor Share." *American Economic Review: Insights* 4, no. 3 (September): 323–340. ISSN: 2640-205X. https://doi.org/10.1257/aeri.20210268.

Straub, Ludwig, and Iván Werning. 2020. "Positive Long-Run Capital Taxation: Chamley-Judd Revisited." *American Economic Review* 110, no. 1 (January): 86–119. ISSN: 0002-8282. https://doi.org/10.1257/aer.20150210.

Tirole, J., and R. Guesnerie. 1981. "Tax reform from the gradient projection viewpoint." *Journal of Public Economics* 15, no. 3 (June): 275–293. ISSN: 00472727. https://doi.org/10.1016/0047-2727(81)90012-8.

Vergara, Damian, and Dustin Swonder. 2025. "A Sufficient Statistics Approach to Optimal Corporate Taxes."

Appendices

Online Appendix

A Proofs

A.1 MCPF Equalization

The government chooses $\boldsymbol{\tau}$ to minimize social cost subject to a revenue requirement:

$$\min_{\boldsymbol{\tau}} W(\boldsymbol{\tau}) \quad \text{subject to} \quad R(\boldsymbol{\tau}) = G. \tag{A.1}$$

Proposition 2 (Ramsey first-order condition)**.** *Suppose an interior optimum $\boldsymbol{\tau}^* \in int(\mathcal{T})$ exists with $B_k(\boldsymbol{\tau}^*) > 0$ and $r_k(\boldsymbol{\tau}^*) > 0$ for all k, and constraint qualification holds. Then there exists $\lambda \in \mathbb{R}$ such that*

$$g(\boldsymbol{\tau}^*) = \lambda\, r(\boldsymbol{\tau}^*). \tag{A.2}$$

Equivalently, marginal costs of public funds are equalized across instruments:

$$MCPF_i(\boldsymbol{\tau}^*) := \frac{g_i(\boldsymbol{\tau}^*)}{r_i(\boldsymbol{\tau}^*)} = \lambda \quad \text{for all } i. \tag{A.3}$$

Proof. Form the Lagrangian $\mathcal{L}(\boldsymbol{\tau}, \lambda) = W(\boldsymbol{\tau}) - \lambda(R(\boldsymbol{\tau}) - G)$. Interior first-order conditions give $\nabla_{\boldsymbol{\tau}} \mathcal{L} = 0$, i.e., $g(\boldsymbol{\tau}^*) = \lambda\, r(\boldsymbol{\tau}^*)$. Dividing componentwise by $r_i(\boldsymbol{\tau}^*) > 0$ yields (A.3). □

A.2 Proof of Proposition 1

Proof. The problem is $\max_{d\boldsymbol{\tau}} -g^\top d\boldsymbol{\tau}$ subject to $r^\top d\boldsymbol{\tau} = 0$. Form the Lagrangian:

$$\mathcal{L}(d\boldsymbol{\tau}, \mu) = -g^\top d\boldsymbol{\tau} + \mu r^\top d\boldsymbol{\tau}. \tag{A.4}$$

First-order conditions yield $-g + \mu r = 0$. This cannot be satisfied for finite μ when MCPFs differ (which would require $g = \mu r$, i.e., g parallel to r). The constraint $r^\top d\boldsymbol{\tau} = 0$ implies $r^\top(-g + \mu r) = 0$, giving:

$$\mu = \frac{r^\top g}{r^\top r}. \tag{A.5}$$

Thus:

$$d\boldsymbol{\tau}^* \propto -\left(g - \frac{r^\top g}{r^\top r} r\right) = -P_{\perp r} g. \tag{A.6}$$

The welfare gain is $\Delta W = g^\top(\alpha d\boldsymbol{\tau}^*) = -\alpha g^\top P_{\perp r} g = -\alpha (P_{\perp r} g)^\top (P_{\perp r} g) = -\alpha \|P_{\perp r} g\|^2$, where the third equality uses symmetry of $P_{\perp r}$ and the fourth uses the definition of the norm. Since $P_{\perp r} g$ is the component of g orthogonal to r, we have $\|P_{\perp r}g\| = \|g\| \sin\theta$, yielding $\Delta W = -\alpha\|g\|^2 \sin^2\theta$. □

A.3 Interpreting the Alignment Metric

Table A.1: Interpreting the Alignment Metric

$\cos\theta$	θ	$\sin^2\theta$	Interpretation
1.00	0°	0.00	Ramsey optimum
0.87	30°	0.25	Near optimum; 25% of gains RN
0.50	60°	0.75	Moderate misalignment; 75% of gains RN
0.00	90°	1.00	Orthogonal; all gains RN
−0.50	120°	0.75	Negative alignment; 75% of gains RN
−0.87	150°	0.25	Severe misalignment; 25% of gains RN
−1.00	180°	0.00	Perfect anti-alignment

Note: Revenue-neutral potential $\sin^2\theta$ is maximized at $\theta = 90°$ and symmetric around it. For example, $\theta = 60°$ and $\theta = 120°$ both yield $\sin^2\theta = 0.75$ because the perpendicular component $\|g\| \sin\theta$ has equal magnitude in both cases.

B Data

The variables and their construction are an extension of Mertens and Ravn (2013a) and Cloyne et al. (2022) from 1947Q1-2019Q4:

- **Personal Income Tax Base.** The personal income tax base is National Income and Product Accounts (NIPA) personal income (Table 2.1 Line 1) plus contributions for government social insurance (3.2 Line 11) less transfers (2.1 Line 17).
- **Personal Income Tax Revenue.** This comes from summing Table 3.2 Lines 3 and 11.
- **Average Personal Income Tax Rate.** The APITR is revenue divided by the lagged base.
- **Corporate Income Tax Base.** The corporate tax base is table 1.2 Line 13 (corporate profits with inventory valuation and capital consumption adjustments) less Federal Reserve Bank profits (Historical Tables 6.16 B-C-D).
- **Corporate Income Tax Revenue.** Corporate revenue is federal taxes on corporate income excluding Federal Reserve banks (Table 3.2 Line 8).

- **Average Corporate Income Tax Rate.** The ACITR is revenue divided by the lagged base.
- **Government Spending.** Real federal government consumption expenditures and gross investment (1.1.3 Line 23).
- **GDP.** Real GDP is from NIPA Table 1.1.3 Line 1.
- **Debt.** Debt is spliced together using Federal debt held by the public from Favero and Giavazzi (2012) and the series FYGFDPUN from FRED.
- **Narrative Instruments.** See Cloyne et al. (2022) for an extension to the Mertens and Ravn (2013a) corporate and personal income tax shock series.

All series in levels are deflated by the GDP deflator (NIPA Table 1.1.9 Line 1) and population, then log-transformed. Population is total population over age 16 from the Bureau of Labor Statistics (CNP16OV in FRED). Because this series starts in 1948, we use a simple linear extrapolation to extend back four quarters to 1947Q1.

C Estimation Details

This appendix provides details on the two-stage Bayesian local projection procedure used to estimate impulse response functions. Many of the estimation details follow Cloyne et al. (2022), Appendix G.

C.1 Stage 1: Bayesian VAR with Hierarchical Priors

The first stage estimates a Bayesian VAR to obtain draws from the posterior distribution of the impact matrix A_0. The VAR is:

$$Z_t = c + B_1 Z_{t-1} + \cdots + B_p Z_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma) \tag{A.7}$$

where Z_t is the vector of endogenous variables and $p = 4$ lags.

Priors. Following Giannone, Lenza, and Primiceri (2015), I treat the hyperparameters governing prior tightness as random and estimate them jointly with the VAR coefficients. The prior combines three components:

- *Minnesota prior*: The prior mean for the coefficient on the first own lag is set to the AR(1) coefficient from a preliminary regression; all other coefficients have prior mean zero. The prior variance for the coefficient on lag ℓ of variable j in equation i is:

$$V_{ij\ell} = \frac{\lambda^2}{\ell^2} \cdot \frac{s_i^2}{s_j^2} \tag{A.8}$$

where s_i and s_j are residual standard deviations from preliminary AR(1) regressions and λ controls overall tightness. The prior on λ is $\lambda \sim N^+(0.2, 0.4^2)$ truncated to $[10^{-4}, 5]$.

- *Sum-of-coefficients prior*: This prior, introduced by Doan, Litterman, and Sims (1984), shrinks the sum of lag coefficients toward unity for persistent variables, helping capture unit root behavior.

- *Single-unit-root prior*: This prior implements the dummy initial observation approach of Sims (1993), providing additional shrinkage toward a random walk.

The hyperparameters governing the sum-of-coefficients and single-unit-root priors are also treated as random, with priors $\Gamma(1, 1)$ truncated to $[10^{-4}, 50]$.

Posterior simulation. The posterior is approximated via Markov chain Monte Carlo. Conditional on hyperparameters, the posterior for VAR coefficients and the covariance matrix Σ is Normal-inverse-Wishart and can be sampled directly. The hyperparameters are updated using a Metropolis-Hastings step. I use 20,000 iterations with 10,000 discarded as burn-in.

C.2 Stage 1b: Proxy Identification of A_0

For each draw of the VAR coefficients and covariance matrix from the Stage 1 posterior, I apply the proxy identification procedure of Mertens and Ravn (2013a) to obtain a draw of the impact matrix A_0.

Let $u_t = A_0 \varepsilon_t$ where ε_t contains the structural shocks with $\mathrm{Var}(\varepsilon_t) = I$. Partition the variables so that the first k correspond to the tax rates (instrumented by the narrative measures m_t) and the remaining $n - k$ are non-tax variables. The reduced-form covariance matrix $\Sigma = A_0 A_0'$ can be partitioned conformably:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{A.9}$$

The proxy relevance condition $E[m_t \varepsilon_{1,t}'] = \Phi \neq 0$ and exogeneity condition $E[m_t \varepsilon_{2,t}'] = 0$ identify the subspace spanned by the tax shocks. Regressing the non-tax residuals on the tax residuals, using the proxies as instruments, recovers the ratio $\beta_{21}\beta_{11}^{-1}$. The remaining elements of A_0 corresponding to the tax shocks follow from the algebra in Mertens and Ravn (2013a).

To separate the personal and corporate tax shocks within the identified subspace, I apply a Cholesky factorization. When estimating the impulse response to shock j, I order tax j last. This imposes that shock j has no direct contemporaneous effect on the other tax rate.

Each draw of A_0 is normalized so that the impact effect on the relevant tax rate equals -1 percentage point (a tax cut).

C.3 Stage 2: Bayesian Local Projections

The second stage estimates local projections at each horizon $h \geq 1$. For outcome variable y (e.g., a tax rate or log tax base), the local projection is:

$$y_{t+h} = c^{(h)} + B_1^{(h)} Z_{t-1} + \sum_{\ell=2}^{p} b_\ell^{(h)} Z_{t-\ell} + u_{t+h} \tag{A.10}$$

where $p = 4$ lags. Let $\omega^{(h)}$ collect the coefficients and let $X_t = (Z'_{t-1}, \ldots, Z'_{t-p}, 1)'$ denote the regressors.

Priors. The prior for $\omega^{(h)}$ is Normal with mean ω_0 and variance S_0. The prior mean implies that the outcome variable follows an AR(1) process: the element corresponding to the first own lag equals the AR(1) coefficient from a preliminary regression, with all other elements set to zero. The prior variance takes the Minnesota form:

$$S_{0,j\ell} = \frac{\lambda^2}{\ell^2} \cdot \frac{s_y^2}{s_j^2} \tag{A.11}$$

where s_y is the standard deviation of the outcome and s_j is the standard deviation of regressor j. The tightness parameter is $\lambda = 10$, reflecting weak prior information. The prior variance on the intercept is 100.

Non-Gaussian errors. As discussed in Jorda (2005), local projection residuals are generally heteroskedastic when $h > 0$. To accommodate this, I model the errors as a scale mixture of normals. Specifically:

$$u_{t+h} | \phi_t \sim N(0, \sigma^2/\phi_t), \qquad \phi_t \sim \Gamma(\nu/2, \nu/2) \tag{A.12}$$

Integrating out the mixing variable ϕ_t yields a Student-t distribution for u_{t+h} with ν degrees of freedom (Geweke 1993). This formulation accommodates heteroskedasticity of unknown form without imposing a specific structure.

The degrees of freedom ν receives an exponential prior:

$$p(\nu) \propto \exp(-\nu/\nu_0) \tag{A.13}$$

with $\nu_0 = 10$. Small values of ν allow heavy tails; the prior places mass on moderate tail thickness while permitting approximate normality.

Posterior simulation. I approximate the posterior using Gibbs sampling with a Metropolis-Hastings step for ν. The sampler iterates over the following blocks, conditioning on all other parameters:

1. *Mixing weights* ϕ_t: Given the current residuals and ν, each ϕ_t is drawn independently from a Gamma distribution with shape $(\nu+1)/2$ and rate $(u_{t+h}^2/\sigma^2 + \nu)/2$.

2. *Degrees of freedom* ν: The conditional posterior is nonstandard. I update ν via random walk Metropolis-Hastings, proposing $\nu' = \nu + \eta$ where $\eta \sim N(0, c)$. The proposal is accepted with the usual Metropolis probability. The step size c is tuned to achieve an acceptance rate between 30% and 50%.

3. *Error variance* σ^2: Defining transformed residuals $\tilde{u}_{t+h} = u_{t+h}\sqrt{\phi_t}$, the conditional posterior for σ^2 is inverse Gamma with shape $T/2$ and scale $\sum_t \tilde{u}_{t+h}^2/2$.

4. *Coefficients* $\omega^{(h)}$: Applying the same transformation to the dependent variable and regressors, $\tilde{y}_{t+h} = y_{t+h}\sqrt{\phi_t}$ and $\tilde{X}_t = X_t\sqrt{\phi_t}$, yields a standard conjugate updating formula. The conditional posterior is Normal with variance

$$V = \left(S_0^{-1} + \sigma^{-2}\tilde{X}'\tilde{X}\right)^{-1} \tag{A.14}$$

and mean

$$\bar{\omega} = V\left(S_0^{-1}\omega_0 + \sigma^{-2}\tilde{X}'\tilde{y}\right) \tag{A.15}$$

I run 12,000 iterations, discarding the first 3,000 as burn-in and retaining every third draw, yielding 3,000 posterior draws for inference.

Impulse response construction. At horizon $h = 0$, the impulse response equals the A_0 draw from Stage 1b. At horizons $h \geq 1$, the impulse response is constructed following Jorda (2005):

$$IRF_h = B_1^{(h-1)}A_0 \tag{A.16}$$

where $B_1^{(h)}$ contains the coefficients on Z_{t-1} from the horizon-h local projection. For each posterior draw, I pair the A_0 draw from Stage 1 with the $B_1^{(h)}$ draw from Stage 2, ensuring that uncertainty in the impact matrix propagates through to longer horizons. By construction, at $h = 0$, $IRF_0 = A_0$.

Revenue gradients. As discussed in Section 4, the present-value revenue gradient is constructed from the impulse responses of retention rates and tax bases. For tax instrument j, equation (21) gives:

$$r_j^{PV} = B_{j,0} \cdot A \cdot \left[1 - \frac{\tau_j}{1-\tau_j}\varepsilon_{jj} - \frac{\tau_k}{1-\tau_j}\frac{B_k}{B_j}\varepsilon_{kj}\right], \tag{A.17}$$

where $A = \sum_{h=0}^{H} \beta^h$ is the annuity factor. Since the elasticities ε_{ij} are permanent-equivalent (cumulative base responses divided by the cumulative retention rate response κ_j), the revenue gradient uses the common annuity factor rather than instrument-specific persistence. This ensures that g^{PV} and r^{PV} correspond to the same policy experiment: a permanent marginal reform. This calculation is performed for each posterior draw, yielding a posterior distribution for the revenue gradient.

D Additional Alignment Results

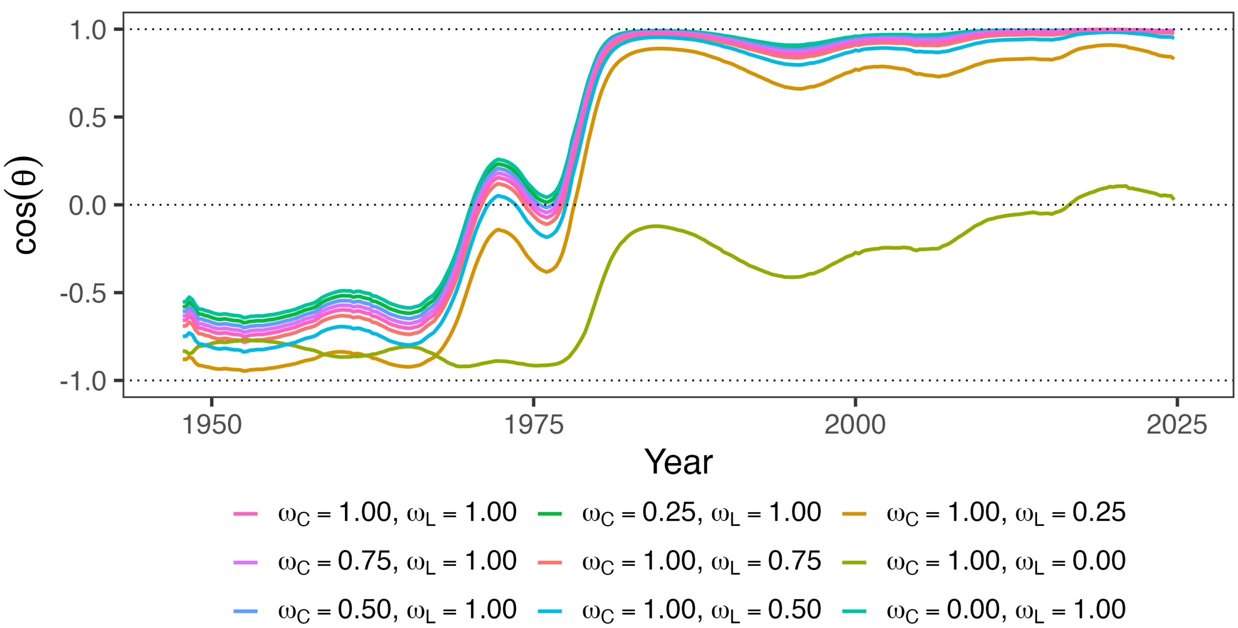
Figure D.1: The Marginal Cost of Public Funds

Note: This figure plots the marginal cost of public funds for corporate and personal income taxes. Following Finkelstein and Hendren (2020), I denote the MCPF as infinite if it is on the wrong side of the Laffer curve.

E Four-Variable System Robustness

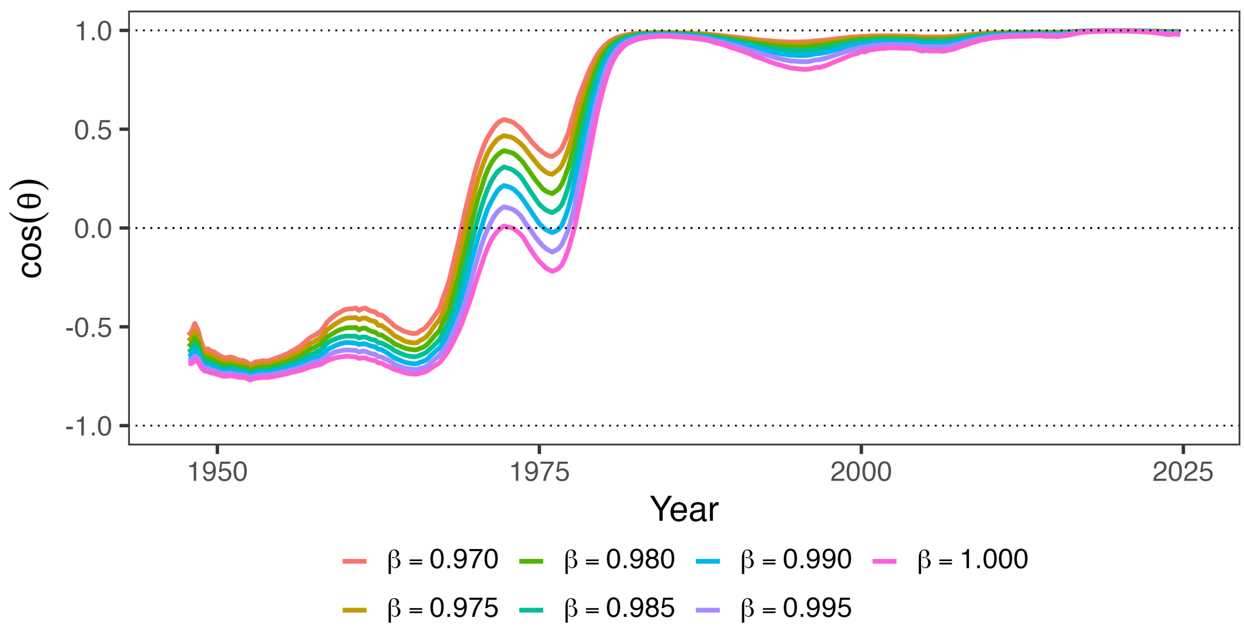
E.1 Robustness of Alignment Statistics

Figure E.1: Robustness Across Welfare Weights



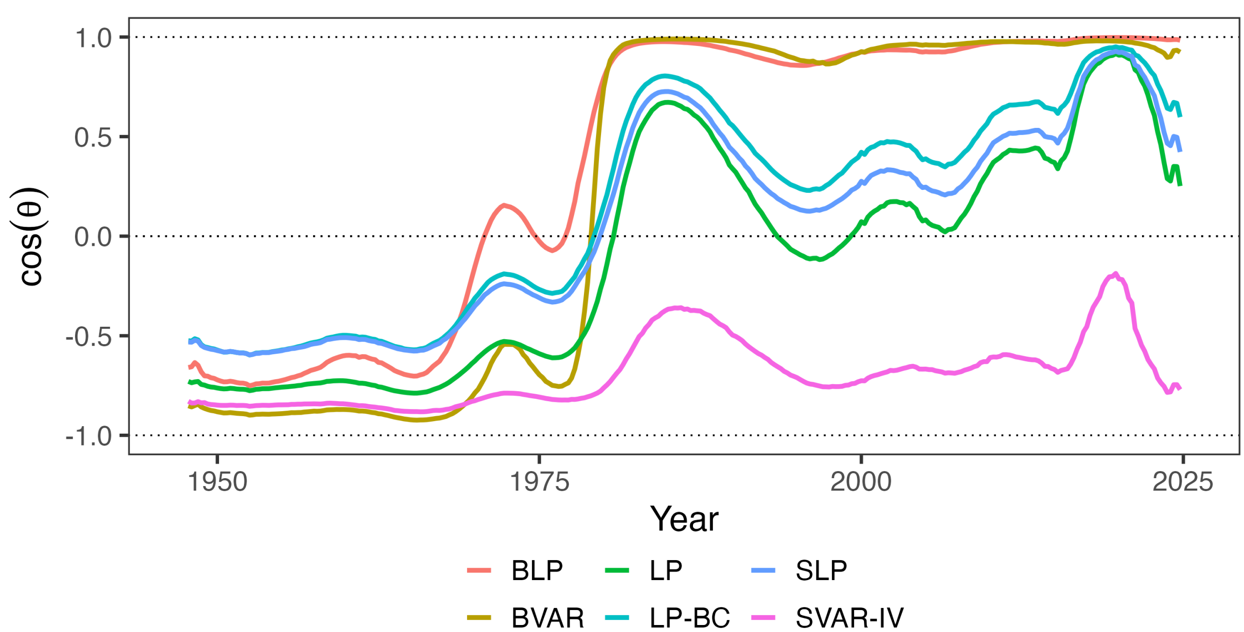
Note: Each panel varies one welfare weight while holding the other at 1. Left panel varies ω_C (weight on corporate taxpayers) with $\omega_L = 1$; right panel varies ω_L (weight on personal taxpayers) with $\omega_C = 1$. The utilitarian baseline has $\omega_C = \omega_L = 1$.

Figure E.2: Robustness Across Discount Factors



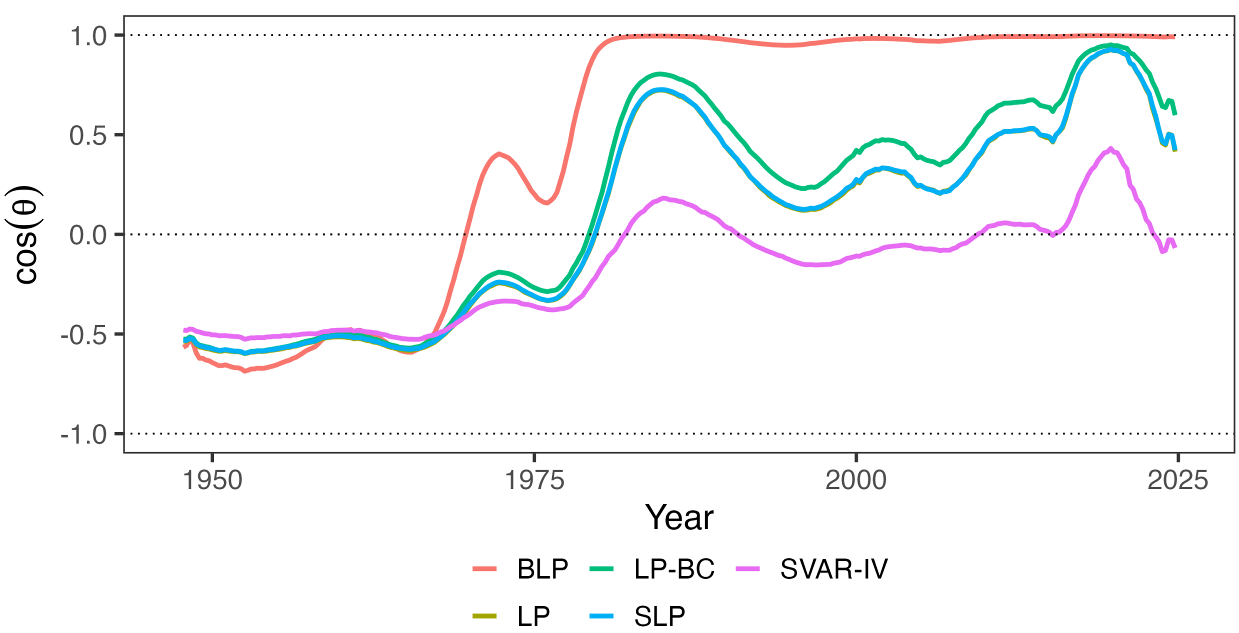
Note: This figure plots alignment $\cos\theta$ across different discount rate regimes. Under the baseline, $\beta = 0.9926$.

Figure E.3: Robustness Across Estimation Methodologies



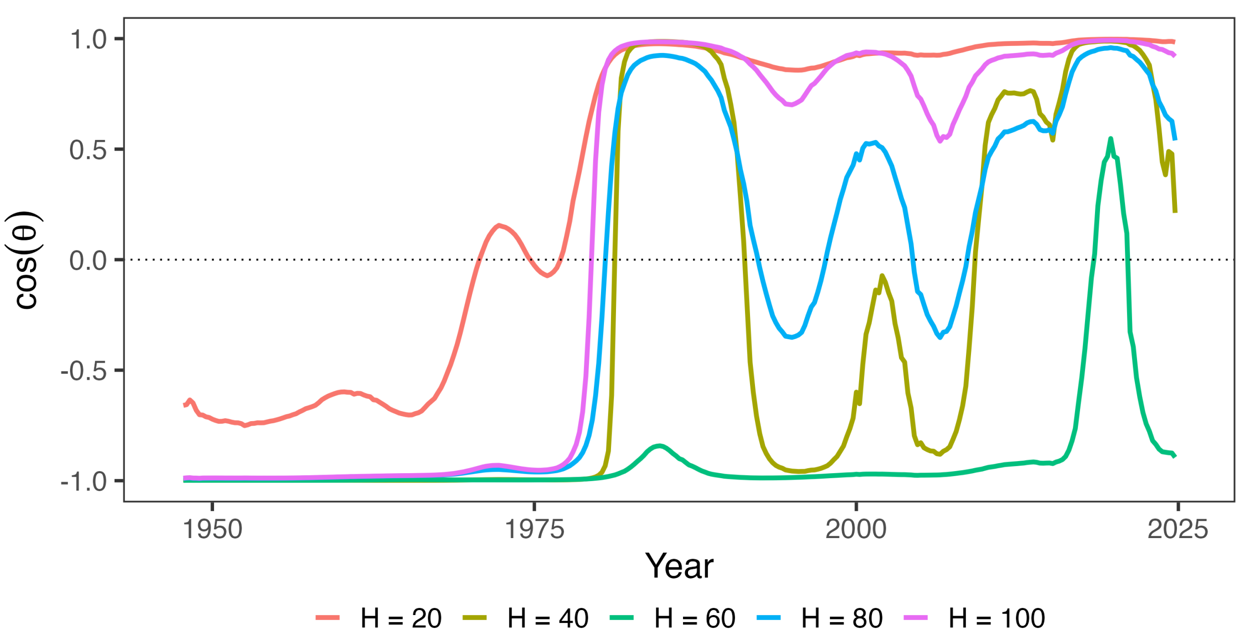
Note: This figure plots alignment $\cos\theta$ across estimation methodologies. See Appendix E for further estimation details.

Figure E.4: Robustness Across Estimation Methodologies



Note: This figure plots alignment $\cos\theta$ across estimation methodologies and adds a linear trend. See Appendix E for further estimation details.

Figure E.5: Robustness Across Horizon Specifications



Note: This figure plots alignment $\cos\theta$ for select models in which vary the horizon length, holding fixed $\beta = 0.9926$.

E.2 Underlying Elasticities and CIRFs

This appendix shows robustness for the own and cross-revenue effects of corporate and personal income tax shocks. Table E.1 in the main text shows the present value of each, while this Appendix shows the IRFs that generate the present value results. All specifications use the same set of variables. The results are consistent across local projections specifications but somewhat noisier with the VAR specifications.

Local projections. This variant uses standard local projections with the Mertens-Ravn structural shocks as instruments. Again, the results are similar, though somewhat smaller. I propagate uncertainty about A_0 using the wild bootstrap proxy SVAR.

Smooth LP. Because the local projections results are lumpy, I show another variant which penalizes deviations across horizons using Barnichon and Brownlees (2019) in Figure E.8.

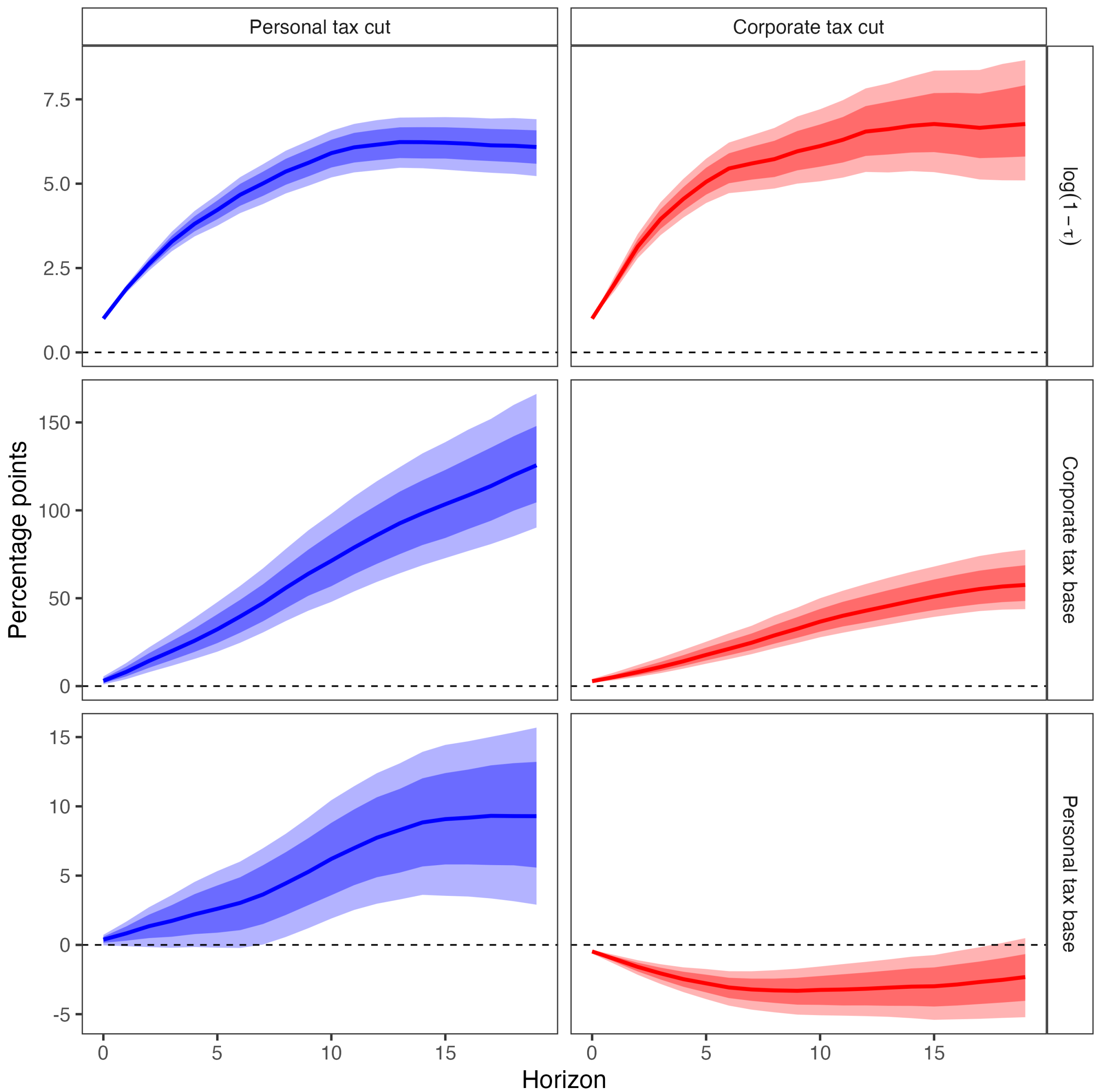
Bias-Corrected LP. Figure E.9 shows the bias-corrected discounted cumulative IRF from Herbst and Johannsen (2024).

Bayesian VAR. The main results come from a Bayesian local projection with the instruments generated from a Bayesian VAR. Figure E.10 shows the impulse responses generated by the Bayesian VAR along with 68% and 90% credible intervals. Following Cloyne et al. (2022), there are $p = 32$ lags. The IRFs are qualitatively similar, though noticeably larger than the baseline BLP.

Proxy SVAR. Figure E.11 shows the impulse responses generated by corporate and personal income tax shocks in a frequentist structural vector autoregression. The specification is the exact same as in the baseline Mertens and Ravn (2013b); it is just extended to 2019Q4. Following their lead, I employ a wild bootstrap. The results are more precise but otherwise qualitatively and quantitatively quite similar to the baseline.

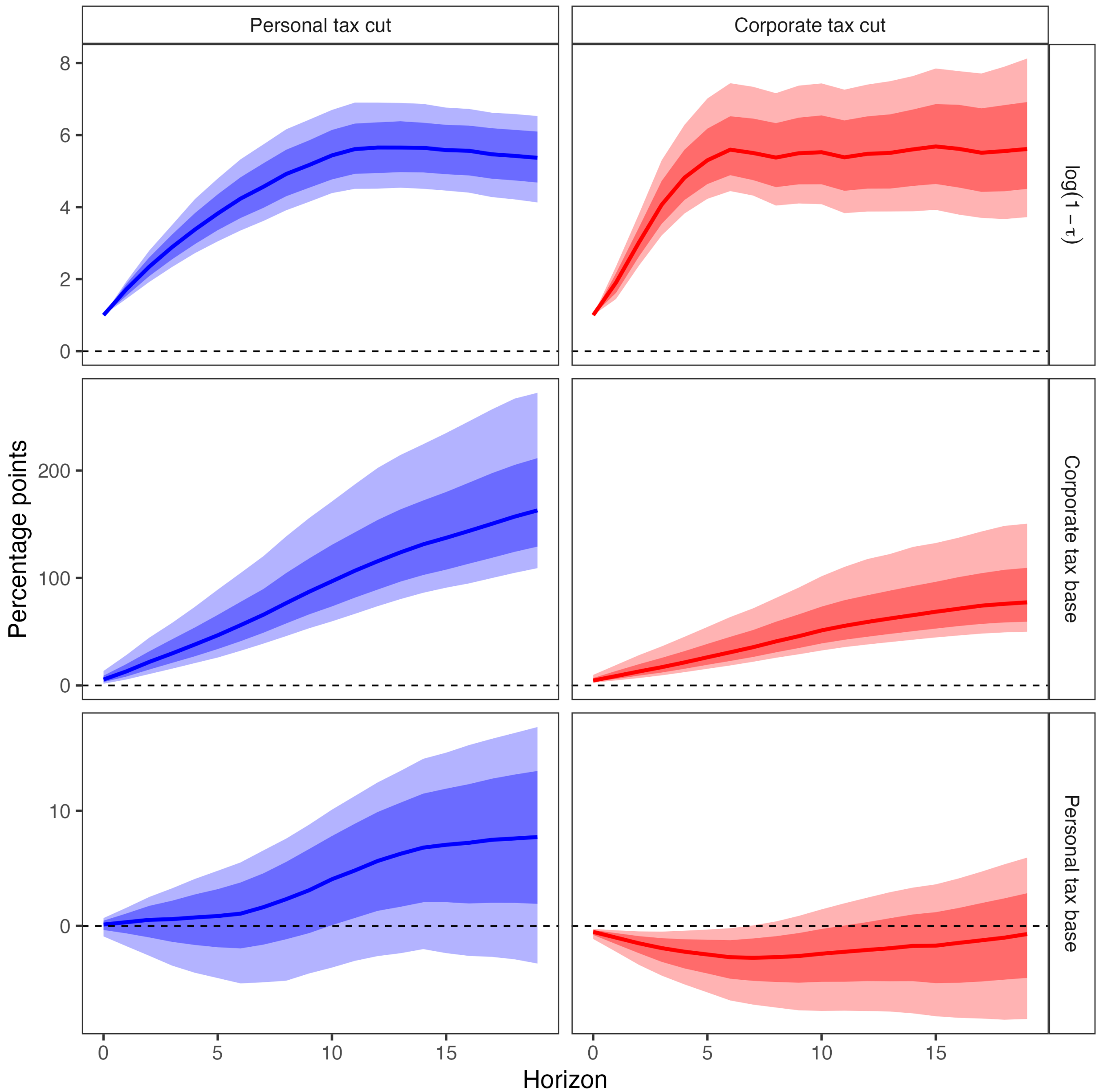
Linear Trends. Figures E.12-E.16 repeat the above specifications but add linear trends.

Figure E.6: Dynamic Effects of Corporate and Personal Income Tax Cuts on Tax Rates and Bases



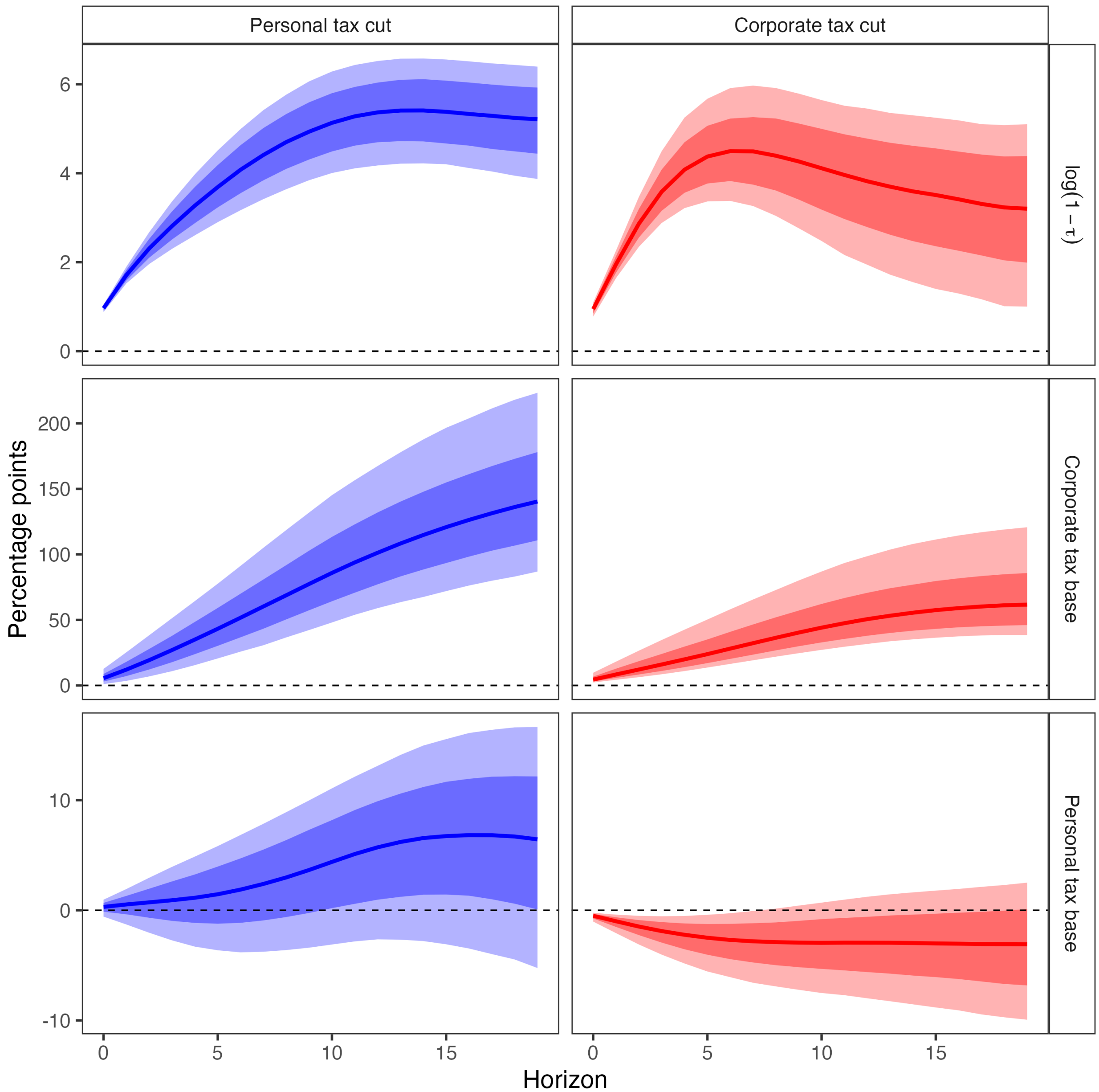
Note: Discounted cumulative impulse responses ($\beta = 0.9926$) to a unit increase in the log retention rate $\log(1 - \tau)$. Sample: 1947Q1–2019Q4. Posterior medians with 68% and 90% credible intervals.

Figure E.7: Dynamic Effects of Narrative Tax Shocks on Revenue (LP)



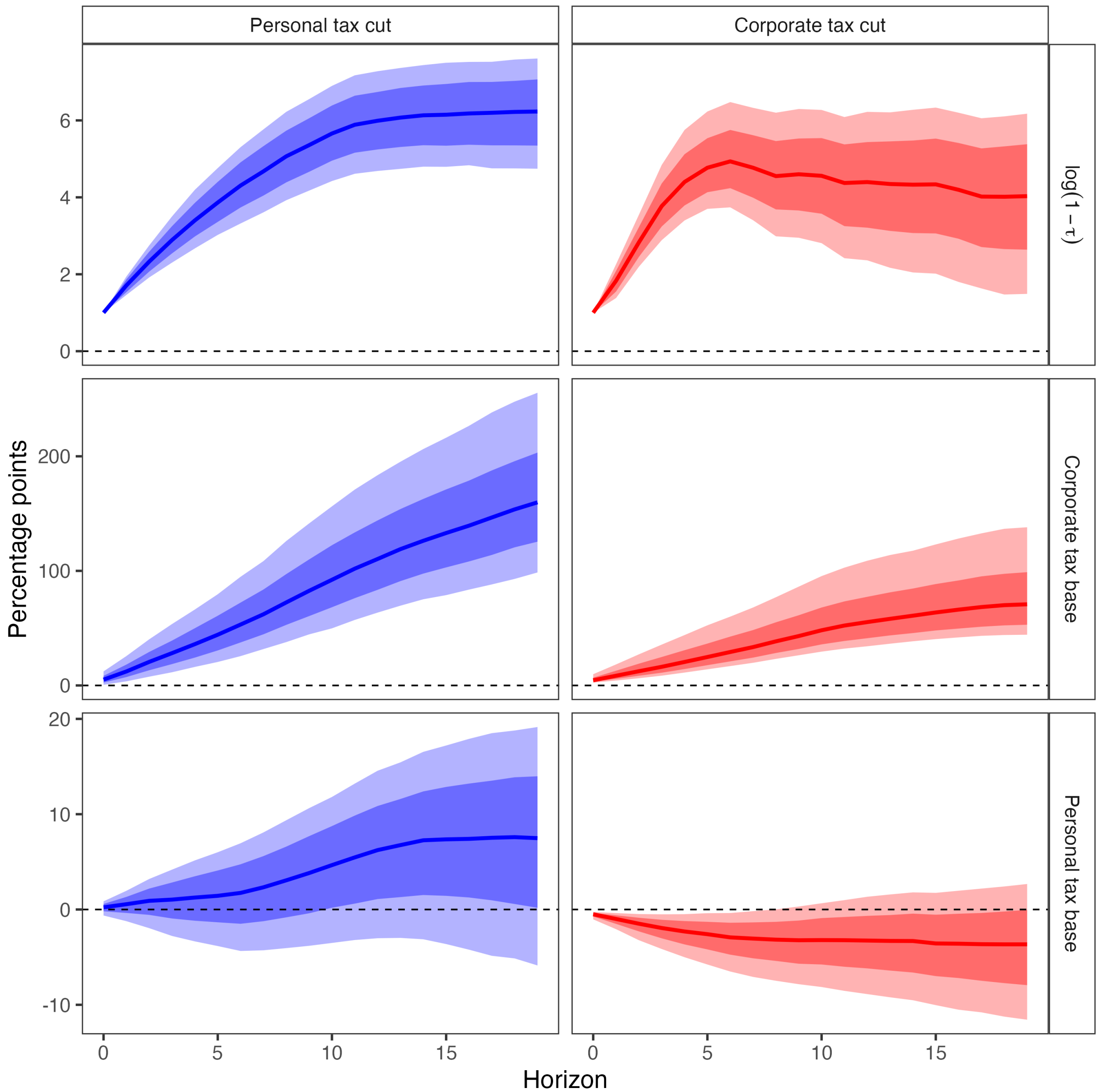
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure E.8: Dynamic Effects of Narrative Tax Shocks on Revenue (Smooth LP)



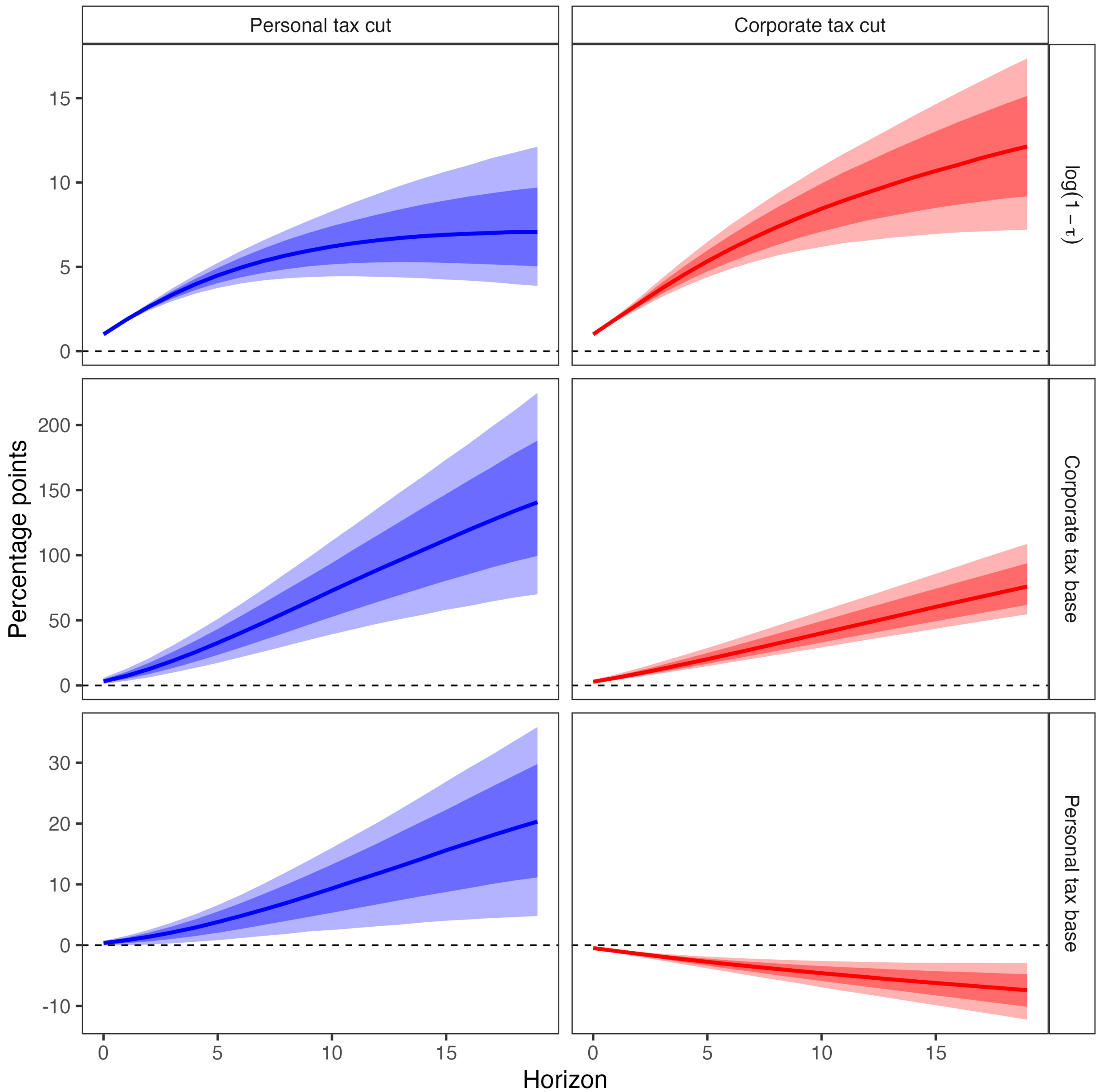
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure E.9: Dynamic Effects of Narrative Tax Shocks on Revenue (Bias-Corrected LP)



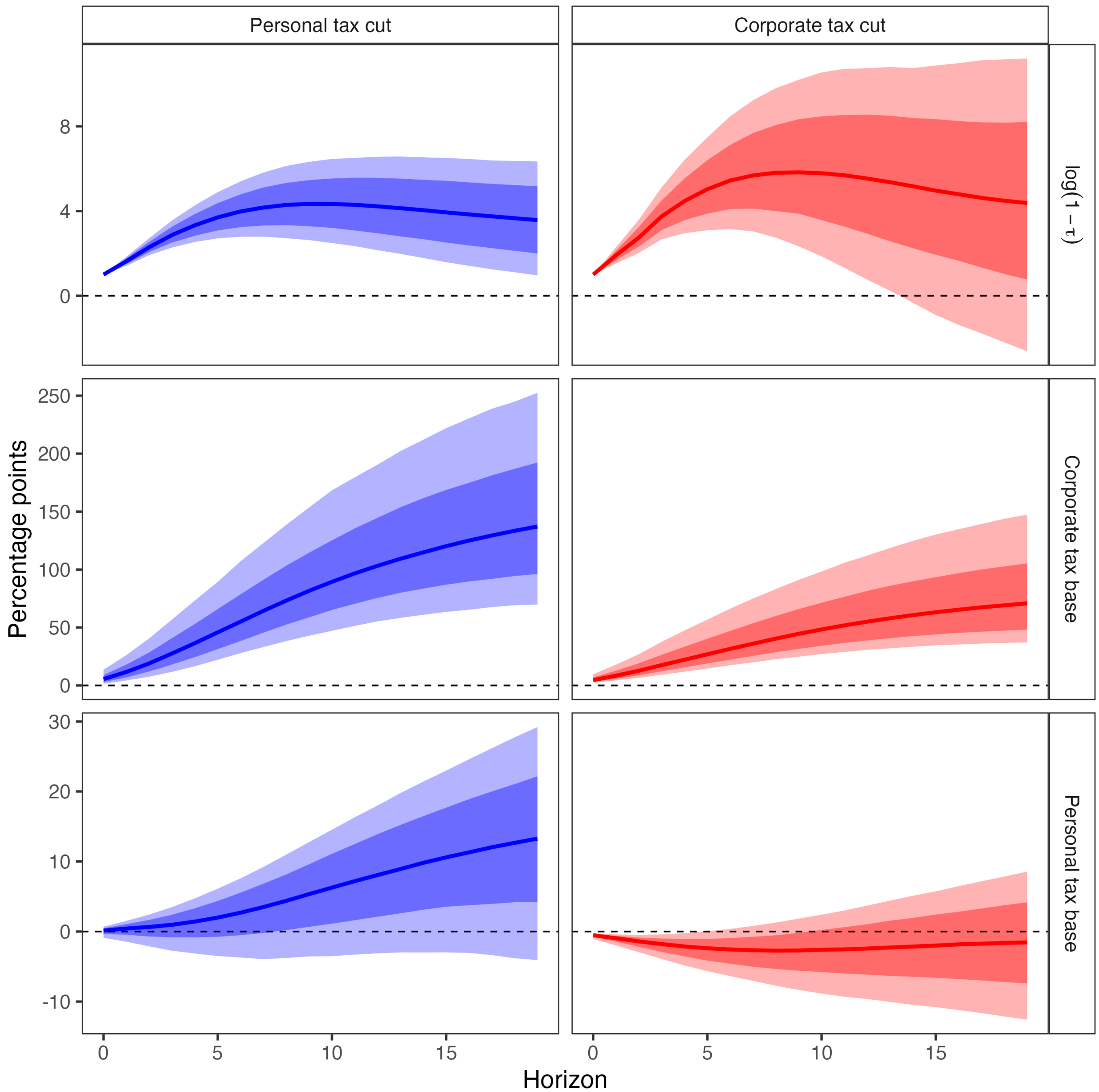
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure E.10: Dynamic Effects of Narrative Tax Shocks on Revenue (Bayesian VAR)



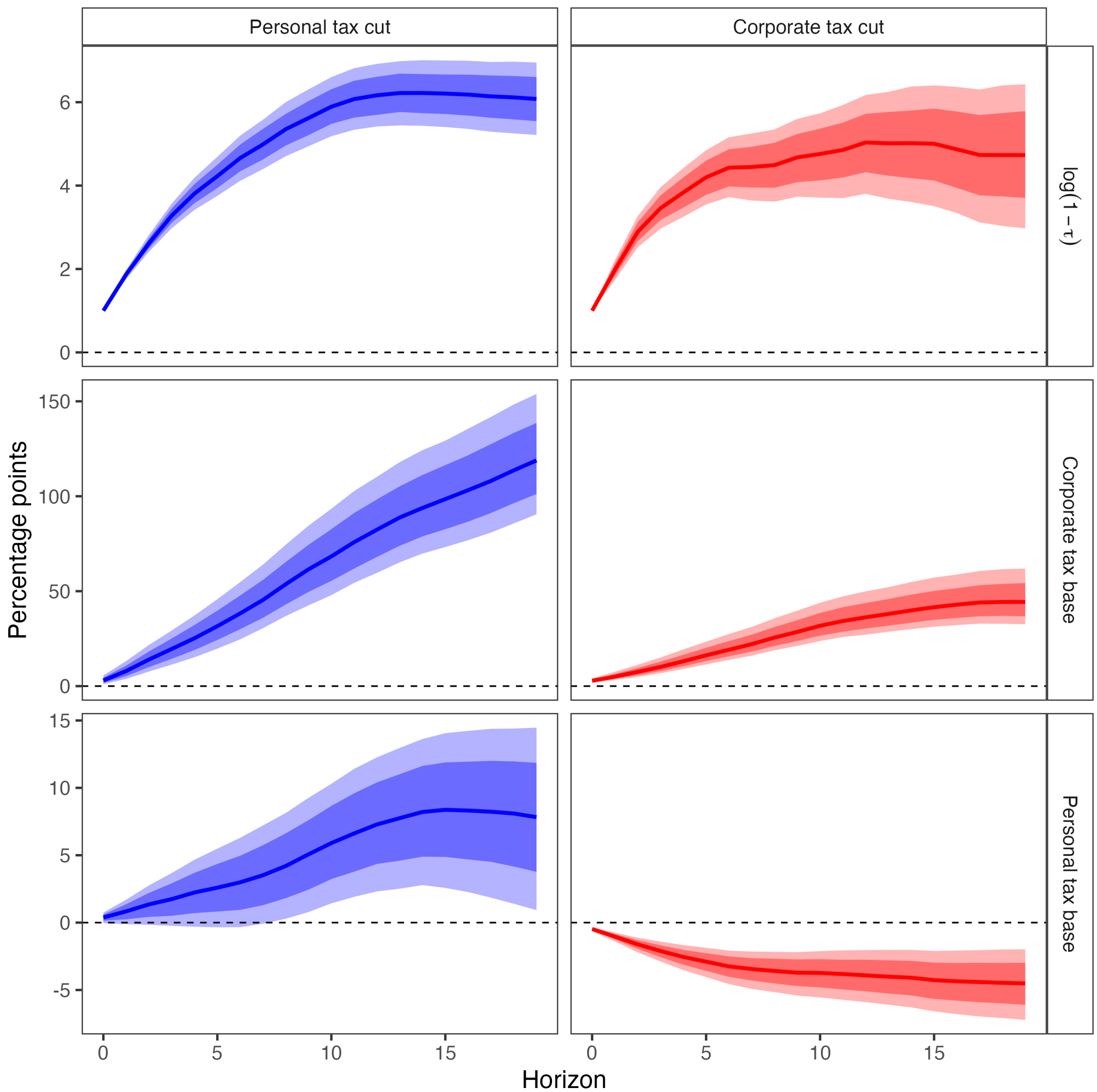
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure E.11: Dynamic Effects of Narrative Tax Shocks on Revenue (Proxy SVAR)



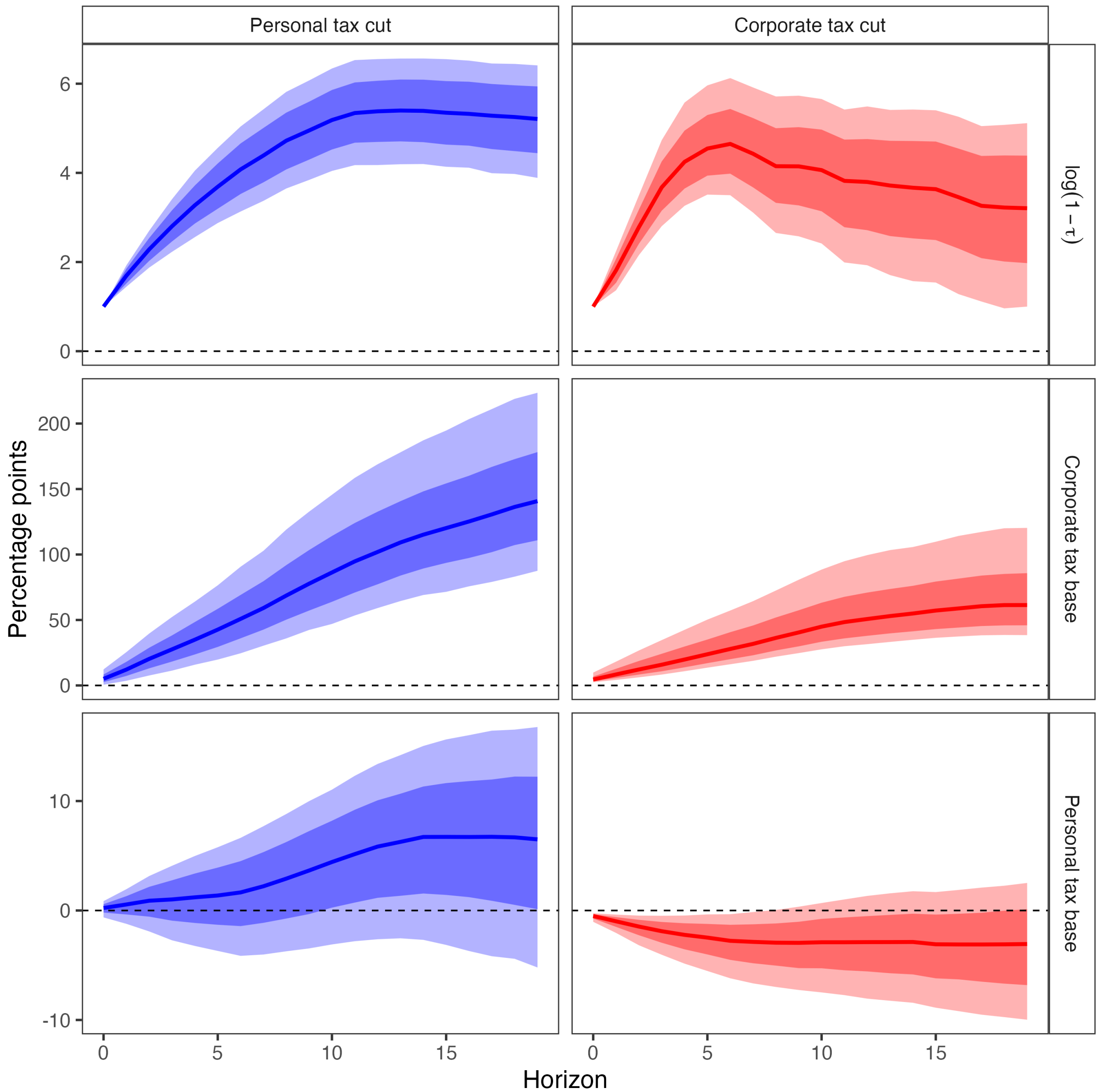
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure E.12: Dynamic Effects of Corporate and Personal Income Tax Cuts on Tax Rates and Bases (BLP with Linear Trend)



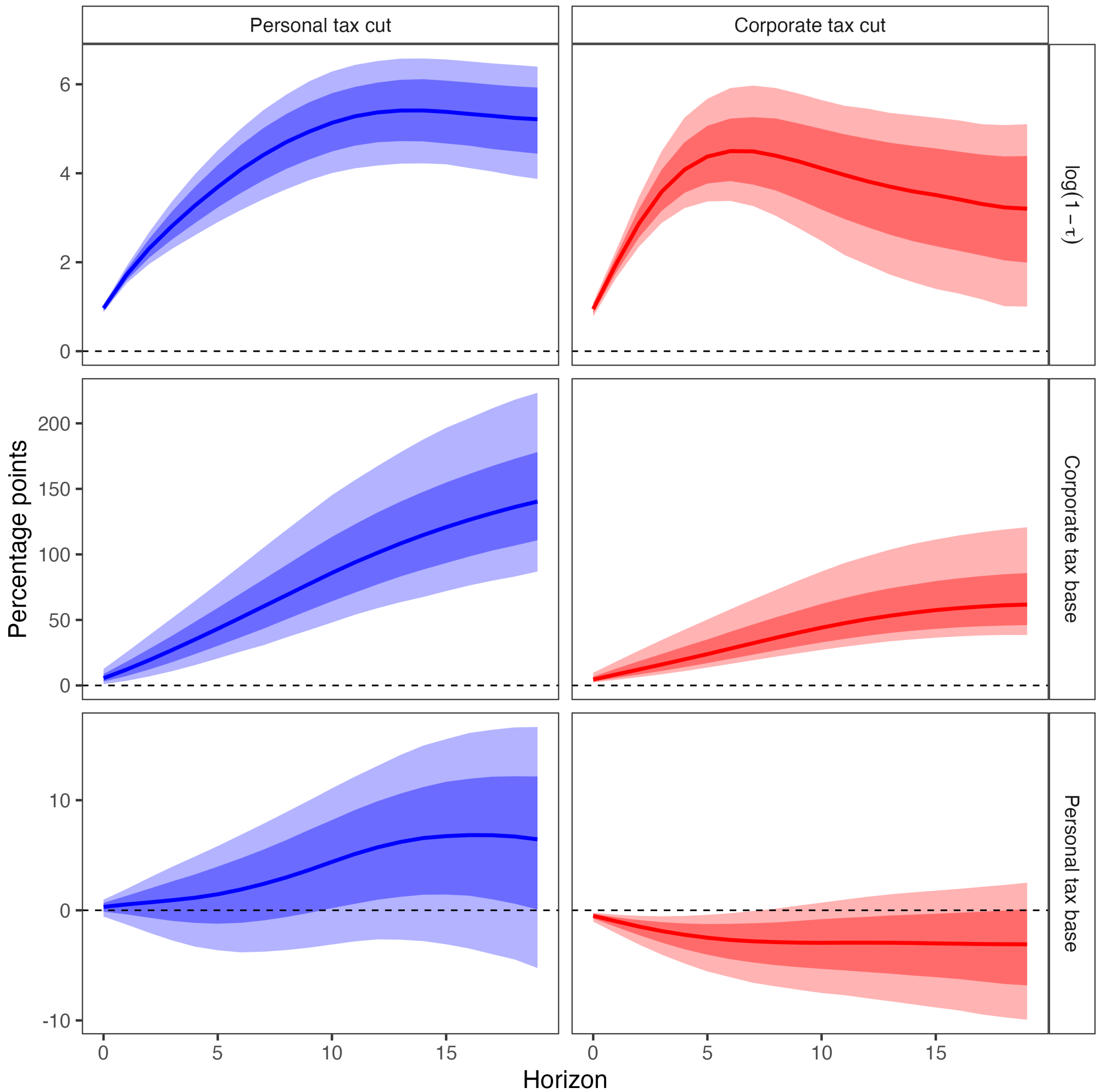
Note: Discounted cumulative impulse responses ($\beta = 0.9926$) to a unit increase in the log retention rate $\log(1-\tau)$. Sample: 1947Q1–2019Q4. Posterior medians with 68% and 90% credible intervals. This specification has a linear trend.

Figure E.13: Dynamic Effects of Narrative Tax Shocks on Revenue (LP with Linear Trend)



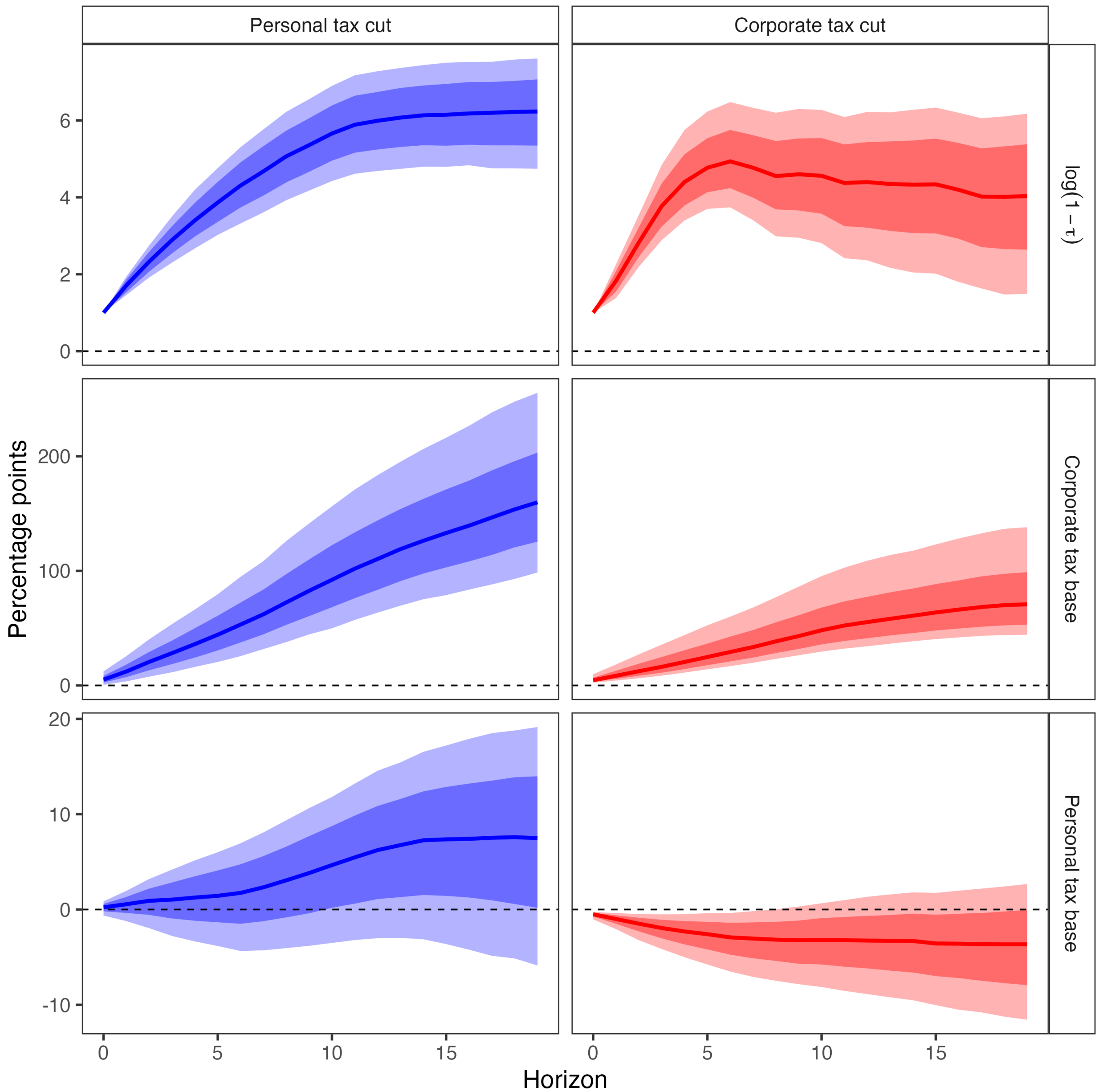
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$. This specification has a linear trend.

Figure E.14: Dynamic Effects of Narrative Tax Shocks on Revenue (Smooth LP with Linear Trend)



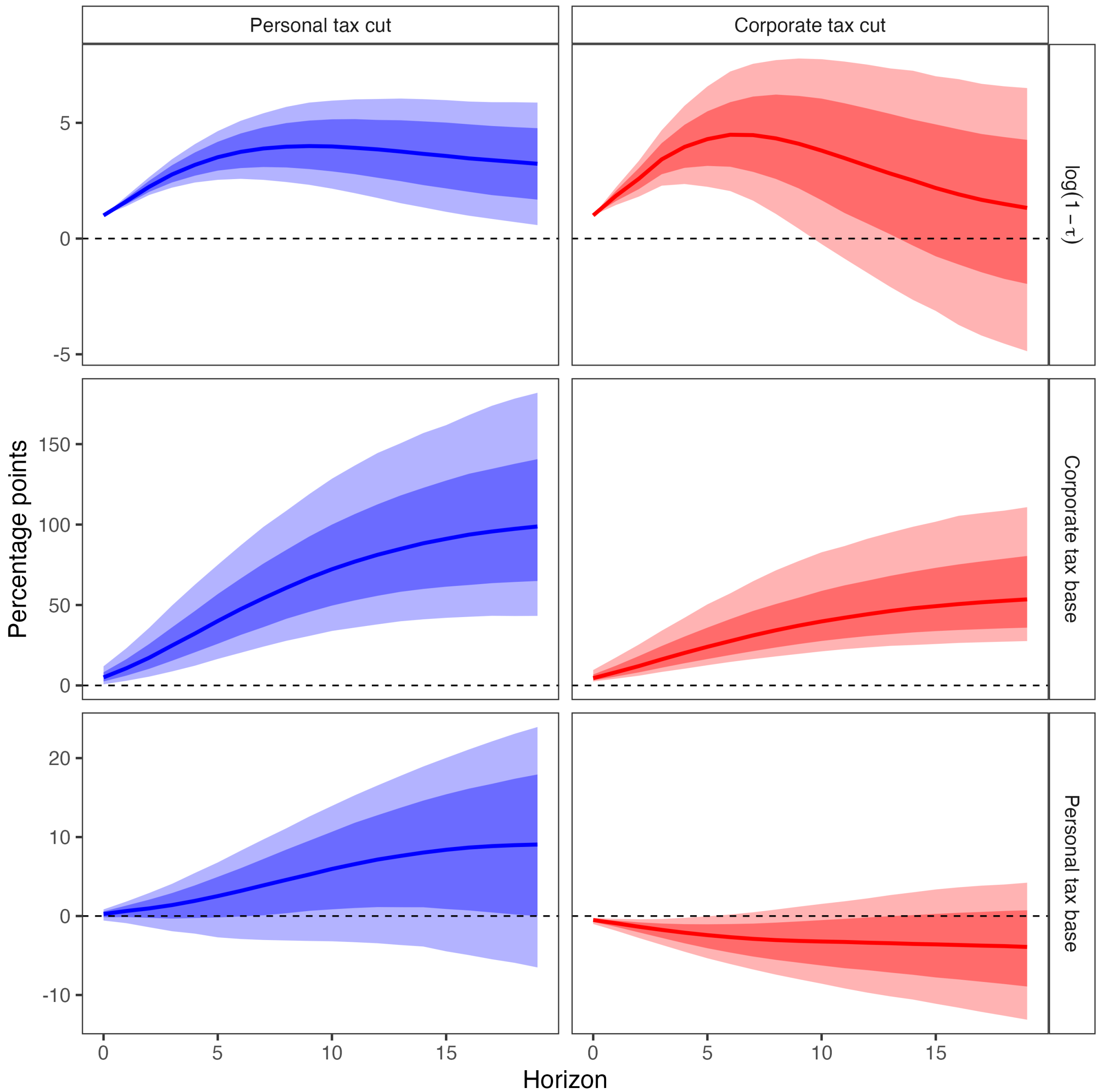
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$. This specification has a linear trend.

Figure E.15: Dynamic Effects of Narrative Tax Shocks on Revenue (Bias-Corrected LP with Linear Trend)



Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$. This specification has a linear trend.

Figure E.16: Dynamic Effects of Narrative Tax Shocks on Revenue (Proxy SVAR with Linear Trend)



Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$. This specification has a linear trend.

Table E.1: Present Value Elasticities of Tax Bases with Respect to Retention Rates (Four Variable System)

	Corporate Tax Shock		Personal Tax Shock	
Method	ε_{CC}	ε_{LC}	ε_{CL}	ε_{LL}
BLP	8.47	-0.34	20.88	1.52
	(5.85, 13.54)	(-0.83, 0.07)	(14.53, 29.60)	(0.39, 2.65)
LP	13.80	-0.14	30.11	1.43
	(8.51, 28.17)	(-1.45, 1.07)	(18.84, 57.05)	(-0.65, 3.86)
SLP	13.88	-0.15	30.06	1.43
	(8.53, 28.03)	(-1.47, 1.05)	(18.79, 57.63)	(-0.68, 3.86)
LP-BC	13.16	-0.19	28.52	1.34
	(8.14, 26.10)	(-1.42, 0.95)	(17.97, 53.30)	(-0.66, 3.65)
BVAR	6.18	-0.60	19.42	2.94
	(3.66, 11.23)	(-1.05, -0.27)	(8.30, 37.45)	(0.68, 7.11)
SVAR-IV	12.54	-0.16	38.27	3.50
	(-58.70, 85.52)	(-5.49, 4.66)	(16.75, 122.56)	(-1.34, 18.70)
With Linear Trend				
BLP	9.44	-0.96	19.73	1.24
	(5.64, 18.78)	(-1.88, -0.43)	(14.16, 27.64)	(0.08, 2.35)
LP	18.18	-0.90	27.04	1.27
	(8.73, 69.03)	(-4.00, 1.30)	(16.11, 48.39)	(-0.85, 3.85)
SLP	18.29	-0.90	26.99	1.26
	(8.72, 69.60)	(-4.01, 1.32)	(16.01, 48.75)	(-0.85, 3.86)
LP-BC	16.89	-0.86	25.65	1.22
	(8.39, 59.63)	(-3.36, 1.05)	(15.41, 45.26)	(-0.81, 3.68)
SVAR-IV	10.48	-0.72	30.34	2.55
	(-110.86, 152.72)	(-12.73, 9.28)	(10.45, 109.97)	(-2.27, 18.84)

Note: Each elasticity ε_{ij} measures the percent change in base i per one percent increase in retention rate $(1 - \tau_j)$. Bayesian specifications have 90% credible intervals in parentheses, while frequentist ones are 90% confidence intervals. This table reports elasticities for the four-variable system comprised solely of tax bases and rates.

F Seven-Variable System Robustness

The baseline specification uses four variables: personal and corporate tax rates and their corresponding bases. This appendix explains why the four-variable specification is appropriate for estimating revenue gradients, and shows robustness to a seven-variable specification that adds output, government spending, and debt.

The revenue gradient in the projection framework is the total derivative of present-value revenue with respect to the tax rate:

$$r_j = \frac{\partial}{\partial \tau_j} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t R_t,$$

where $R_t = \tau_L B_L + \tau_C B_C$ is total income tax revenue. The key object is $\frac{\partial B_k}{\partial \tau_j}$, the total derivative of base k with respect to tax rate j. This total derivative includes all channels through which a tax change affects the base: direct behavioral responses, income shifting between bases, and general equilibrium effects through output, investment, and employment. The theory does not ask for $\frac{\partial B_k}{\partial \tau_j}\big|_Y$, the partial derivative holding output constant. It asks for the full response that a policymaker would observe if they changed the tax rate.

The four-variable specification delivers this total derivative. When we estimate the impulse response of the corporate base to a personal tax shock, the response captures all channels, including any effects operating through GDP. We are not conditioning on GDP, so the full reduced-form relationship between tax shocks and bases is preserved.

A seven-variable specification that includes GDP, government spending, and debt estimates a different object. By including lagged GDP as a regressor in the local projection, we ask: what is the effect of a tax shock on the corporate base, controlling for GDP? This partials out GDP-mediated effects. If a personal tax cut raises GDP, and higher GDP raises corporate profits, then including GDP in the system absorbs some of the variation we want to attribute to the tax shock. The resulting elasticity is a conditional object that understates the total revenue response.

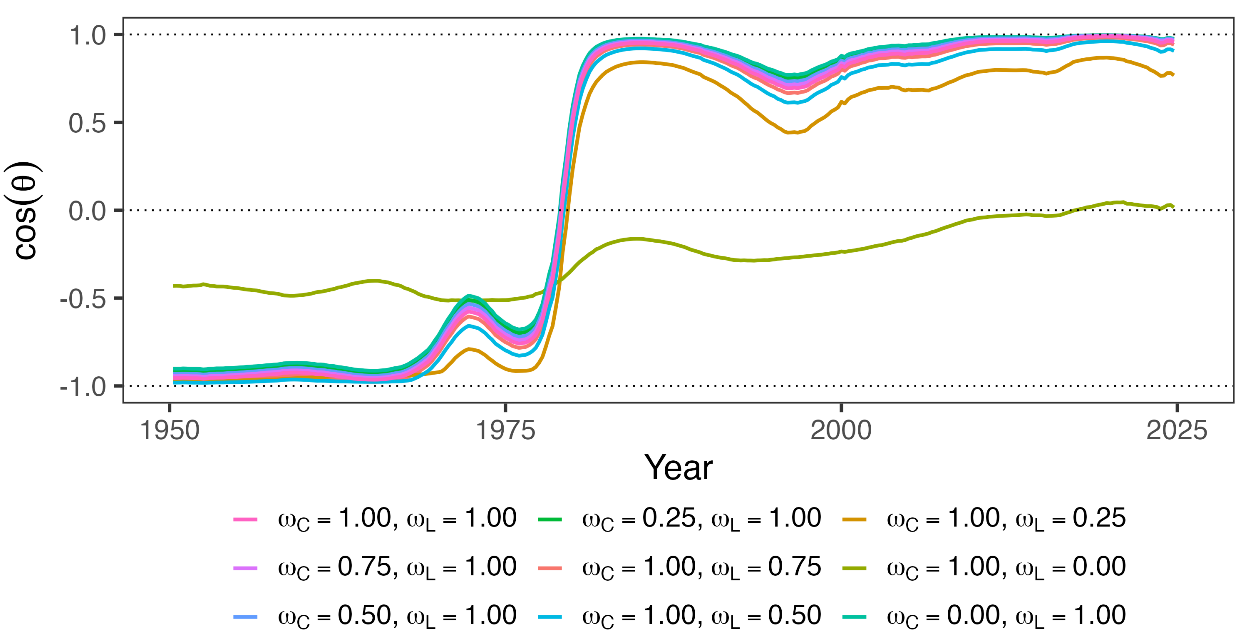
One might worry that omitting GDP leaves the tax shock correlated with other structural disturbances, biasing the estimates. But narrative identification does not require controlling for other variables. The Mertens-Ravn proxy SVAR achieves identification through moment conditions requiring that the narrative instruments correlate with structural tax shocks and are uncorrelated with other structural shocks. These conditions are satisfied by construction: Romer and Romer (2010) selected tax changes legislated for reasons unrelated to current economic conditions. This selection is what ensures exogeneity, not the inclusion of control variables. Adding variables to the VAR does not improve identification; it changes what object is being estimated.

Mertens and Ravn (2013a) included output, government spending, and debt because they were interested in output multipliers and fiscal sustainability. For their research question, the seven-variable specification was appropriate. For the present question—what are the revenue effects of tax changes for evaluating Ramsey alignment?—the total elasticity is required, and the additional variables are unnecessary.

Table F.1 reports elasticities from both specifications. The seven-variable specification yields qualitatively similar point estimates but wider credible intervals, as expected given the additional parameters. The alignment metrics are robust to this choice: $\cos\theta$ in the 2010s exceeds 0.8 under both specifications. The four-variable specification is the appropriate baseline because it estimates the object the theory requires; the seven-variable specification serves as a robustness check demonstrating that results are not sensitive to the conditioning set.

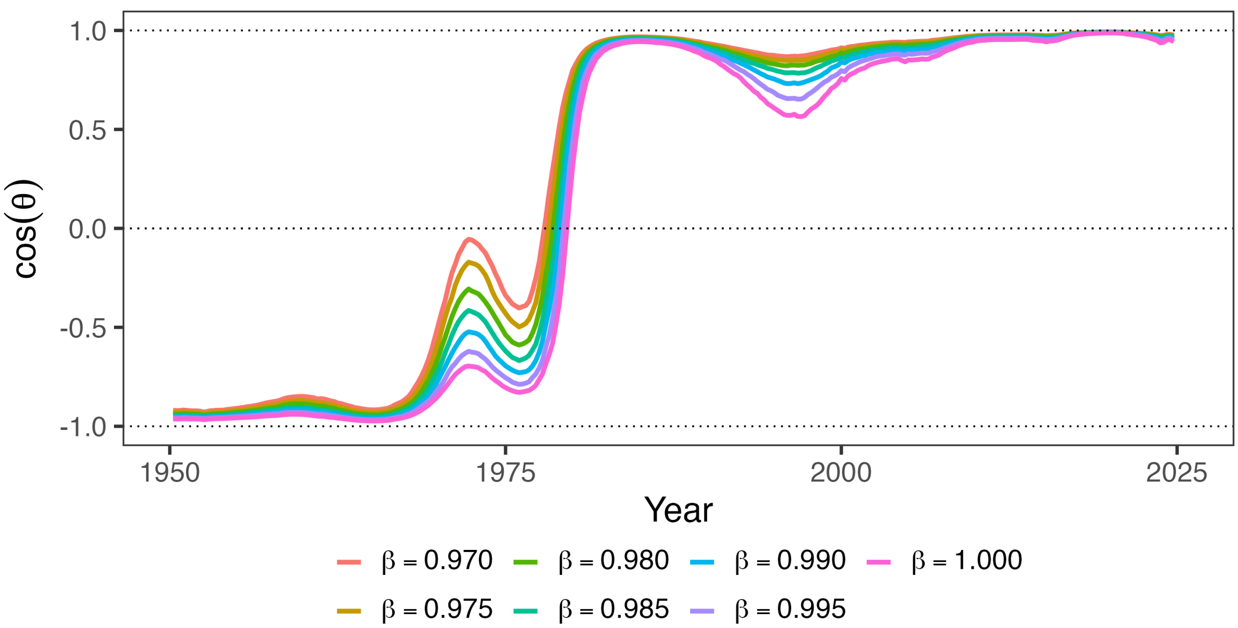
F.1 Robustness of Alignment Statistics

Figure F.1: Robustness Across Welfare Weights



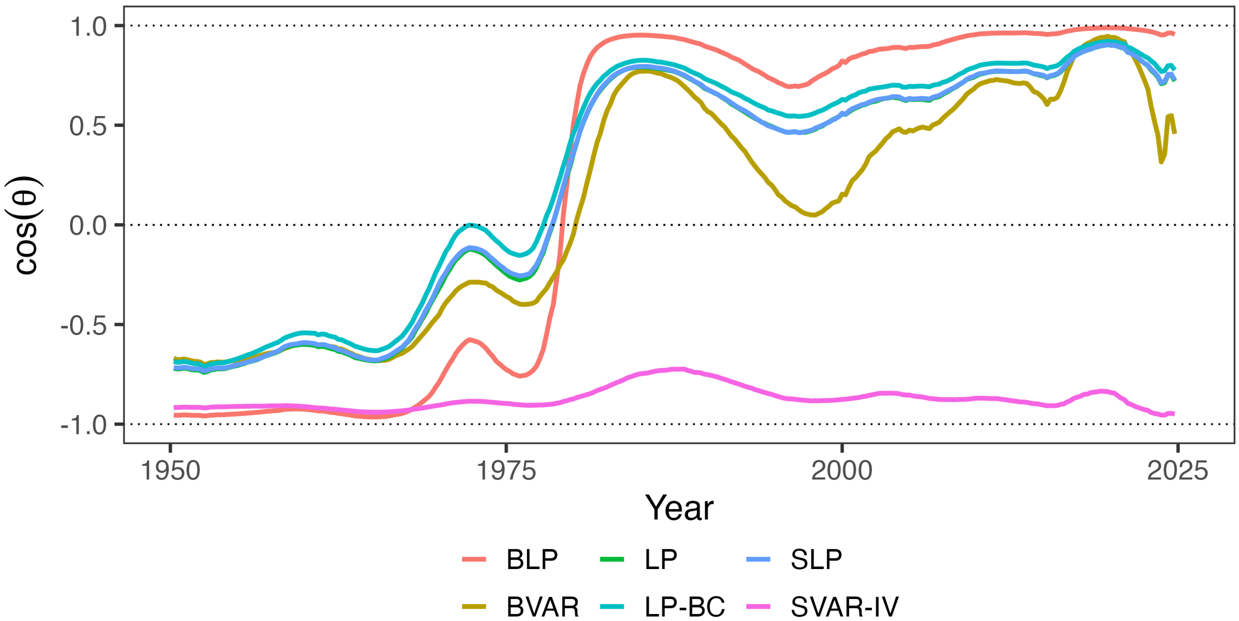
Note: Each panel varies one welfare weight while holding the other at 1. Left panel varies ω_C (weight on corporate taxpayers) with $\omega_L = 1$; right panel varies ω_L (weight on personal taxpayers) with $\omega_C = 1$. The utilitarian baseline has $\omega_C = \omega_L = 1$. Each line corresponds to an estimate with the seven-variable system.

Figure F.2: Robustness Across Discount Factors



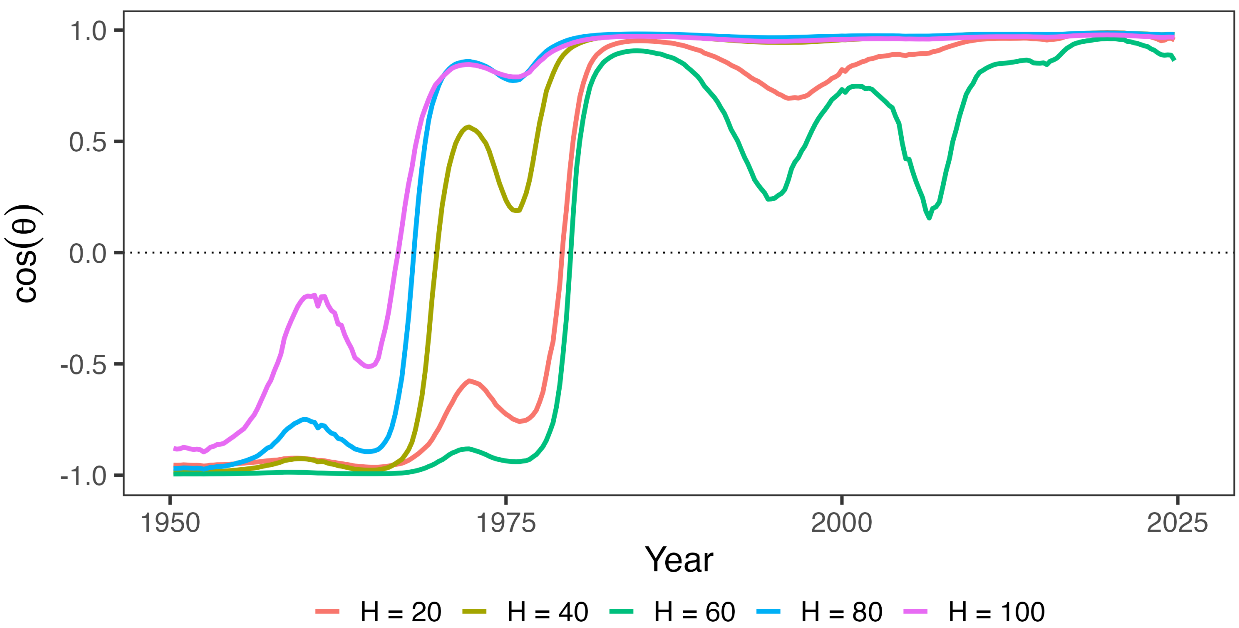
Note: This figure plots alignment $\cos\theta$ across different discount rate regimes for the seven-variable system. Under the baseline, $\beta = 0.9926$.

Figure F.3: Robustness Across Estimation Methodologies



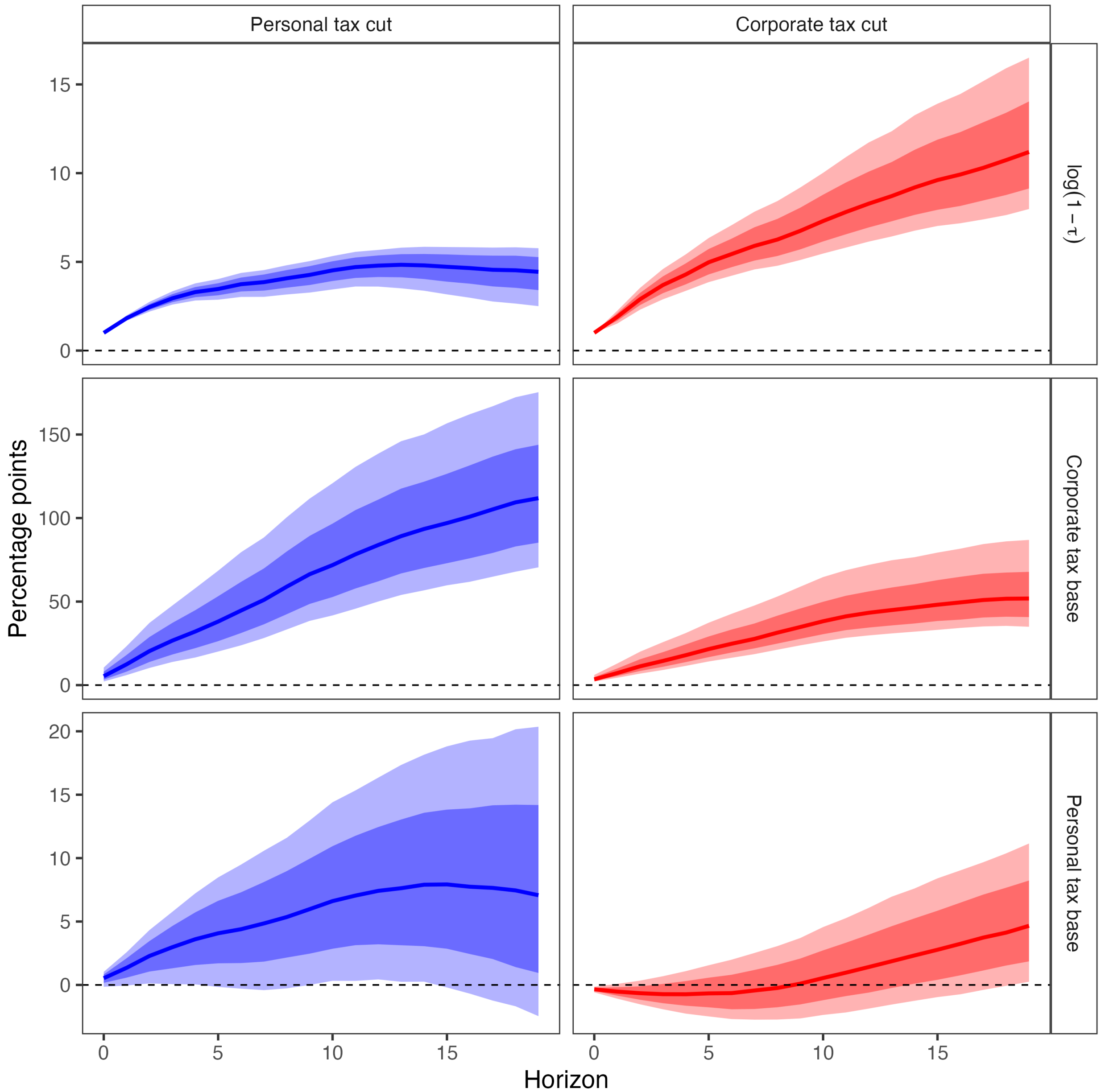
Note: This figure plots alignment $\cos\theta$ across estimation methodologies for the seven-variable system. See Appendix E for further estimation details.

Figure F.4: Robustness Across Horizon Specifications



Note: This figure plots alignment $\cos\theta$ for select models in which vary the horizon length, holding fixed $\beta = 0.9926$. This is for the seven-variable system.

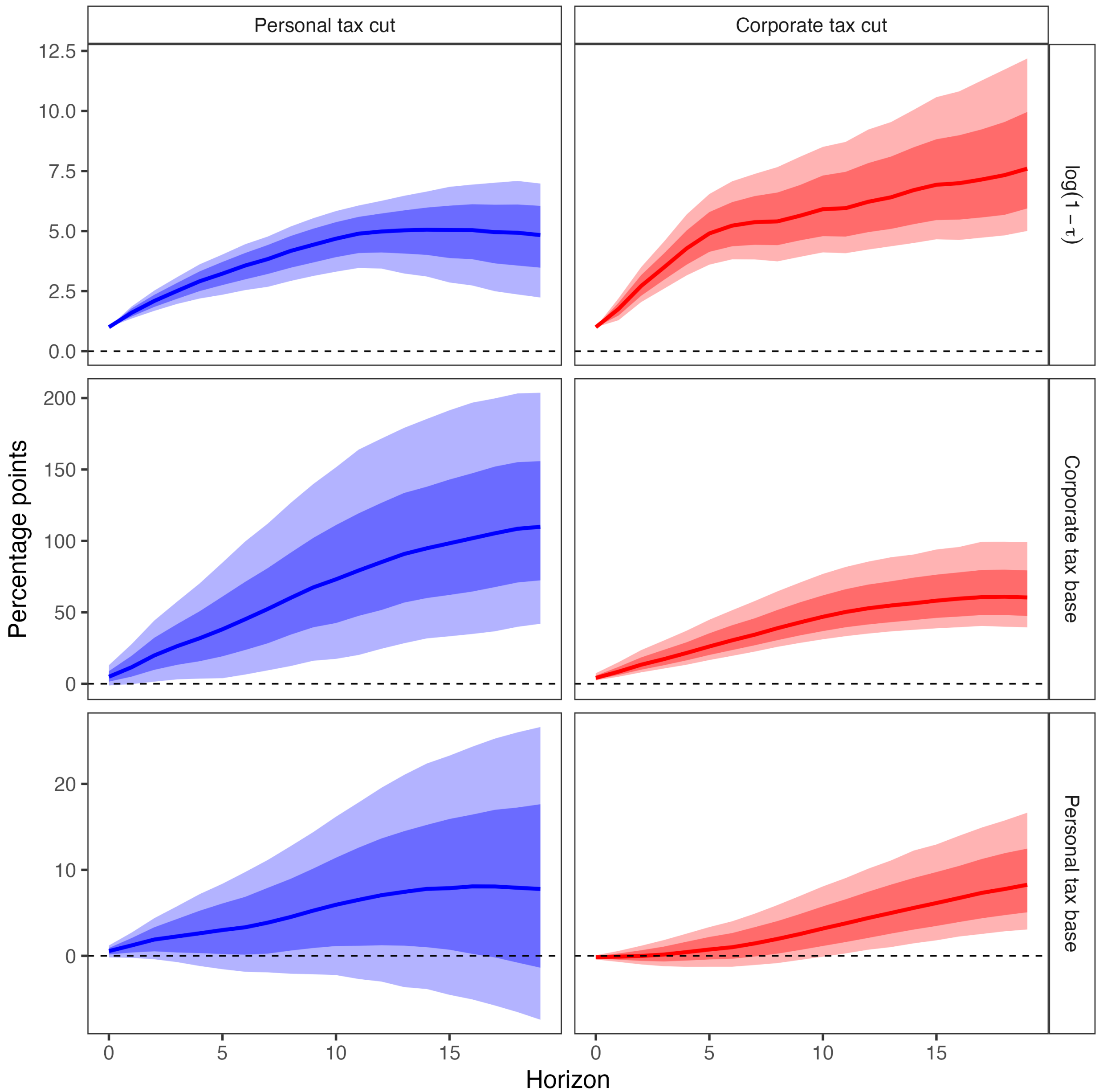
Figure F.5: Dynamic Effects of Corporate and Personal Income Tax Cuts on Tax Rates and Bases



Note: Discounted cumulative impulse responses ($\beta = 0.9926$) to a unit increase in the log retention rate $\log(1-\tau)$. Sample: 1947Q1–2019Q4. Posterior medians with 68% and 90% credible intervals.

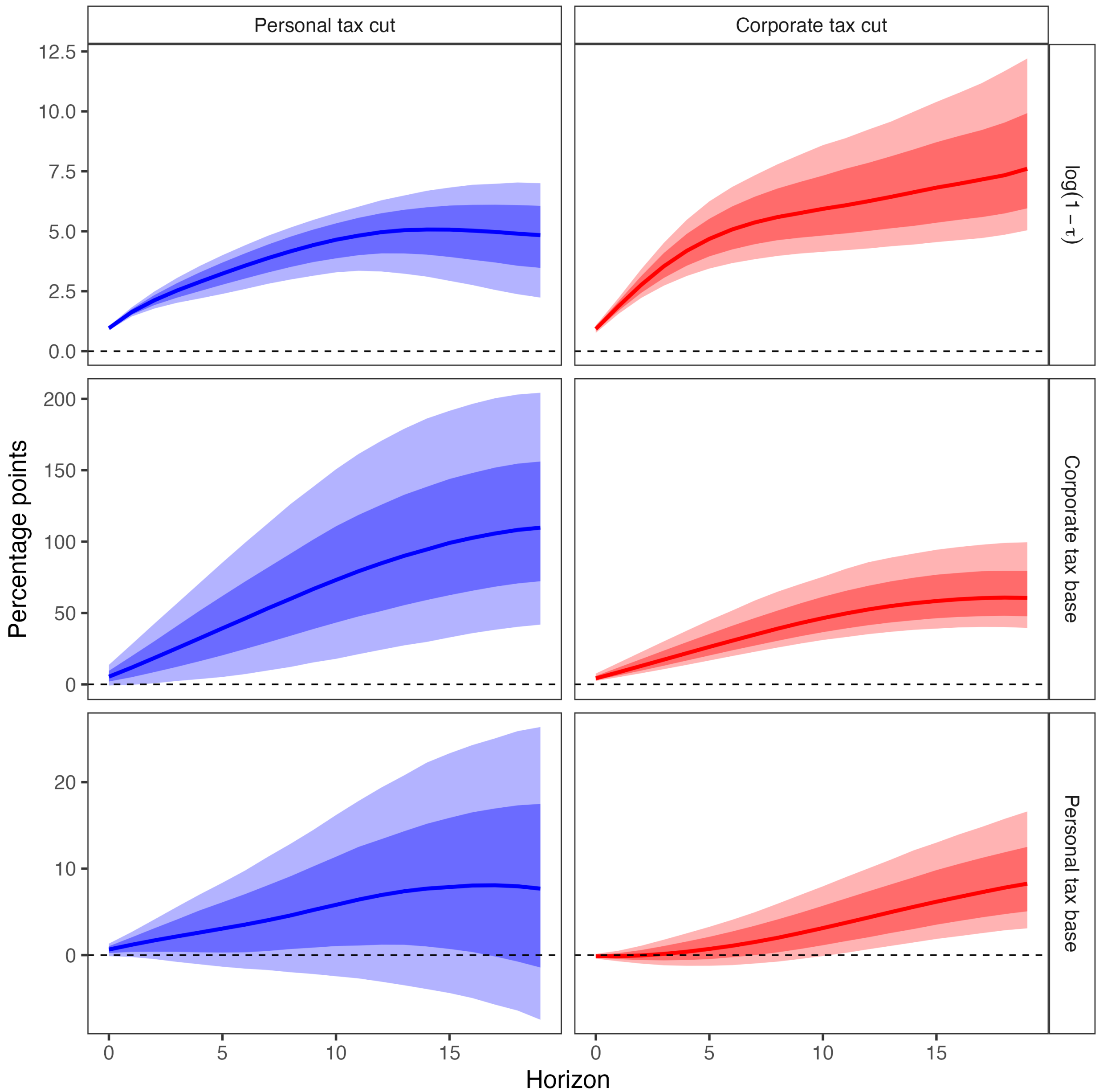
F.2 Underlying Elasticities and CIRFs

Figure F.6: Dynamic Effects of Narrative Tax Shocks on Revenue (LP)



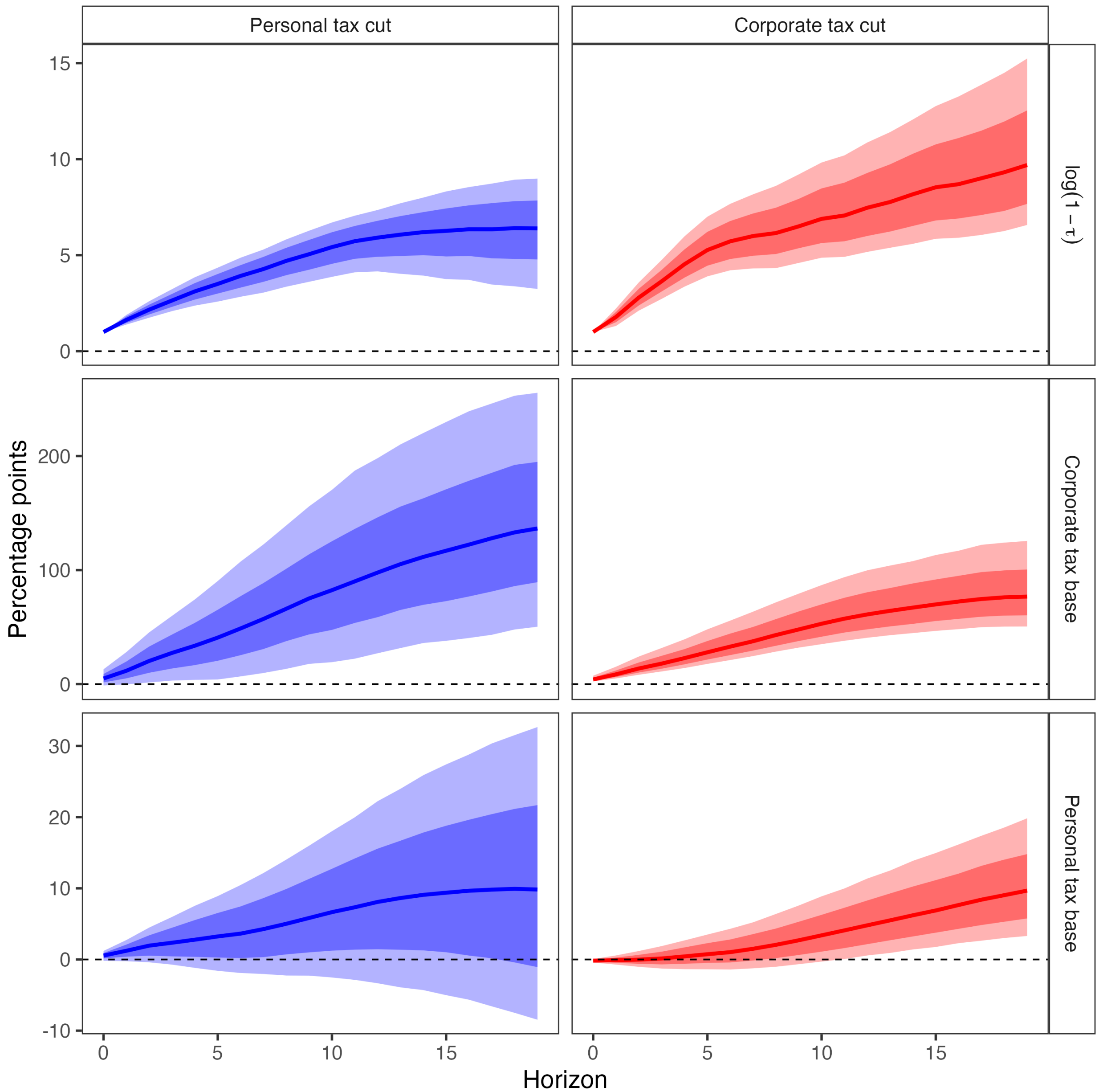
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure F.7: Dynamic Effects of Narrative Tax Shocks on Revenue (Smooth LP)



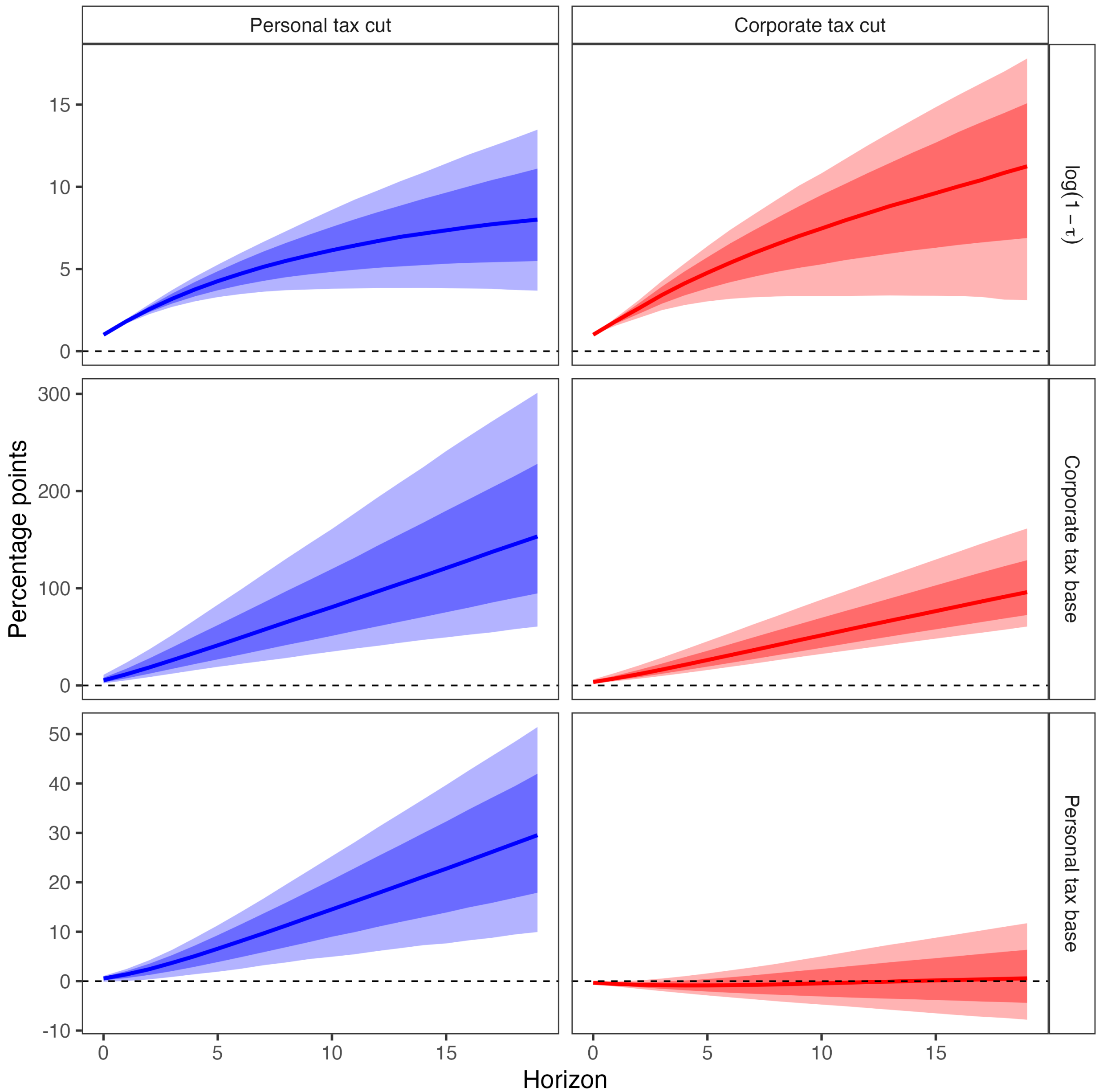
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure F.8: Dynamic Effects of Narrative Tax Shocks on Revenue (Bias-Corrected LP)



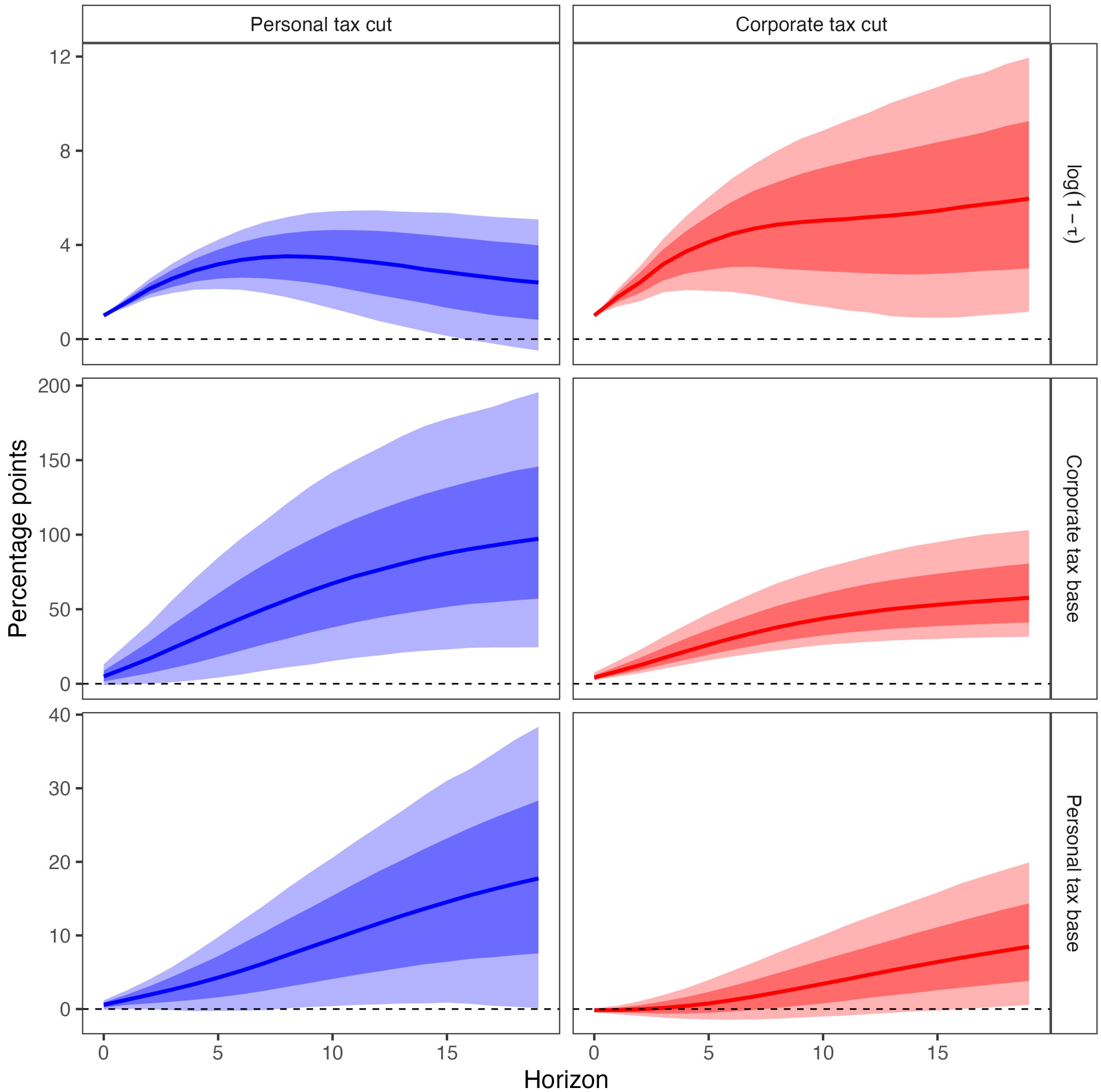
Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure F.9: Dynamic Effects of Narrative Tax Shocks on Revenue (Bayesian VAR)



Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Figure F.10: Dynamic Effects of Narrative Tax Shocks on Revenue (Proxy SVAR)



Note: Dynamic effects of a unit increase in each instrument's average log retention rate from 1947Q1-2019Q4. The response is a discounted cumulative IRF with $\beta = 0.9926$.

Table F.1: Present Value Elasticities of Tax Bases with Respect to Retention Rates (Seven Variable System)

	Corporate Tax Shock		Personal Tax Shock	
Method	ε_{CC}	ε_{LC}	ε_{CL}	ε_{LL}
BLP	4.63	0.42	24.94	1.57
	(3.03, 7.21)	(0.02, 0.97)	(13.52, 59.06)	(-0.57, 6.65)
LP	7.96	1.09	22.41	1.58
	(5.12, 12.96)	(0.39, 2.13)	(5.88, 76.29)	(-1.29, 10.01)
SLP	7.99	1.09	22.32	1.55
	(5.14, 12.96)	(0.39, 2.13)	(5.80, 76.83)	(-1.28, 10.04)
LP-BC	7.94	1.00	21.18	1.52
	(5.19, 12.82)	(0.34, 2.01)	(5.58, 67.70)	(-1.11, 8.39)
BVAR	8.41	0.04	18.47	3.61
	(4.06, 31.51)	(-0.83, 1.65)	(6.62, 55.55)	(0.89, 11.09)
SVAR-IV	9.23	1.35	34.56	5.92
	(3.37, 37.65)	(-0.20, 6.40)	(-87.70, 214.00)	(-24.88, 48.88)

Note: Each elasticity ε_{ij} measures the percent change in base i per one percent increase in retention rate $(1 - \tau_j)$. Bayesian specifications have 90% credible intervals in parentheses, while frequentist ones are 90% confidence intervals. This table reports elasticities for the seven-variable system, which adds government spending, debt, and GDP to the four-variable system.

G Evaluating Deficit-Financed Reforms

This appendix extends the projection framework to evaluate deficit-financed reforms when the repayment rule is partially identified. The approach parallels the classic results in Appendix H: we show that optimal debt policy emerges as a special case of the intertemporal projection problem from Section 3.1, and develop bounds when the repayment path is uncertain.

Debt as Intertemporal Tax Reallocation

The transversality condition from equation (19) reveals that debt financing is equivalent to an intertemporal reallocation of the tax burden. From the perspective of the sequence-space Jacobian introduced in Section 3.1, issuing debt at $t = 0$ creates lower-diagonal entries: future revenue must rise to service the debt. If agents are forward-looking and anticipate the future tax increases required for repayment, this also generates upper-diagonal feedbacks, as current behavior responds to expected future policy. The present-value revenue gradient r_d^{PV} in equation (15) aggregates these intertemporal spillovers, exactly as the dynamic tax gradients aggregate spillovers across instruments and time. This connection clarifies that debt is not a distinct financing instrument but rather a repackaging of future tax obligations.

The welfare consequences of this deferral depend on whether the marginal cost of public funds differs across time. If current taxes are highly distortionary while future taxes are less distortionary, deferring taxation via debt improves welfare by economizing on distortions. Conversely, if future taxes are more distortionary than current taxes, deferring taxation reduces welfare. The optimal financing mix should equalize the welfare cost per dollar of financing across all margins—both across instruments and across time.

The dynamic Ramsey problem with transversality. Imposing the transversality condition, the government chooses the tax path $\{\boldsymbol{\tau}_t\}_{t=0}^{\infty}$ to maximize:

$$\max_{\{\boldsymbol{\tau}_t\}} \sum_{t=0}^{\infty} \beta^t W_t(\boldsymbol{\tau}_t) \quad \text{subject to} \quad \sum_{t=0}^{\infty} \gamma^t R_t(\boldsymbol{\tau}_t) = \sum_{t=0}^{\infty} \gamma^t G_t + D_0, \tag{A.18}$$

where $\beta = (1+\rho)^{-1}$ is the social discount factor. The debt path $\{D_t\}$ is residual, determined by the budget identity $D_t = D_{t-1}(1 + r_{\mathrm{debt}}) + G_t - R_t$. Debt is not a choice variable; the transversality condition eliminates it as a degree of freedom. The only choice is the tax path.

The Lagrangian is:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t W_t(\boldsymbol{\tau}_t) - \lambda \left[\sum_{t=0}^{\infty} \gamma^t R_t(\boldsymbol{\tau}_t) - C \right], \tag{A.19}$$

where λ is the multiplier on the intertemporal budget constraint and $C = \sum_t \gamma^t G_t + D_0$ is the right-hand side. The first-order condition for $\tau_{i,t}$ is:

$$\beta^t \frac{\partial W_t}{\partial \tau_{i,t}} = \lambda \gamma^t \frac{\partial R_t}{\partial \tau_{i,t}}. \tag{A.20}$$

Rearranging:

$$\mathrm{MCPF}_{i,t} = \lambda \left(\frac{1+\rho}{1+r_{\mathrm{debt}}} \right)^t \tag{A.21}$$

At the Ramsey optimum, MCPFs must be equalized after adjusting for the wedge between the social discount rate and the government's borrowing rate. When $r_{\mathrm{debt}} = \rho$, MCPFs should be constant across time—this recovers the tax smoothing result derived in Appendix H.2. When

$r_{\text{debt}} \neq \rho$, optimal MCPFs tilt over time. If $r_{\text{debt}} < \rho$ (as in recent decades), the fiscal constraint underweights future costs relative to social welfare. MCPFs should fall over time at rate $(1+\rho)/(1+r_{\text{debt}}) > 1$, favoring back-loaded taxation.

Dual discounting and the projection formula. For local reforms around baseline $\bar{\boldsymbol{\tau}} = (\bar{\tau}_{1,0}, \ldots, \bar{\tau}_{n,T})$, the optimal direction satisfies:

$$d\boldsymbol{\tau}^* = -P_{\perp\tilde{r}^{\text{PV}}} g^{\text{PV}}, \tag{A.22}$$

where $g^{\text{PV}} = (\beta^0 g_0, \beta^1 g_1, \ldots, \beta^T g_T)$ stacks welfare gradients with social discounting and $\tilde{r}^{\text{PV}} = (\gamma^0 r_0, \gamma^1 r_1, \ldots, \gamma^T r_T)$ stacks revenue gradients with fiscal discounting. The projection operator $P_{\perp\tilde{r}^{\text{PV}}} = I - \tilde{r}^{\text{PV}}(\tilde{r}^{\text{PV}\top}\tilde{r}^{\text{PV}})^{-1}\tilde{r}^{\text{PV}\top}$ enforces the fiscal constraint: reforms must satisfy $(\tilde{r}^{\text{PV}})^\top d\boldsymbol{\tau} = 0$ in present value, discounted at the government's borrowing rate.

The projection formula from Section 2.2 carries over with one critical modification: the revenue gradient must be discounted at γ, not β. The welfare gradient uses social discounting because we evaluate welfare from society's perspective; the revenue constraint uses fiscal discounting because the government's ability to borrow and repay is determined by market interest rates, not social preferences. When $r_{\text{debt}} = \rho$, the two coincide and the formula simplifies to $d\boldsymbol{\tau}^* = -P_{\perp r^{\text{PV}}} g^{\text{PV}}$ from Section 3.1. When they diverge, the dual-discounting structure in equation (A.22) is required.

This formulation does not treat debt as a separate policy instrument. There is no "MCPF of debt" in the sense of a standalone financing margin. Debt is the residual determined by the tax path. The question "should we issue more debt?" is equivalent to "should we lower taxes today and raise them tomorrow?" The answer depends on the relative MCPFs across time, adjusted for the interest rate wedge in equation (A.21).

Evaluating Incomplete Reforms

The projection formula delivers the globally optimal tax path $\{\boldsymbol{\tau}_t^*\}$ satisfying the intertemporal budget constraint. In practice, we rarely observe complete reform paths. The Tax Cuts and Jobs Act specified tax changes at $t = 0$—corporate rate cuts of 14 percentage points and personal rate cuts of 2.6 percentage points—but left the future adjustment path implicit. The deficit created by these cuts must eventually be closed, but the timing and composition of that closure are uncertain. Evaluating such incomplete reforms requires comparing them to feasible alternatives.

The comparison. Consider two reform paths:

Reform A (deficit-financed, as implemented): At $t = 0$, cut taxes by $d\boldsymbol{\tau}_0^A = (\Delta\tau_C, \Delta\tau_L)$, creating a revenue loss. In future periods, raise taxes to satisfy the transversality condition.

Reform B (revenue-neutral counterfactual): At $t = 0$, cut taxes by $d\boldsymbol{\tau}_0^B = (\Delta\tau_C, \Delta\tau_L + \text{offset})$, where the offset is chosen to satisfy $r_0^\top d\boldsymbol{\tau}_0^B = 0$. No future adjustment is needed.

The welfare comparison reduces to: which path has lower present-value cost? Reform A provides larger immediate tax relief but requires costly future repayment. Reform B provides smaller immediate relief but avoids future distortions. The trade-off depends on the relative MCPFs: if current taxes are highly distortionary and future taxes are cheap, Reform A dominates. If the reverse holds, Reform B dominates.

The repayment problem. Reform A creates a fiscal gap. The present-value revenue loss (discounted at the fiscal rate γ) is:

$$\Delta R^{\text{PV}} = \sum_{h=0}^{H} \gamma^h \left[r_{C,h} \Delta\tau_C(h) + r_{L,h} \Delta\tau_L(h) \right], \tag{A.23}$$

where $\Delta\tau_i(h) = \Delta\tau_i(0) \times \text{IRF}(\tau_i, h)$ captures the persistence of tax rates.[10]

This gap must be closed by raising taxes in future periods. Define a repayment rule W that specifies the path $\{\Delta\tau_{k,t}\}_{t\geq 1}$ satisfying:

$$\sum_{t\geq 1} \sum_{k} \gamma^t r_{k,t} \Delta\tau_{k,t} = \Delta R^{\text{PV}}. \tag{A.24}$$

The present-value welfare cost of this repayment is:

$$\text{Welfare cost}(W) = \sum_{t\geq 1} \sum_{k} \beta^t g_{k,t} \Delta\tau_{k,t}. \tag{A.25}$$

The marginal cost of public funds for deferral under rule W is:

$$\text{MCPF}^{\text{PV}}(W) = \frac{\text{Welfare cost}(W)}{\Delta R^{\text{PV}}} = \sum_{t,k} \omega_{t,k}(W) \left(\frac{\beta^t}{\gamma^t} \right) \text{MCPF}_{k,t}, \tag{A.26}$$

where $\omega_{t,k}(W) = \gamma^t r_{k,t} \Delta\tau_{k,t} / \Delta R^{\text{PV}}$ are fiscal-PV-normalized repayment shares (summing to one across (t, k)) and $\text{MCPF}_{k,t} = g_{k,t}/r_{k,t}$ is the standard MCPF for tax k at time t.

Equation (A.26) reveals that the MCPF of deferral is a weighted average of future tax MCPFs, where the weights depend on both the repayment timing and the discount wedge β^t/γ^t. When $r_{\text{debt}} < \rho$ (as in recent decades), $\beta^t/\gamma^t < 1$ for $t > 0$, mechanically lowering the effective MCPF of

10. This is a local approximation using baseline revenue gradients and policy persistence. The preferred approach for Section 4 directly estimates revenue IRFs for the reform package: $\Delta R^{\text{PV}} = \sum_h \gamma^h \cdot \text{IRF}_{\text{Rev}}(h|\text{TCJA shock})$. Equation (A.23) is valid locally but does not capture general equilibrium revenue responses to the full reform.

deferral relative to the case where $r_{\text{debt}} = \rho$. Intuitively, low government borrowing rates make deferral cheaper because the fiscal constraint underweights future costs relative to social welfare accounting.

Bounds without specifying the repayment rule. The repayment rule W is unobserved. Equation (A.26) expresses $\text{MCPF}^{\text{PV}}(W)$ as a convex combination of atoms:

$$a_{t,k} := \left(\frac{\beta^t}{\gamma^t}\right) \text{MCPF}_{k,t}. \tag{A.27}$$

When discount rates are constant over time, this simplifies to $a_{t,k} = \left(\frac{1+r_{\text{debt}}}{1+\rho}\right)^t \text{MCPF}_{k,t}$.

Since the weights $\omega_{t,k}(W)$ are non-negative and sum to one, $\text{MCPF}^{\text{PV}}(W)$ must lie within the convex hull of these atoms. For any plausible repayment horizon $\mathcal{T} = [t_1, t_2]$ and set of instruments $\mathcal{K}$, this implies:

$$\min_{t\in\mathcal{T},k\in\mathcal{K}} a_{t,k} \;\leq\; \text{MCPF}^{\text{PV}}(W) \;\leq\; \max_{t\in\mathcal{T},k\in\mathcal{K}} a_{t,k}. \tag{A.28}$$

These bounds allow robust evaluation without specifying the repayment rule. Let j denote the tax that would be raised under a revenue-neutral alternative (typically the less distortionary instrument at $t = 0$, chosen to minimize the MCPF of immediate financing). Then:

Case 1: Deferral robustly justified. If $\max_{t,k} a_{t,k} < \text{MCPF}_{j,0}$, then $\text{MCPF}^{\text{PV}}(W) < \text{MCPF}_{j,0}$ for all possible repayment rules. Deficit financing is unambiguously superior to revenue-neutral financing, regardless of how or when the debt is repaid.

Case 2: Deferral robustly dominated. If $\min_{t,k} a_{t,k} > \text{MCPF}_{j,0}$, then $\text{MCPF}^{\text{PV}}(W) > \text{MCPF}_{j,0}$ for all possible repayment rules. Revenue-neutral financing is unambiguously superior, regardless of repayment details.

Case 3: Inconclusive. If $\min_{t,k} a_{t,k} < \text{MCPF}_{j,0} < \max_{t,k} a_{t,k}$, the answer depends on the actual repayment rule. The identified range is $[\min a_{t,k}, \max a_{t,k}]$. If the range is narrow, the conclusion is approximately robust. If the range is wide, specifying or estimating the repayment rule W becomes necessary for a definitive evaluation.

This bounding approach follows directly from the structure of the problem. It requires no auxiliary assumptions about repayment beyond the horizon $\mathcal{T}$ and the set of feasible instruments $\mathcal{K}$. All primitives—$\text{MCPF}_{k,t}$, β, γ—are estimated from the empirical framework in Section 4.

H Classic Optimal Tax Results as Projections

This appendix demonstrates that the projection framework nests several canonical results from optimal tax theory. While the main text focuses on policy evaluation, these derivations show the framework is consistent with standard normative prescriptions under various structural assumptions. The first two subsections restate the Ramsey and Barro conditions in the notation and geometry of this paper. The subsequent subsections show how the Atkinson–Stiglitz theorem and the zero-capital-tax result can be interpreted as special configurations of the welfare and revenue gradients under their usual assumptions.

H.1 The Ramsey Inverse Elasticity Rule

The Ramsey problem seeks the tax structure that minimizes welfare costs subject to raising a fixed revenue requirement. At the optimum, the first-order conditions imply that marginal costs of public funds are equalized across instruments. The projection framework recovers this result and clarifies its geometric interpretation.

Proposition 3 (Ramsey as Perfect Alignment)**.** *At a Ramsey optimum $\bar{\boldsymbol{\tau}}$, the welfare and revenue gradients are perfectly aligned:*

$$g(\bar{\boldsymbol{\tau}}) = \mu r(\bar{\boldsymbol{\tau}})$$

for some scalar $\mu > 0$. Equivalently, $\cos\theta = 1$ and the alignment is perfect.

Proof. The Ramsey problem is:

$$\min_{\boldsymbol{\tau}} W(\boldsymbol{\tau}) \quad \text{subject to} \quad R(\boldsymbol{\tau}) = \bar{R}.$$

The Lagrangian is $\mathcal{L} = W(\boldsymbol{\tau}) - \mu[R(\boldsymbol{\tau}) - \bar{R}]$. The first-order condition is:

$$\frac{\partial W}{\partial \tau_i} = \mu \frac{\partial R}{\partial \tau_i} \quad \text{for all } i,$$

which in vector form is $g(\bar{\boldsymbol{\tau}}) = \mu r(\bar{\boldsymbol{\tau}})$. Since g and r point in the same direction, the angle between them is zero: $\cos\theta = 1$. □

When gradients are perfectly aligned, the projection $P_{\perp r} g = 0$ is the zero vector: there is no component of the welfare gradient perpendicular to revenue. All welfare improvements require changing total revenue. This is the geometric interpretation of "no further revenue-neutral gains are available."

The Ramsey optimum also implies MCPF equalization. From $g_i = \mu r_i$, we have:

$$\text{MCPF}_i = \frac{g_i}{r_i} = \mu \quad \text{for all } i.$$

All MCPFs equal the Lagrange multiplier μ, the shadow cost of the revenue constraint.

The inverse elasticity rule. Under additional structure, the Ramsey formula yields the classic inverse elasticity result. Suppose:

1. Utilitarian welfare: $g_i = B_i$ (the tax base).
2. Ad valorem taxation with instrument-specific rates τ_i.
3. No cross-base spillovers: $\Lambda_i = 0$ (partial equilibrium).

Then from equation (3), the revenue gradient is:

$$r_i = B_i \left[1 - \frac{\tau_i}{1-\tau_i}\varepsilon_{ii}\right].$$

MCPF equalization $g_i/r_i = \mu$ implies:

$$\frac{B_i}{B_i\left[1 - \frac{\tau_i}{1-\tau_i}\varepsilon_{ii}\right]} = \mu \quad \Rightarrow \quad \frac{\tau_i}{1-\tau_i}\varepsilon_{ii} = 1 - \frac{1}{\mu}.$$

Since the right-hand side is constant across i, we have $\frac{\tau_i}{1-\tau_i}\varepsilon_{ii} = \text{constant}$. For small tax rates, this implies:

$$\tau_i \propto \frac{1}{\varepsilon_{ii}}.$$

Taxes should be inversely proportional to elasticities. Highly elastic bases receive low taxes; inelastic bases receive high taxes. This is Ramsey's (1927) classic result, recovered as the shadow of perfect alignment under no cross-base spillovers.

H.2 Barro Tax Smoothing

Barro (1979) shows that when the government's borrowing rate equals the social discount rate, optimal tax policy smooths distortions over time. This result emerges directly from the intertemporal projection framework in Section 3.1.

Proposition 4 (Barro as Temporal MCPF Equalization)**.** *Consider the dynamic Ramsey problem from equation* (A.18)*. Suppose (1) the government's borrowing rate equals the social discount rate:*

$r_{debt} = \rho$, so $\gamma = \beta$; and (2) Tax bases and behavioral responses are stationary: $B_{i,t} = B_i$ and $\varepsilon_{ik,t} = \varepsilon_{ik}$ for all t. Then at the Ramsey optimum, MCPFs are constant over time:

$$MCPF_{i,t} = MCPF_{i,s} \quad \textit{for all } t, s.$$

Proof. From equation (A.20), the first-order condition for the dynamic Ramsey problem is:

$$\beta^t \frac{\partial W_t}{\partial \tau_{i,t}} = \lambda \gamma^t \frac{\partial R_t}{\partial \tau_{i,t}}.$$

When $\gamma = \beta$, this simplifies to:

$$\frac{\partial W_t}{\partial \tau_{i,t}} = \lambda \frac{\partial R_t}{\partial \tau_{i,t}}.$$

The discount factors cancel. Under stationarity, $\partial W_t/\partial \tau_{i,t} = g_i$ and $\partial R_t/\partial \tau_{i,t} = r_i$ are constant across t. Therefore:

$$\text{MCPF}_{i,t} = \frac{g_i}{r_i} = \lambda \quad \text{for all } t.$$

MCPFs are equalized not only across instruments (the static Ramsey condition) but also across time. ◻

This is Barro's tax smoothing result: optimal policy equalizes distortions intertemporally. When $r_{\text{debt}} = \rho$, there is no wedge between fiscal and social accounting, so the government should smooth MCPFs over time just as it smooths them across instruments. Temporary shocks to revenue needs should be financed primarily by debt, with small permanent tax adjustments to service the debt, maintaining equalized MCPFs across time. Permanent shocks should be financed by permanent tax increases.

When $r_{\text{debt}} \neq \rho$, the result no longer holds. From equation (A.21):

$$\text{MCPF}_{i,t} = \lambda \left(\frac{1 + r_{\text{debt}}}{1 + \rho} \right)^t.$$

If $r_{\text{debt}} < \rho$ (as in recent decades), the fiscal constraint underweights future costs relative to social welfare, creating an incentive to front-load taxation. MCPFs should rise over time at rate $(1 + r_{\text{debt}})/(1+\rho) < 1$. Conversely, if $r_{\text{debt}} > \rho$, MCPFs should decline over time, favoring deferral. The projection framework generalizes Barro's result to environments with discount wedges, showing that optimal policy adjusts for the gap between market rates and social preferences.

H.3 The Atkinson–Stiglitz Theorem

Atkinson and Stiglitz (1976) show that under weak separability between consumption and labor, and with access to an unrestricted non-linear income tax, differential commodity taxation has no role in the second-best. In the projection framework, this theorem implies a particular structure for the welfare and revenue gradients in the commodity-tax subspace.

Let the tax vector be (τ^y, τ^c), where τ^y is a non-linear income tax and $\tau^c \in \mathbb{R}^m$ are commodity taxes. Partition the welfare and revenue gradients as

$$g = \begin{pmatrix} g_y \\ g_c \end{pmatrix}, \qquad r = \begin{pmatrix} r_y \\ r_c \end{pmatrix}.$$

Corollary 3 (Atkinson–Stiglitz as a projection result)**.** *Under the Atkinson–Stiglitz assumptions (weak separability and identical subutility across individuals), evaluated at the second-best optimum* $(\bar{\tau}^y, \bar{\tau}^c)$ *with an optimally chosen non-linear income tax, there exists a scalar* λ *such that*

$$g_c = \lambda r_c.$$

Hence any reform that (i) changes only commodity taxes and (ii) is revenue-neutral must satisfy $r_c^\top d\tau^c = 0$*, and therefore yields zero first-order welfare effect:*

$$g_c^\top d\tau^c = \lambda r_c^\top d\tau^c = 0.$$

Within the commodity-tax subspace there is no component of the welfare gradient orthogonal to the revenue gradient, so revenue-neutral commodity-tax tilts cannot improve welfare.

Proof. Atkinson and Stiglitz show that with an unrestricted non-linear income tax and weak separability, any allocation achievable with differential commodity taxes can be replicated, for the same revenue, using the non-linear income tax alone and a uniform commodity tax. Thus differential commodity taxation has no screening value at the second-best optimum. Locally, this means that small commodity-tax perturbations affect welfare only through their effect on revenue. Formally, there exists λ such that for all $d\tau^c$,

$$dW = g_c^\top d\tau^c = \lambda\, r_c^\top d\tau^c.$$

If a commodity-only reform is revenue-neutral, then $r_c^\top d\tau^c = 0$, and the expression above implies $g_c^\top d\tau^c = 0$ as well. Such reforms therefore have zero first-order welfare effect. This is exactly the Atkinson–Stiglitz conclusion in the projection framework. ◻

The key geometric point is that g_c is collinear with r_c: the welfare gradient in the commodity-tax subspace points fully in the revenue direction. The revenue-neutral hyperplane cuts this subspace along directions orthogonal to r_c, and welfare is flat along those directions.

H.4 Zero Capital Taxation as a Limiting Case

Chamley (1986) and Chari, Nicolini, and Teles (2020) show that under standard macroeconomic assumptions (infinitely lived agents, perfect capital markets, homothetic preferences, additively separable preferences over time), the optimal long-run tax on capital is zero. In those environments, a permanent change in the capital tax rate induces such a large long-run response of the capital stock that the present-value revenue raised by the tax tends to zero. The projection framework makes this logic transparent by expressing it in terms of the present-value revenue gradient.

In the notation of Section 3.1, consider a permanent increase in the capital tax $\tau_{K,0}$. Its present-value revenue effect is

$$r_{K,0}^{PV} = \sum_{t=0}^{\infty} \gamma^t \frac{\partial R_t}{\partial \tau_{K,0}},$$

and the present-value marginal cost of public funds for capital taxation is

$$\mathrm{MCPF}_K^{PV} = \frac{g_{K,0}^{PV}}{r_{K,0}^{PV}}.$$

In a Chamley-type environment, long-run capital supply is effectively infinitely elastic: a permanent increase in τ_K causes the capital tax base to shrink so much over time that

$$r_{K,0}^{PV} \to 0.$$

Given a finite welfare gradient $g_{K,0}^{PV}$, this implies $\mathrm{MCPF}_K^{PV} \to \infty$. In the projection formula, instruments with very high MCPFs are pushed toward their lower bounds. When the present-value revenue gradient for capital taxation vanishes, the optimal reform direction drives τ_K toward zero. The familiar zero-capital-tax result thus appears as a limiting case of the general projection framework, corresponding to an extreme configuration of the present-value revenue gradient.

This limiting case is not generic. If the present-value elasticity of capital supply is large but finite—because of adjustment costs, firm-specific capital, borrowing constraints, or heterogeneous discount rates—then $r_{K,0}^{PV}$ remains strictly positive and so does MCPF_K^{PV}. Straub and Werning (2020) show that relaxing the standard assumptions can overturn the zero-capital-tax result. In the projection framework, this fragility appears as sensitivity of the optimal capital tax to the shape of r_K^{PV}: whenever $r_{K,0}^{PV}$ does not collapse to zero, the optimal reform sets a finite capital tax determined by the same MCPF equalization logic as for any other instrument.