Minimum-Distortion Wealth Taxation, I: Information-Theoretic versus Transport-Geometric Optimality on the Proportional Class

Anders G. Frøseth*

July 8, 2026

Abstract

We characterise minimum-distortion wealth taxation under two contrasting normative criteria within a Fokker–Planck framework on log-wealth: the JKO free-energy gap, an information-theoretic measure aligned with the Mirrleesian decision-distortion tradition, and the squared 2-Wasserstein distance from the no-tax distribution at horizon T, a transport-geometric measure aligned with the Saez–Zucman distributional-compression tradition. Restricting to the neutrality-preserving (C1)–(C3) schedule class of Frøseth (2026c), both optima admit closed forms in the two-dimensional design plane parametrised by the corporate–dividend retention $k = (1-\tau_c)(1-\tau_d)$ and the proportional wealth-tax rate τ_w. The JKO optimum partitions the regime axis into three phases as a function of the dimensionless ratio $\rho = \Sigma_0\, m_0/\sigma^2$, with $m_0 = \mu - \sigma^2/2$ the geometric mean log-return: a pure wealth-tax phase at low ρ, a mixed-instrument phase at intermediate ρ, and a pure flow-tax phase at high ρ. The W_2 optimum, by contrast, is degenerate in this calibration: it pins to the pure flow-tax corner across the whole regime axis. The criterion contrast admits an economically meaningful reading via a *bluntness index* $B(m_0) = b/(am_0)$ that measures the wealth-tax channel's mean-displacement-per-revenue overshoot relative to the flow-tax channel; JKO weights B linearly, W_2 weights it quadratically, and the two normative traditions correspond to this difference in weighting. Norwegian-flavoured calibrations sit inside the JKO mixed-instrument phase under stock-heavy portfolio volatility but move into the pure flow-tax phase under the realised effective volatility of typical real-estate-heavy households. We do not in this paper compare directly to the bracket-based wealth-tax schedules in actual use; that comparison requires a piecewise-Gaussian extension of the present analysis, sketched as a companion paper.

1 Introduction

1.1 The question

The political debate around wealth taxation has been arrayed along a familiar normative axis for decades, with four arguments structuring the standard exchange. Proponents typically advance two claims. *First*, that capital owners accumulate disproportionate influence and political power

*Independent Researcher. E-mail: indrefjorden@pm.me.

relative to wage earners, and that taxation of the capital *stock* — not merely the income flow it generates — is the appropriate instrument through which the wealthier should contribute a larger share to public revenue (Piketty, 2014; Piketty and Zucman, 2014). *Second*, that a wealth tax is the most direct and effective instrument for distributional redistribution: where income taxes track flows that capital owners can defer or restructure, a wealth tax acts on the stock itself, and is therefore the central redistributive tool in any serious progressive programme (Saez and Zucman, 2019; Blanchet et al., 2022). Opponents typically advance two countervailing claims. *Third*, that a wealth tax punishes entrepreneurship and discourages the deployment of risky capital, because it taxes the position rather than the realised return and falls heaviest on risk-bearing assets at exactly the points in the wealth lifecycle where allocation decisions are most consequential (Mirrlees et al., 2011; Saez and Stantcheva, 2018). *Fourth*, that a wealth tax distorts capital allocation and disadvantages domestic investors relative to foreign investors who do not face the same levy in their home jurisdictions, with empirical responses documented in Jakobsen et al. (2020) and the Norwegian Bø reform debate.

These four arguments share a common structure once translated into the language of distortion-minimisation. The two pro-tax arguments say, in effect, that a wealth tax should be designed to minimise the *distributional* perturbation from the no-tax counterfactual — to compress the upper tail, or to move households toward greater equality at finite horizon. The two anti-tax arguments say, in effect, that a wealth tax should be designed to minimise the *allocational* perturbation — not to redirect the household's portfolio decisions, and not to favour foreign over domestic capital. Both normative positions are defensible. Neither is entailed by the other, and as we will show, they correspond to two mathematically distinct distortion criteria. The distributional-compression position aligns with the 2-Wasserstein distance between post-tax and no-tax distributions (W_2^2) — the natural geometric metric on probability measures — and with the Saez–Zucman tradition in optimal-tax theory. The allocational-neutrality position aligns with the JKO free-energy gap — the natural information-theoretic distortion in the Fokker–Planck framework — and with the Mirrleesian decision-distortion tradition. Section 3.3 develops the criterion choice in detail; Section 6.2 develops the connection to the two normative traditions. The political debate is then a debate between two implicit objective functions, both well-defined, and the choice between them is a normative one that the present paper does not adjudicate.

What the paper does is make those two objectives precise, characterise the wealth-tax schedule that minimises each, and prove that the two optima are governed by a single dimensionless ratio that distinguishes regimes in which they coincide from regimes in which they diverge.

1.2 Two distortion criteria

Working within the Fokker–Planck framework of Frøseth (2026g,c) on log-wealth, we restrict attention to the *(C1)–(C3) class* of Frøseth (2026c) — tax schedules satisfying three conditions on the underlying ownership-tax system: the capital-income tax rate matches the corporate tax rate ($\tau_k = \tau_c$); the shielding rate matches the risk-free rate ($r_s = r_f$); and the wealth-tax assessment is uniform across asset classes ($\alpha_i = \alpha$). Table 2 of Section 2.3 states each condition formally with its effect. Together the three conditions imply that the post-tax wealth

process inherits the multiplicative structure of the pre-tax one, with the effect on log-wealth at finite horizon being a *drift-shift-and-rescale*: a uniform shift from the wealth tax and a uniform rescaling of excess drifts by the factor $k = (1-\tau_c)(1-\tau_d)$ from the corporate–dividend flow-tax pair. The design space is parametrised by this pass-through factor $k \in (0,1]$ together with the proportional wealth-tax rate $\tau_w \in [0, \tau_w^{\max}]$. By Frøseth (2026e,b,c) this is exactly the schedule class that preserves *neutrality* with respect to portfolio choice under homogeneous returns and CRRA preferences: the household's risk-bearing decisions are not redirected by a tax in this class. Schedules outside (C1)–(C3) break neutrality.

The design parameters are direct policy levers. The substantive importance of this decomposition cannot be overstated: the design parameters (k, τ_w) are not abstract optimisation variables, but the actual policy levers in the ownership-tax debate. The wealth-tax rate τ_w is the proportional rate the public discussion centres on. The pass-through factor $k = (1-\tau_c)(1-\tau_d)$ is the share of pre-tax returns that survives the corporate–dividend tax sequence: $k = 1$ corresponds to no flow taxation, $k = 0.5$ corresponds to a combined corporate–dividend burden of about half pre-tax returns. Choosing a point in $\mathcal{S}_C = (0,1] \times [0, \tau_w^{\max}]$ *is* choosing an actual ownership-tax package satisfying (C1)–(C3). This is the link between the framework's analytical content and the political-economy debate of Section 1.1: arguments about the wealth tax and arguments about corporate–dividend taxation are arguments about the two axes of $\mathcal{S}_C$.

Within this class, we characterise the schedule that minimises each of two distortion criteria.

The *Jordan–Kinderlehrer–Otto (JKO) free-energy gap* (Jordan et al., 1998)

$$\Delta F[\tau] \;:=\; F[p_T^\tau] - F[p_T^0], \qquad F[p] \;=\; \int p \log p \,\mathrm{d}y + \int V_Y^0 \, p \,\mathrm{d}y, \tag{1}$$

is the information-theoretic distortion measure native to the Fokker–Planck framework. It measures how far the post-tax population distribution at horizon T has been pushed from the no-tax counterfactual in entropy-plus-potential terms. The JKO criterion is the natural formalisation of *minimum decision-distortion*: the policy that minimises ΔF at matched revenue is the policy that least disturbs the natural Fokker–Planck flow of household wealth, and therefore the one that least redirects the household's allocation and risk-bearing decisions. The two anti-tax arguments above are implicit JKO concerns.

The *squared 2-Wasserstein distance* to the no-tax distribution

$$W_2^2(p_T^\tau, p_T^0) \;=\; \int_0^1 \left[F_T^{\tau,-1}(u) - F_T^{0,-1}(u)\right]^2 \mathrm{d}u, \tag{2}$$

introduced in this context by Frøseth (2026h), is the transport-geometric distortion measure: it measures how much "mass" must be moved from the no-tax distribution to reach the post-tax distribution, percentile by percentile. The W_2 criterion is the natural formalisation of *minimum distributional perturbation*: the policy that minimises W_2^2 at matched revenue is the policy that keeps every percentile of the wealth distribution closest to its no-tax counterpart. The two pro-tax arguments above are implicit W_2 concerns when restricted to revenue-matched schedules.

These two criteria are not the same. JKO weights perturbations to the distribution linearly in both the mean shift and the (relative) variance shift, while W_2 weights both quadratically. The quantitative consequence is that the criteria pick different optima within the same (C1)–(C3) design space.

1.3 Results

We establish four results. *First* (Theorems 1 and 2), the constrained minimisation of each criterion within the (C1)–(C3) class admits a unique closed-form optimum; the JKO optimum has the structure "linear in τ_w, log-derivative in k," and the W_2 optimum is quadratic in both. *Second* (Theorem 4), the JKO optimum partitions the regime axis into three phases as a function of the dimensionless ratio

$$\rho \;:=\; \Sigma_0 \cdot m_0/\sigma^2 \;=\; \sqrt{v_0 + \sigma^2 T}\,\frac{\mu - \sigma^2/2}{\sigma^2} \tag{3}$$

combining the horizon variance Σ_0, the log-wealth drift m_0, and the volatility squared: a pure wealth-tax phase at low ρ, a mixed-instrument phase at intermediate ρ, and a pure flow-tax phase at high ρ. The two optima coincide as $\rho \to \infty$ and diverge as $\rho \to 0$. *Third*, the phase structure has a phase-transition reading (Section 5.3): the wealth-tax revenue share $\phi^\star = a\,\tau_w^\star/R^\star$ acts as an order parameter, ρ as a control parameter, and the phase boundaries $\rho_{\text{low}}, \rho_{\text{high}}$ as kinks in $\phi^\star(\rho)$. The corresponding W_2 analysis on (C1)–(C3) yields a degenerate diagram with no transition — a "phase-richness contrast" between the two criteria. *Fourth* (calibration in Section 7), the Norwegian-flavoured regime sits at $\rho \approx 0.23$, inside the mixed-instrument phase under stock-heavy portfolio volatility, but moves toward the corner phases when effective volatility is computed from typical real-estate-heavy household portfolios.

The paper does not adjudicate between the two criteria, and it does not in this paper attempt a direct comparison between the JKO recommendation and the actual schedules in force in Norway, Switzerland, Spain, or France — those are bracket-based and lie outside the proportional (C1)–(C3) class on which the closed forms here are derived. The matching empirical exercise lives on the bracket sub-class, the subject of a companion paper sketched in Section 8.6. What this paper does is articulate the normative split between two distortion criteria as a precise mathematical split, characterise the optimal proportional schedule under each, and identify when the two sides converge and when they diverge across the regime axis.

1.4 Placement in the corpus

The present paper is the FP-side companion to Frøseth (2026h) ("the Wasserstein paper"). Frøseth (2026h) establishes the W_2 optimality of the progressive schedule $\tau^\star(x) = \lambda x^2$ over the full admissible class $\mathcal{S}$, including schedules outside (C1)–(C3) that break neutrality. The present paper restricts to (C1)–(C3) and asks how the JKO and W_2 criteria compare *within* the neutrality-preserving subclass. Together the two papers identify a 2×2 grid of design problems: criterion JKO or W_2, schedule class (C1)–(C3) or full $\mathcal{S}$. The four cells fill in characterisations of minimum-distortion taxation under different combinations of normative criterion and admissible-schedule restrictions.

Part I of a two-paper series. The present paper treats the proportional sub-class — the (C1)–(C3) instruments parametrised by $(k, \tau_w) \in \mathcal{S}_C$ — and develops the phase structure of the JKO optimum in closed form. It mirrors Frøseth (2026e) in establishing the theory cleanly on a restricted class, with the same restriction-to-extension structure that Frøseth (2026b) bears to Frøseth (2026e). A companion paper, in preparation, extends the analysis to the *2-bracket sub-class* (k, X_0, r), the simplest finite-dimensional schedule class that captures the structure actually used by Norway, Switzerland, Spain, and France's old ISF (an exemption threshold combined with marginal rates above). The extension is technically tractable via piecewise-Gaussian FP solutions matched at the bracket boundaries. Section 8.6 sketches the recipe and the relationship to the present paper. Direct comparisons between the recommendations of the present analysis and the actual schedules in force in any specific jurisdiction are deferred to that companion paper, since the proportional class on which the present results are derived is the wrong functional space in which to make such comparisons.

The paper also draws on Frøseth (2026f) for the drift-design language used in the discussion of (C3) violations, and on a forthcoming empirical pipeline for the calibration in Section 7. Heterogeneous-returns extensions (in the spirit of Frøseth (2026d)) and migration-channel considerations (Frøseth (2026a)) are noted as orthogonal extensions in the discussion.

2 Setup

The symbols used throughout the paper are summarised in Table 1; the rest of Section 2 introduces them in context.

2.1 Pre-tax dynamics

We work throughout in log-wealth coordinates $Y = \log X$. Pre-tax wealth X_t follows GBM with drift μ and volatility σ:

$$\mathrm{d}X_t \;=\; \mu X_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t, \qquad X_0 \sim \pi_0. \tag{4}$$

By Itô's lemma, Y satisfies

$$\mathrm{d}Y_t \;=\; m_0\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad m_0 := \mu - \sigma^2/2, \tag{5}$$

with constant diffusion in log-wealth.

2.2 Notation: rate versus schedule

The corpus uses two different conventions for tax-schedule notation, and reconciling them needs to be explicit. Frøseth (2026e,b,c) use τ (or τ_w) for the *wealth-tax rate*: a holder of wealth x pays $\tau_w \cdot x$ per unit time. Frøseth (2026h) (the Wasserstein paper) uses $\tau(x)$ for the *tax schedule* (total amount paid).

We adopt an explicit two-symbol convention. Let $r(x)$ denote the wealth-tax rate per unit wealth per unit time, consistent with the τ_w-as-rate convention of Frøseth (2026e). Let $\tau(x) := r(x) \cdot x$

Wealth process and parameters	
X_t	Pre-tax wealth process; X_0 has distribution π_0.
$Y_t = \log X_t$	Log-wealth process.
μ, σ	GBM drift and volatility on raw wealth.
$m_0 := \mu - \sigma^2/2$	Drift on log-wealth under no taxation; equivalently the *expected log-wealth growth rate* (volatility-corrected geometric mean return) of pre-tax wealth.
T	Finite-horizon design parameter.
$\pi_0 \in \mathcal{P}_2(\mathbb{R}_+)$	Initial wealth distribution.
μ_0, v_0	Mean and variance of $\log X_0$ when X_0 is log-normal.
Population distributions	
p_t^0	No-tax distribution of Y_t at time t (Fokker–Planck flow under no tax).
p_t^τ	Post-tax distribution of Y_t at time t under schedule τ.
M_t^τ, Σ_t^τ	Mean and standard deviation of p_t^τ when Gaussian.
$\Sigma_0 := \sqrt{v_0 + \sigma^2 T}$	Standard deviation of Y_T under no tax.
Tax instrument and schedule classes	
$r(x)$	Wealth-tax rate per unit wealth (rate convention of Frøseth (2026e,c)).
$\tau(x) := r(x) \cdot x$	Tax schedule: total tax paid by a holder of wealth x per unit time (schedule convention of Frøseth (2026h) and the present paper).
τ_w	Proportional wealth-tax rate (the constant rate in the (C1)–(C3) class).
$k = (1-\tau_c)(1-\tau_d)$	Pass-through factor (corporate–dividend flow-tax product), $k \in (0,1]$. Post-tax log-wealth volatility is $\sqrt{k}\,\sigma$.
$\mathcal{S}$	Full admissible schedule class.
$\mathcal{S}_C := (0,1] \times [0, \tau_w^{\max}]$	(C1)–(C3) design space.
$\mathcal{S}_R \subset \mathcal{S}_C$	Revenue-matched constraint set ($\bar{R} = R^\star$).
$\mathcal{N}$	Neutral schedule class of Frøseth (2026e).
Distortion criteria and reference quantities	
$V_Y^0(y) := -m_0 y/\sigma^2$	Reference potential (no-tax FP gradient flow).
$F[p] = \int p \log p + \int V_Y^0 p$	JKO free-energy functional.
$\Delta F[\tau] := F[p_T^\tau] - F[p_T^0]$	Terminal-gap JKO criterion.
$W_2^2(p,q)$	Squared 2-Wasserstein distance between distributions p and q.
$R, R^\star$	Time-averaged revenue rate; revenue target.
a, b	Linear-approximation revenue weights: $R \approx a\,\tau_w + b\,(1-k)$.
Crossover	
$\rho := \Sigma_0\, m_0/\sigma^2$	Dimensionless crossover ratio governing the JKO/W_2 split (Section 5).

Table 1: Summary of notation. Variables introduced once and used locally are not listed; see the surrounding text at first appearance.

denote the tax schedule (total tax paid by a holder of wealth x per unit time), consistent with the variational object of Frøseth (2026h). Each policy can be described cleanly via either object:

Policy	Rate $r(x)$	Schedule $\tau(x)$
Proportional wealth tax	τ_w (constant)	$\tau_w\, x$
W_2-optimal of Frøseth (2026h)	λx	λx^2
Two-bracket (threshold x_0, rate r_0)	$r_0\, \mathbf{1}\{x \geq x_0\}$	$r_0\, (x - x_0)_+$

The drift modification on log-wealth from a stock tax is determined by the rate: $\delta v_Y(y) = -r(e^y)$.

2.3 Tax instruments and the (C1)–(C3) class

A flow + stock tax instrument τ lies in the *(C1)–(C3) class* of Frøseth (2026c) if it satisfies the three conditions summarised in Table 2.

Condition	**Statement**	**Effect**
(C1)	Capital-income tax rate equals corporate tax rate: $\tau_k = \tau_c$.	Removes the differential between taxation of dividends paid out versus capital income retained, so that the after-tax excess return is independent of how a unit of pre-tax return is realised.
(C2)	Shielding rate equals the risk-free rate: $r_s = r_f$.	The shielding mechanism (Norwegian *skjermingsfradrag*) acts as a pure deduction at the risk-free rate, equating after-tax excess returns across assets up to the wealth-tax channel. Together with (C1) this ensures the equity–debt split is undistorted.
(C3)	Uniform wealth-tax assessment across assets: $\alpha_i = \alpha$ for all assets i.	No asset-class-specific drift modification; the wealth tax acts as a uniform drift shift on log-wealth. Violations distort the tangency portfolio.

Table 2: The three conditions of the (C1)–(C3) class, restated from Frøseth (2026c). *Together* they imply that the combined tax system acts on the Fokker–Planck equation as a uniform drift shift (from the wealth tax) plus a uniform rescaling of excess drifts by the factor $k = (1 - \tau_c)(1-\tau_d)$ (from the corporate–dividend flow-tax pair). The household's portfolio-choice problem is unchanged by the tax up to this rescaling; Frøseth (2026c) calls this property *generalised neutrality*.

Specifically, Frøseth (2026c) establishes that under (C1)–(C3) the post-tax log-wealth process takes the drift-shift-and-rescale form

$$\mathrm{d}Y_t \;=\; [k\, m_0 - \tau_w]\, \mathrm{d}t + \sqrt{k}\, \sigma\, \mathrm{d}W_t, \tag{6}$$

where the pass-through factor $k = (1 - \tau_c)(1 - \tau_d) \in (0, 1]$ is determined by the corporate and dividend tax rates (with $k = 1$ corresponding to the no-flow-tax case $\tau_c = \tau_d = 0$, and $k \to 0$ corresponding to full pass-through), and the proportional wealth-tax rate $\tau_w \in [0, \tau_w^{\max}]$ enters

as a uniform drift shift on log-wealth.

The link between policy levers and Fokker–Planck channels. Equation 6 makes precise an idea that motivates everything that follows. The two design parameters (k, τ_w) enter the Fokker–Planck equation through *distinct channels*: the wealth-tax rate τ_w acts purely as a uniform drift shift, while the corporate–dividend pass-through k acts purely as a multiplicative rescaling of both drift and diffusion. The drift-shift mechanism is the substance of Frøseth (2026e)'s neutrality result for the proportional wealth tax in the single-asset GBM setting; the rescaling-by-k mechanism is Frøseth (2026c)'s generalisation when the full ownership-tax system (corporate + dividend + capital-income + wealth) is in play. The (C1)–(C3) class is therefore not a *technical* restriction so much as the *natural decomposition* of how policy operates on the wealth process: every ownership-tax package satisfying (C1)–(C3) reduces to some choice of (k, τ_w), and conversely every choice of (k, τ_w) in $\mathcal{S}_C$ corresponds to an actual implementable ownership-tax package.

This decomposition is what allows the design problem of this paper to be stated cleanly. We are not optimising over abstract function spaces; we are choosing a single point in the two-dimensional (k, τ_w)-plane, with each axis a familiar policy lever and the constraint set $\mathcal{S}_R$ a one-dimensional curve corresponding to matched-revenue trade-offs. The substantive question — how much of public revenue should come from the wealth-tax channel (τ_w) and how much from the corporate–dividend channel (k) — is exactly the question the JKO and W_2 optima of Section 4 answer, each from its own normative criterion.

We denote the resulting design space by

$$\mathcal{S}_C \;:=\; (0, 1] \times [0, \tau_w^{\max}]. \tag{7}$$

The (C1)–(C3) class is at once a substantive and a technical restriction. Substantively, it is the class of *neutrality- preserving* schedules in the sense of Frøseth (2026e) — schedules under which the household's risk-bearing decisions are not redirected by the tax. Technically, it reduces the infinite-dimensional schedule space to the two-parameter design space $\mathcal{S}_C$ above, in which existence and uniqueness of optima follow from standard convex-analysis arguments (Section 4.1). Schedules outside (C1)–(C3) break neutrality and can be analysed in the broader framework of Frøseth (2026h) or Frøseth (2026f); the present paper restricts to (C1)–(C3) by design.

2.4 Population measure and Fokker–Planck dynamics

The population is a Borel probability measure $\pi_0 \in \mathcal{P}_2(\mathbb{R}_+)$ on initial wealth, with finite second moments on log-wealth so that $Y_0 \in \mathcal{P}_2(\mathbb{R})$. Specific results that involve closed-form ratios additionally assume X_0 to be log-normal with log-volatility σ_0; we flag this assumption at the point of use rather than treating it as standing.

The population law of log-wealth at time $t \in [0, T]$ under no taxation is the pushforward $p_t^0 :=$

$\mathrm{Law}(Y_t \mid (5))$, satisfying the Fokker–Planck equation

$$\partial_t p_t^0 \;=\; -m_0\,\partial_y p_t^0 + \tfrac{\sigma^2}{2}\,\partial_y^2 p_t^0. \tag{8}$$

Under a schedule $\tau \in \mathcal{S}_C$ the post-tax log-wealth has modified drift and diffusion as in (6), and its population law p_t^τ satisfies the FP equation with the modified coefficients.

The free-energy functional F of (12) is the JKO gradient flow object for (8): the no-tax FP equation is the steepest-descent flow of F in the 2-Wasserstein metric on $\mathcal{P}_2(\mathbb{R})$ (Jordan et al., 1998; Santambrogio, 2017; Villani, 2009). This is the structural reason why F is the natural information-theoretic distortion measure: any tax modification of the FP coefficients perturbs the flow away from the gradient flow of F, and the gap $\Delta F[\tau]$ measures the magnitude of that perturbation at horizon T.

2.5 Revenue constraint

A schedule τ collects revenue at rate

$$R_t[\tau] \;:=\; \mathbb{E}[\tau(X_t)] \;=\; \mathbb{E}[r(X_t)\cdot X_t] \tag{9}$$

per unit time, where the expectation is under the post-tax law. For matched-revenue analysis we use the time-averaged revenue rate

$$\bar{R}[\tau] \;:=\; \frac{1}{T}\int_0^T R_t[\tau]\,\mathrm{d}t, \tag{10}$$

and the revenue-matched constraint set

$$\mathcal{S}_R \;:=\; \{\tau \in \mathcal{S}_C \;:\; \bar{R}[\tau] = R^\star\}, \tag{11}$$

where $R^\star > 0$ is a target revenue rate. Within $\mathcal{S}_C$ the set $\mathcal{S}_R$ is one-dimensional (a curve in the (k, τ_w)-plane), and the JKO and W_2 minimisation problems are to find points on this curve that minimise the corresponding distortion criterion.

For analytical convenience we sometimes work with a linear approximation $R[\tau] \approx a\,\tau_w + b\,(1-k)$ to the revenue functional, valid in the small-tax regime. The linear approximation is sufficient for the closed-form FOCs of Section 4.2; the full log-quadratic form is recovered for the calibration of Section 7.

3 The two distortion criteria

3.1 The JKO free-energy gap

The Fokker–Planck equation (8) is the gradient flow, in the 2-Wasserstein metric on $\mathcal{P}_2(\mathbb{R})$, of the *Jordan–Kinderlehrer–Otto free-energy functional* (Jordan et al., 1998)

$$F[p] \;=\; \int p(y)\,\log p(y)\,\mathrm{d}y + \int V_Y^0(y)\,p(y)\,\mathrm{d}y, \qquad V_Y^0(y) \;:=\; -\frac{m_0}{\sigma^2}\,y. \tag{12}$$

The reference potential V_Y^0 is chosen so that the no-tax FP flow (8) is the steepest-descent flow of F: $\partial_t p_t^0 = -\nabla_{W_2} F[p_t^0]$. The tax modification of the FP coefficients perturbs the flow away from this gradient flow, and the JKO terminal gap

$$\Delta F[\tau] \;:=\; F[p_T^\tau] - F[p_T^0] \tag{13}$$

measures the magnitude of that perturbation at horizon T. We use ΔF throughout as the JKO criterion. For the convention discussion of fixed-reference-potential versus moving-reference, which distinguishes ΔF from Kullback–Leibler divergence, see Section 3.3.

Closed form for Gaussian distributions. Under (C1)–(C3) and a log-normal initial wealth X_0 with $\log X_0 \sim \mathcal{N}(\mu_0, v_0)$, the post-tax log-wealth at horizon T is Gaussian: $Y_T^\tau \sim \mathcal{N}(M_T^\tau, {\Sigma_T^\tau}^2)$ with

$$M_T^\tau \;=\; \mu_0 + (k\, m_0 - \tau_w)\, T, \qquad {\Sigma_T^\tau}^2 \;=\; v_0 + k\,\sigma^2\, T. \tag{14}$$

The no-tax distribution is the same with $(k, \tau_w) = (1, 0)$: $M_T^0 = \mu_0 + m_0 T$, ${\Sigma_T^0}^2 = v_0 + \sigma^2 T$. Direct computation using $H(\mathcal{N}(M, \Sigma^2)) = \frac{1}{2}\log(2\pi e \Sigma^2)$ and the linear potential V_Y^0 gives

$$\boxed{\Delta F[(k, \tau_w)] \;=\; -\tfrac{1}{2}\log\frac{v_0 + k\sigma^2 T}{v_0 + \sigma^2 T} \;-\; \frac{m_0^2}{\sigma^2}\,(k-1)\,T \;+\; \frac{m_0}{\sigma^2}\,\tau_w\, T.} \tag{15}$$

Sanity check and properties. $\Delta F[(1, 0)] = 0$, as required: the no-tax point makes the gap vanish. The first term is strictly convex in k on $(0, 1]$ (its second derivative is $\sigma^4 T^2/[2(v_0 + k\sigma^2 T)^2] > 0$); the second term is linear in k; the third is linear in τ_w. So ΔF is strictly convex in k and jointly convex in (k, τ_w), which underwrites the existence-and-uniqueness arguments of Section 4.1.

Linear-in-perturbation scaling. Around the no-tax point, write $k = 1 - \epsilon_k$ and $\tau_w = \epsilon_w$ with both small. The leading-order expansion of (15) is

$$\Delta F \;\approx\; \frac{\sigma^2 T}{2(v_0 + \sigma^2 T)}\,\epsilon_k \;+\; \frac{m_0^2 T}{\sigma^2}\,\epsilon_k \;+\; \frac{m_0 T}{\sigma^2}\,\epsilon_w \;=\; \alpha_k\,\epsilon_k + \alpha_w\,\epsilon_w \tag{16}$$

for coefficients $\alpha_k, \alpha_w > 0$. The JKO gap is *linear* in both perturbations at leading order: a small variance reduction ϵ_k costs $\alpha_k\,\epsilon_k$, a small mean shift ϵ_w costs $\alpha_w\,\epsilon_w$, and their relative weight is set by the dimensionless ratio $\rho := \Sigma_0\, m_0/\sigma^2$ that drives the crossover of Section 5.

3.2 The 2-Wasserstein distance to the no-tax distribution

The squared 2-Wasserstein distance between two probability measures $p, q \in \mathcal{P}_2(\mathbb{R})$ is

$$W_2^2(p, q) \;=\; \inf_{\pi \in \Pi(p,q)} \int (y_1 - y_2)^2\, \pi(\mathrm{d}y_1, \mathrm{d}y_2), \tag{17}$$

where $\Pi(p, q)$ is the set of joint laws with marginals p, q (Villani, 2009). In one dimension, the infimum is attained by the comonotonic coupling — the unique coupling under which both

marginals are mapped via their inverse CDFs to the same uniform random variable — giving the explicit formula

$$W_2^2(p,q) \;=\; \int_0^1 \left[F_p^{-1}(u) - F_q^{-1}(u)\right]^2 \mathrm{d}u, \tag{18}$$

where F_p, F_q are the cumulative distribution functions. The W_2 distance to the no-tax distribution at horizon T,

$$W_2^2[\tau] \;:=\; W_2^2(p_T^\tau, p_T^0), \tag{19}$$

is the transport-geometric distortion measure introduced in this context by Frøseth (2026h).

Closed form for Gaussian distributions. For Gaussian measures the comonotonic coupling reduces to a linear map and (18) evaluates to

$$W_2^2(\mathcal{N}(M_p,\Sigma_p^2), \mathcal{N}(M_q,\Sigma_q^2)) \;=\; (M_p - M_q)^2 + (\Sigma_p - \Sigma_q)^2. \tag{20}$$

Applying this to (14),

$$\boxed{W_2^2[(k,\tau_w)] \;=\; \left[(k-1)\,m_0\,T - \tau_w\,T\right]^2 \;+\; \left[\sqrt{v_0 + k\sigma^2 T} - \sqrt{v_0+\sigma^2 T}\right]^2.} \tag{21}$$

Quadratic-in-perturbation scaling. Around the no-tax point ($k = 1-\epsilon_k$, $\tau_w = \epsilon_w$), the leading-order expansion of (21) is

$$W_2^2 \;\approx\; \frac{(\sigma^2 T)^2}{4(v_0+\sigma^2 T)}\,\epsilon_k^2 \;+\; (m_0\,T\,\epsilon_k + T\,\epsilon_w)^2 \;=\; \beta_k\,\epsilon_k^2 + 2\,\beta_{kw}\,\epsilon_k\epsilon_w + \beta_w\,\epsilon_w^2. \tag{22}$$

The W_2 criterion is *quadratic* in both perturbations at leading order. This is the structural difference from the JKO criterion that drives the crossover: the same trade-off in the (ϵ_k, ϵ_w)-plane is weighted differently by linear and quadratic objective surfaces.

Convexity. The closed form (21) is the sum of a quadratic term in the means and a squared difference of standard deviations. The first is convex in (k, τ_w) jointly. The second is convex in k on $[0,1]$ as the squared difference of two concave functions of k that meet at $k=1$. So W_2^2 is jointly convex in (k,τ_w), with the same well-posedness consequences for existence and uniqueness as for ΔF (see Section 4.1).

3.3 Why these two

Several other distortion criteria are natural in the Fokker–Planck setting. We comment briefly on each and explain why the JKO/W_2 pair is the canonical choice for the present analysis.

Kullback–Leibler divergence. The relative entropy $\mathrm{KL}(p_T^\tau \,\|\, p_T^0) = \int p_T^\tau \log(p_T^\tau/p_T^0)\,\mathrm{d}y$ is a tempting candidate: it has a clean information-theoretic interpretation, it is nonnegative, it vanishes if and only if $p_T^\tau = p_T^0$. The relationship to ΔF is subtle. Both have the form "entropy of the post-tax distribution plus a linear functional of it," but the linear functional differs: ΔF uses a *fixed* potential $V_Y^0(y) = -m_0 y/\sigma^2$ (the natural gradient-flow potential of the no-tax FP

equation), while KL uses the *moving* potential $-\log p_T^0(y) = (y - M_0)^2/(2\Sigma_0^2) + \text{const}$ (the negative log of the no-tax density). The two coincide on the no-tax point and at the linear level when V_Y^0 matches $-\partial_y \log p_T^0$; they diverge for non-trivial perturbations. We work with ΔF rather than KL because the gradient-flow structure of (8) privileges the fixed potential: ΔF measures perturbation *against the natural FP flow*, whereas KL measures it *against the no-tax distribution*, which is itself moving. The two criteria pick different optima within $\mathcal{S}_R$ in general, although they agree on the qualitative neutrality-vs-distributional split discussed in Section 6.

Time-integrated free energy. A natural alternative to the terminal gap (13) is the time-integrated version $\Delta F^{\text{int}}[\tau] := \int_0^T (F[p_t^\tau] - F[p_t^0])\,\mathrm{d}t$, which weights perturbations across the entire horizon rather than only at the end. One might worry that the terminal-only criterion underweights early-horizon distortion. Numerical evidence (Section 8.1) shows the two variants pick the *same* optimum on the policy-relevant $\mathcal{S}_R$ contour: the integrand $F[p_t^\tau] - F[p_t^0]$ grows roughly linearly in t for fixed schedule, so $\Delta F^{\text{int}} \approx (T/2)\,\Delta F$ and the ranking of schedules is preserved. The terminal-gap variant is therefore the natural canonical form, and we work with it throughout.

Other transport-geometric distances. Higher-order Wasserstein distances W_2^p for $p > 2$ and divergence-form transport costs (Sinkhorn, entropy-regularised optimal transport) all yield criteria related to but distinct from W_2^2. We work with $p = 2$ because it is the natural metric of the JKO scheme on $\mathcal{P}_2(\mathbb{R})$ and because the comonotonic reduction (18) gives a closed-form one-dimensional expression. Other p values are tractable but produce the same qualitative split as W_2^2 at the leading order; the choice $p = 2$ is conventional in this context.

Variance and Gini-style criteria. Direct distributional measures — the variance of post-tax log-wealth, the Gini coefficient, top-share metrics — are common in the policy literature but are not naturally embedded in the FP framework. They can be related to ΔF and W_2^2 via moment arguments under Gaussian distributions (e.g., the Gini coefficient of $\mathcal{N}(M, \Sigma^2)$ is $2\Phi(\Sigma/\sqrt{2}) - 1$, monotone in Σ), but the relation is non-canonical and the FP-native pair is cleaner for the variational analysis. Discussion of how the JKO and W_2 optima interact with Gini-style metrics is deferred to Section 8.3 and to the redistribution-design companion paper Frøseth (2026f).

The chosen pair. ΔF and W_2^2 together span the natural distinction between information-theoretic and transport-geometric distortion in the Fokker–Planck framework. Both are well-defined on $\mathcal{P}_2(\mathbb{R})$, both are convex on $\mathcal{S}_C$, and both reduce to clean closed-form objectives under (C1)–(C3) with Gaussian distributions at horizon T. Their structural difference — linear versus quadratic scaling in perturbations of (k, τ_w) — is exactly what produces the regime-dependent crossover at the heart of this paper.

Mathematical canonicity of the pair. A reader sceptical that the JKO/W_2 contrast might be an artefact of two arbitrarily-chosen criteria should be reminded that the two are not independent picks. The Jordan et al. (1998) variational scheme establishes the Fokker–Planck

equation as the gradient flow of relative entropy on the Wasserstein-2 manifold $(\mathcal{P}_2(\mathbb{R}), W_2)$: the FP solution at time $t = n\tau$ minimises $\frac{1}{2\tau} W_2^2(p, p_{(n-1)\tau}) + \mathrm{Ent}(p)$ over $p \in \mathcal{P}_2$ for each step. The two functionals ΔF and W_2^2 are the two pieces of *the same scheme*: ΔF is the time-derivative of relative entropy along that gradient flow, and W_2^2 is the metric on the manifold. Selecting them together as the criterion pair is not picking two of many; it is reading the JKO scheme as a normative optimal-tax instrument. Other distortion criteria (KL against the moving distribution, W_2^p for $p \neq 2$, Sinkhorn divergences, Gini-style metrics) lack this canonical relationship to the FP equation. The mathematical-canonicity argument stands even for a reader who rejects the normative-tradition framing of Section 6.

Falsifiability of the contrast. The phase-transition crossover and the bluntness mechanism that produces it (Section 5.5) make falsifiable claims rather than tautological ones. Specifically: the JKO and W_2 optima *coincide* at high ρ (both pin to the lower corner), constraining the contrast at one end of the regime axis. The phase boundaries $\rho_{\text{low}}, \rho_{\text{high}}$ have specific locations derivable from the GBM primitives and the revenue weights $(\mu, \sigma, T, v_0, R^\star, a, b)$, not free parameters fitted to a desired result. The mechanism is named and algebraic: the bluntness index $B(m_0) = b/(am_0)$ enters the JKO penalty linearly and the W_2 penalty quadratically, and a reader who doubts the contrast can verify the gradient calculation directly. Both criteria pick optima within the same (k, τ_w) plane on the same matched-revenue contour — a considerable structural constraint — and disagree only on *where* on the contour the optimum sits. Each of these properties could in principle fail; that they hold is content, not assumption.

4 Variational problem within (C1)–(C3)

We now turn to the variational analysis: find the schedule $\tau \in \mathcal{S}_R$ that minimises each criterion, characterise the optimum in closed form, and establish the technical conditions that guarantee well-posedness. Throughout this section we use the linear revenue approximation

$$R[\tau] \;\approx\; a\,\tau_w + b\,(1-k), \qquad a, b > 0, \tag{23}$$

valid in the small-tax regime.[1] The constants a and b are interpretable as the revenue elasticities of the wealth-tax base and the flow-tax base respectively at the no-tax point; they are jurisdiction-specific and known empirically.

4.1 Existence and uniqueness

Theorem 1 (Existence and uniqueness)**.** *The constrained minimisation problems* $\arg\min_{\tau \in \mathcal{S}_R} \Delta F[\tau]$ *(JKO) and* $\arg\min_{\tau \in \mathcal{S}_R} W_2^2(p_T^\tau, p_T^0)$ *(W_2) each admit a unique minimiser.*

Proof. The constraint set $\mathcal{S}_R \subset \mathcal{S}_C$ is a closed line segment: with R linear and $\mathcal{S}_C = (0,1] \times [0, \tau_w^{\max}]$ a closed bounded box, $\mathcal{S}_R = \{(k, \tau_w) \in \mathcal{S}_C : a\tau_w + b(1-k) = R^\star\}$ is the intersection of

[1] The linear approximation is sufficient for the closed-form FOCs below; the full log-quadratic revenue functional $R[\tau] = \tau_w\,\mathbb{E}[X_T] + (1-k)\,\mathbb{E}[\text{flow base}]$ is recovered for the calibration of Section 7, where the moments $\mathbb{E}[X_T] = \exp(M_T + \Sigma_T^2/2)$ are taken under the post-tax law.

a hyperplane with a compact convex set, hence itself compact and convex.

Both objective functions are continuous on $\mathcal{S}_C$: $\Delta F[(k, \tau_w)]$ has explicit form (15), and $W_2^2[(k, \tau_w)]$ has explicit form (21), both manifestly continuous in (k, τ_w). By Weierstrass's theorem, each attains its minimum on the compact set $\mathcal{S}_R$.

For uniqueness it suffices to show strict convexity of each objective on $\mathcal{S}_R$. Both objectives are jointly convex on $\mathcal{S}_C$ (Sections 3.1 and 3.2). ΔF is strictly convex in k and linear in τ_w (the second derivative $\partial_k^2 \Delta F = \sigma^4 T^2/[2(v_0 + k\sigma^2 T)^2] > 0$); W_2^2 is jointly strictly convex in (k, τ_w) except along the no-tax direction, where the leading-order quadratic in (22) has positive definite Hessian away from the origin. Strict convexity in any direction not parallel to the constraint hyperplane is enough: along the one-dimensional constraint set $\mathcal{S}_R$, both objectives are strictly convex functions of the single free parameter, and admit unique minima. □

Remark. Theorem 1 confirms that the optima we characterise below are well-defined, theorem-grade objects — not non-uniqueness artefacts. This matters for the comparison theorem of Section 5: both criteria pick *a single point* in $\mathcal{S}_C$, and the crossover concerns the displacement between those two points as a function of GBM parameters.

4.2 The Lagrangian and first-order conditions

The constrained minimisation problem $\min_{\tau \in \mathcal{S}_R} \mathcal{J}[\tau]$ (where $\mathcal{J}$ stands for either ΔF or W_2^2) has Lagrangian

$$\mathcal{L}(k, \tau_w; \mu) \;=\; \mathcal{J}(k, \tau_w) \;-\; \mu \left[a\,\tau_w + b\,(1 - k) - R^\star \right], \tag{24}$$

with multiplier $\mu \in \mathbb{R}$. The first-order conditions are

$$\partial_k \mathcal{J} \;+\; \mu\, b \;=\; 0, \tag{25}$$

$$\partial_{\tau_w} \mathcal{J} \;-\; \mu\, a \;=\; 0, \tag{26}$$

$$a\,\tau_w + b\,(1 - k) \;=\; R^\star. \tag{27}$$

For the JKO criterion, differentiating (15):

$$\partial_k \Delta F \;=\; -\frac{\sigma^2 T}{2(v_0 + k\sigma^2 T)} - \frac{m_0^2}{\sigma^2}\, T, \qquad \partial_{\tau_w} \Delta F \;=\; \frac{m_0\, T}{\sigma^2}. \tag{28}$$

The τ_w-derivative is *constant in* (k, τ_w), which is the "linear in τ_w" structure noted in Section 3.1: the JKO cost of an extra unit of τ_w does not depend on the operating point. The k-derivative has a hyperbolic-in-k contribution from the entropy term plus a constant from the potential term.

For the W_2 criterion, differentiating (21):

$$\partial_k W_2^2 \;=\; 2\, m_0 T\, \varepsilon_M + \frac{\sigma^2 T}{\Sigma_T^\tau}\, \varepsilon_\Sigma, \qquad \partial_{\tau_w} W_2^2 \;=\; -2T\, \varepsilon_M, \tag{29}$$

where $\varepsilon_M := M_T^\tau - M_T^0 = (k-1)m_0 T - \tau_w T$ and $\varepsilon_\Sigma := \Sigma_T^\tau - \Sigma_T^0$ are the mean and standard-

deviation shifts. Both derivatives are *linear in* $(\varepsilon_M, \varepsilon_\Sigma)$ — the "quadratic in perturbation" structure of W_2 is exactly that its first-order conditions are linear in the linearised quantities.

Structural comparison. The key contrast between the two FOC systems is now visible. The JKO multiplier μ_{JKO} is determined immediately from (26) and the fact that $\partial_{\tau_w}\Delta F$ is constant:

$$\mu_{\text{JKO}} \;=\; \frac{m_0\,T}{a\,\sigma^2}. \tag{30}$$

The constant value of μ_{JKO} depends only on the GBM parameters and the wealth-tax revenue weight a, not on the operating point. This is what makes the JKO problem cleanly solvable in closed form: the τ_w FOC fixes the multiplier, and the k FOC is then a single equation in a single unknown.

The W_2 multiplier μ_{W_2} depends on ε_M via (26), and ε_M depends on the operating point. The system is therefore coupled in (k, τ_w, μ) and solved simultaneously, but the solution is still in closed form because both FOCs are linear in $(\varepsilon_M, \varepsilon_\Sigma)$ and the linearised constraint is linear in (k, τ_w).

4.3 Closed-form optima

Theorem 2 (Closed-form JKO optimum)**.** *Subject to the* existence condition

$$\frac{b}{a} \;>\; m_0, \tag{31}$$

the unique JKO-optimal point in $\mathcal{S}_C$ *is*

$$k^\star_{\text{JKO}} \;=\; \frac{1}{\sigma^2 T}\left[\frac{\sigma^4}{2\,m_0\,(b/a - m_0)} - v_0\right], \qquad \tau^\star_{w,\text{JKO}} \;=\; \frac{R^\star - b\,(1 - k^\star_{\text{JKO}})}{a}. \tag{32}$$

If $b/a \le m_0$ *the unique optimum is the corner* $k^\star_{\text{JKO}} = 1$, $\tau^\star_{w,\text{JKO}} = R^\star/a$ *— the proportional wealth tax with no flow component.*

Proof. Substitute μ_{JKO} from (30) into the k-FOC:

$$-\frac{\sigma^2 T}{2(v_0 + k\sigma^2 T)} - \frac{m_0^2 T}{\sigma^2} + \frac{m_0 T}{a\sigma^2}\,b \;=\; 0,$$

which rearranges to

$$\frac{\sigma^2}{2(v_0 + k\sigma^2 T)} \;=\; \frac{m_0}{\sigma^2}\left(\frac{b}{a} - m_0\right). \tag{33}$$

The left-hand side is positive, so an interior solution requires the right-hand side to be positive, giving the existence condition (31). Inverting (33) for k gives the expression for $k^\star_{\text{JKO}}$ in (32). Substituting into the revenue constraint (27) gives $\tau^\star_{w,\text{JKO}}$. If the existence condition fails, the right-hand side of (33) is non-positive, no interior k satisfies it, and the minimum lies on the boundary $k = 1$, where the FOC reduces to a problem in τ_w alone and the revenue constraint determines $\tau^\star_w = R^\star/a$. □

Theorem 3 (Closed-form W_2 optimum)**.** *Subject to the existence condition $b/a > m_0$, the unique W_2-optimal point in $\mathcal{S}_C$, at leading order in the linearised perturbations $(\varepsilon_M, \varepsilon_\Sigma)$ about the no-tax point, is*

$$k^\star_{W_2} \;=\; 1 \;-\; \frac{4\,\Sigma_0^2\,(b/a - m_0)\,(R^\star/a)}{4\,\Sigma_0^2\,(b/a - m_0)^2 + \sigma^4}, \qquad \tau^\star_{w,W_2} \;=\; \frac{R^\star - b\,(1 - k^\star_{W_2})}{a}, \tag{34}$$

where $\Sigma_0^2 = v_0 + \sigma^2 T$. If $b/a \le m_0$ the optimum is the upper corner $k^\star_{W_2} = 1$, $\tau^\star_{w,W_2} = R^\star/a$. If the closed-form $k^\star_{W_2}$ in (34) falls below the feasibility bound $k_{\mathrm{lo}} = 1 - R^\star/b$ (the constraint $\tau_w \ge 0$), the constrained optimum on $\mathcal{S}_R$ is the lower corner $(k_{\mathrm{lo}}, 0)$.

Proof. Reparametrise the matched-revenue contour by $u = k - 1 \in [k_{\mathrm{lo}} - 1, 0]$, eliminating τ_w via $\tau_w = (R^\star + bu)/a$. Linearise the post-tax mean and spread shifts about the no-tax point $(k, \tau_w) = (1, 0)$:

$$\varepsilon_M \;=\; -T\,\big[\,u\,(b/a - m_0) + R^\star/a\,\big], \qquad \varepsilon_\Sigma \;=\; \frac{\sigma^2 T}{2\,\Sigma_0}\,u + O(u^2),$$

where the second expression follows from a Taylor expansion of $\Sigma_T = \sqrt{v_0 + (1+u)\sigma^2 T}$ about $u = 0$. Writing $q := b/a - m_0$ and $r := R^\star/a$, the leading-order W_2^2 on the contour is

$$W_2^2(u) \;\approx\; T^2(uq + r)^2 \;+\; \left(\tfrac{\sigma^2 T}{2\Sigma_0}\right)^2 u^2.$$

This is a strictly convex quadratic in the single variable u, with first-order condition $T^2 q(uq + r) + (\sigma^2 T)^2/(4\Sigma_0^2) \cdot u = 0$. Solving for u and dividing through by T^2 gives

$$u \;=\; -\,\frac{q\,r}{q^2 + \sigma^4/(4\Sigma_0^2)} \;=\; -\,\frac{4\,\Sigma_0^2\,q\,r}{4\,\Sigma_0^2\,q^2 + \sigma^4},$$

which is (34) after substituting $u = k - 1$. The corresponding $\tau^\star_{w,W_2}$ follows from the revenue constraint $a\tau_w - bu = R^\star$. The corner cases ($b/a \le m_0$ or $k^\star_{W_2} < k_{\mathrm{lo}}$) are the same projection arguments used in the proof of Theorem 2: the unconstrained quadratic minimum lies outside the feasibility interval $[k_{\mathrm{lo}}, 1]$, and the constrained minimum is attained at the nearest boundary. □

Remark (The two optima coincide at the no-tax boundary). At the limit $R^\star \to 0$ both optima reduce to the no-tax point $(k, \tau_w) = (1, 0)$. Away from this limit, they differ, generically. Section 5 characterises how the displacement between $(k^\star_{\mathrm{JKO}}, \tau^\star_{w,\mathrm{JKO}})$ and $(k^\star_{W_2}, \tau^\star_{w,W_2})$ depends on the GBM parameters.

Remark (The existence condition). The condition $b/a > m_0$ admits an economic reading. The ratio b/a is the relative revenue weight of a unit of flow-tax pass-through (the $1 - k$ direction) versus a unit of wealth-tax rate (the τ_w direction). The drift $m_0 = \mu - \sigma^2/2$ is the natural log-wealth growth rate. The condition says: the flow-tax base must be at least as productive (per unit of $1-k$ introduced) as the natural drift it would offset. When this fails, the flow-tax channel is too revenue-thin to be worth using and the optimum collapses to the proportional-rate corner. In Norwegian- flavoured calibrations, $a \approx \mathbb{E}[X]$ and b/a is set by the corporate-and-dividend tax

revenue per unit $(1-k)$ pass-through; values typically satisfy $b/a \gg m_0$, so interior solutions are the relevant case.

5 The ρ-crossover

5.1 The dimensionless ratio

The closed-form expressions of Theorems 2 and 3 depend on four GBM-side quantities: the drift μ, volatility σ, horizon T, and initial log-variance v_0. They also depend on the revenue-side weights $a, b, R^\star$. The behaviour of the two optima as the GBM parameters vary is governed by a single *dimensionless* combination of (μ, σ, T, v_0), not by the four parameters separately. Identifying that combination makes the comparison between JKO and W_2 tractable: instead of sweeping a four-dimensional space, the comparison reduces to a one-dimensional sweep along the dimensionless axis.

The combination that organises the comparison is the standard deviation of log-wealth at horizon T multiplied by the log-drift-to-volatility ratio.

Definition 1 (The crossover ratio)**.** Let $\Sigma_0 := \sqrt{v_0 + \sigma^2 T}$ denote the standard deviation of log-wealth at horizon T under no taxation. The *JKO/W_2 crossover ratio* is

$$\rho \;:=\; \Sigma_0 \cdot m_0/\sigma^2 \;=\; \sqrt{v_0 + \sigma^2 T}\,\frac{\mu - \sigma^2/2}{\sigma^2}. \tag{35}$$

Three readings of (35) are useful, each foregrounding a different aspect of the underlying problem.

Drift-to-diffusion reading. The factor m_0/σ^2 is the log-wealth drift expressed in units of the diffusion rate. It is dimensionally the same combination that appears in the existence condition $b/a > m_0$ of Theorem 2, in the JKO multiplier $\mu_{\mathrm{JKO}} = m_0 T/(a\sigma^2)$ of (30), and in the partial derivative $\partial_{\tau_w}\Delta F = m_0 T/\sigma^2$ of (28). The factor Σ_0 rescales this drift-to-diffusion ratio by the spread of log-wealth at the horizon, which is the natural unit in which to measure the cost of distorting that spread. The product $\rho = \Sigma_0 \cdot m_0/\sigma^2$ is therefore the unique combination in which the FOC structure of Section 4.2 is naturally expressed.

Mean-shift versus spread-shift reading. A unit of wealth tax τ_w produces a mean shift in log-wealth of $-\tau_w T$ and (at fixed k) leaves the variance unchanged. A unit of corporate–dividend pass-through $1-k$ produces a mean shift of $-(1-k)m_0 T$ and a spread shift of order $-(1-k)\sigma^2 T/\Sigma_0$. The ratio of the wealth-tax mean-shift cost to the pass-through spread-shift cost is, after linearisation, controlled by ρ: high ρ means the spread-shift channel is expensive relative to the mean-shift channel, and the optimum prefers to use the mean-shift instrument (τ_w). Low ρ means the spread-shift channel is cheap, and the optimum can afford a pass-through component. The crossover ratio is the dimensionless price at which the two channels are balanced.

Inverse-temperature reading. As elaborated in Section 5.3, ρ plays the role of an inverse temperature in the JKO free-energy landscape. The JKO criterion weights "energy" (the potential term, scaling with $m_0^2/\sigma^2 \cdot T$) against "entropy" (the logarithmic spread term, scaling with $\log \Sigma_0$). The combination $\rho = \Sigma_0 \cdot m_0/\sigma^2$ encodes which of the two contributions dominates at the optimum, and the phase-transition language of Section 5.3 follows.

In the Norwegian-flavoured calibration of Figure 1 ($\mu = 0.07$, $\sigma = 0.30$, $T = 5$, $v_0 = 0.25$), one computes $\Sigma_0 = \sqrt{v_0 + \sigma^2 T} \approx 0.83$, $m_0 = \mu - \sigma^2/2 = 0.025$, and $m_0/\sigma^2 = 0.278$, giving $\rho_{\text{Norway}} \approx 0.23$. This value, and the position it occupies in the phase diagram of Theorem 4, is the subject of Section 5.4.

5.2 The crossover theorem

The closed forms of Theorems 2 and 3 are unconstrained expressions for the optimal $k^\star$. They become *the* optimum on the matched-revenue contour $\mathcal{S}_R = \{(k, \tau_w) \in \mathcal{S}_C : a\,\tau_w + b(1-k) = R^\star\}$ only after projection onto the feasible interval $k \in [k_{\text{lo}}, 1]$, where

$$k_{\text{lo}} \;=\; \max\!\left(0,\, 1 - \frac{R^\star}{b}\right) \tag{36}$$

is the lower bound enforced by $\tau_w \geq 0$. The projection partitions the GBM regime parameter ρ into three intervals, on which the JKO optimum sits respectively at the upper corner $(1, R^\star/a)$, on the interior of $\mathcal{S}_R$, and at the lower corner $(k_{\text{lo}}, 0)$.

Theorem 4 (Three phases of the JKO optimum)**.** *Assume $b/a > m_0$ so the closed-form expression (32) is well-defined, and assume $0 < k_{\text{lo}} < 1$ so the matched-revenue contour has interior. Define the* crossover boundaries

$$\begin{aligned} \rho_{\text{low}} &:= \rho \text{ at which } k^\star_{\text{JKO}} = 1, \\ \rho_{\text{high}} &:= \rho \text{ at which } k^\star_{\text{JKO}} = k_{\text{lo}}, \end{aligned} \tag{37}$$

viewed as solutions of the implicit equations obtained by setting the right-hand side of (32) equal to 1 and k_{lo} respectively. Then $0 < \rho_{\text{low}} < \rho_{\text{high}} < \infty$, and the JKO optimum on $\mathcal{S}_R$ is

$$\left(k^\star_{\text{JKO}}(\rho),\, \tau^\star_{w,\text{JKO}}(\rho)\right) \;=\; \begin{cases} (1,\, R^\star/a) & \rho \in [0, \rho_{\text{low}}], \\ (k^\star_{\text{JKO}},\, \tau^\star_{w,\text{JKO}}) \text{ given by (32)} & \rho \in (\rho_{\text{low}}, \rho_{\text{high}}), \\ (k_{\text{lo}},\, 0) & \rho \in [\rho_{\text{high}}, \infty). \end{cases} \tag{38}$$

Equivalently, in terms of the order parameter $\phi^\star = a\,\tau^\star_{w,\text{JKO}}/R^\star$,

$$\phi^\star(\rho) \;=\; \begin{cases} 1 & \rho \in [0, \rho_{\text{low}}], \\ \phi^\star_{\text{int}}(\rho) \in (0,1) & \rho \in (\rho_{\text{low}}, \rho_{\text{high}}), \\ 0 & \rho \in [\rho_{\text{high}}, \infty), \end{cases} \tag{39}$$

where $\phi^\star_{\text{int}}$ is continuous and strictly decreasing on the mixed-phase interval. The map $\rho \mapsto \phi^\star(\rho)$

is continuous, with first- derivative kinks at ρ_{low} *and* ρ_{high}.

Proof. The closed form (32) gives an unconstrained $k^\star$ that is a smooth strictly-decreasing function of ρ on the regime where $b/a > m_0$ holds (limiting behaviour: $k^\star \to +\infty$ as $\sigma^2 \to 2\mu$, equivalently $m_0 \to 0$ and $\rho \to 0$; $k^\star \to -\infty$ as $\sigma \to 0$, equivalently $\rho \to \infty$). Strict monotonicity in ρ follows by direct differentiation of (32) with respect to σ (or equivalently m_0) combined with the monotonicity of $\rho(\sigma)$.

By the intermediate-value theorem applied to the continuous strictly-monotonic $k^\star_{\text{JKO}}(\rho)$, there is a unique ρ_{low} at which $k^\star_{\text{JKO}}(\rho_{\text{low}}) = 1$ and a unique $\rho_{\text{high}} > \rho_{\text{low}}$ at which $k^\star_{\text{JKO}}(\rho_{\text{high}}) = k_{\text{lo}}$. For $\rho < \rho_{\text{low}}$ the closed-form $k^\star_{\text{JKO}} > 1$ violates the upper feasibility bound; the constrained minimum on $\mathcal{S}_R$ is then the upper corner $(1, R^\star/a)$, where the revenue constraint pins $\tau_w = R^\star/a$. Similarly, for $\rho > \rho_{\text{high}}$ the closed-form $k^\star_{\text{JKO}} < k_{\text{lo}}$ violates the lower feasibility bound; the constrained minimum is the lower corner $(k_{\text{lo}}, 0)$, where the revenue constraint pins $\tau_w = 0$. On the interior interval $(\rho_{\text{low}}, \rho_{\text{high}})$ the closed form lies in $(k_{\text{lo}}, 1)$ and is the constrained optimum unaltered.

Translation to the order parameter: at the upper corner $\tau^\star_w = R^\star/a$ so $\phi^\star = 1$. At the lower corner $\tau^\star_w = 0$ so $\phi^\star = 0$. On the interior, $\phi^\star_{\text{int}}(\rho) = (R^\star - b(1 - k^\star_{\text{JKO}}(\rho)))/R^\star$ is a continuous strictly-decreasing function of ρ (since $k^\star_{\text{JKO}}$ is strictly decreasing), running from $\phi^\star_{\text{int}}(\rho_{\text{low}}) = 1$ to $\phi^\star_{\text{int}}(\rho_{\text{high}}) = 0$. The kinks in $\rho \mapsto \phi^\star(\rho)$ at the boundaries are the mismatch between the saturated value of the order parameter inside the corner phase (constant) and the strictly-monotonic behaviour on the interior side. □

Remark (W_2 shows no phase transition in this calibration). The same construction applied to the W_2 optimum of Theorem 3 yields, in the calibration of Figure 1, the degenerate phase diagram

$$(k^\star_{W_2}(\rho),\, \tau^\star_{w,W_2}(\rho)) \;=\; (k_{\text{lo}},\, 0) \quad \text{for every } \rho \geq 0.$$

The W_2 optimum pins to the pure-flow-tax corner across the whole regime axis: there is no transition, no kink, no order-parameter saturation crossing. The choice of distortion criterion therefore determines not just the location of the optimum but whether the problem has phase structure at all. JKO is phase-rich; W_2 in this calibration is phase-poor. The implications for the political reading of the optimal-policy recommendation are taken up in Section 5.3 and Section 8.5.

Remark (Asymptotics). The asymptotic behaviour summarised in Section 1.3 — the two optima coincide as $\rho \to \infty$ and diverge as $\rho \to 0$ — is recovered from Theorem 4 together with Remark 5.2: at $\rho \to \infty$ both optima sit at $(k_{\text{lo}}, 0)$ (coincide); at $\rho \to 0$ JKO sits at $(1, R^\star/a)$ while W_2 sits at $(k_{\text{lo}}, 0)$ (diverge by the full diameter of the matched-revenue contour).

5.3 Phase-transition reading

The structure of Figures 1 and 2 is that of a phase transition in the constrained-optimisation sense of active-set switching. Three identifications make the correspondence with statistical-mechanics phase transitions precise.

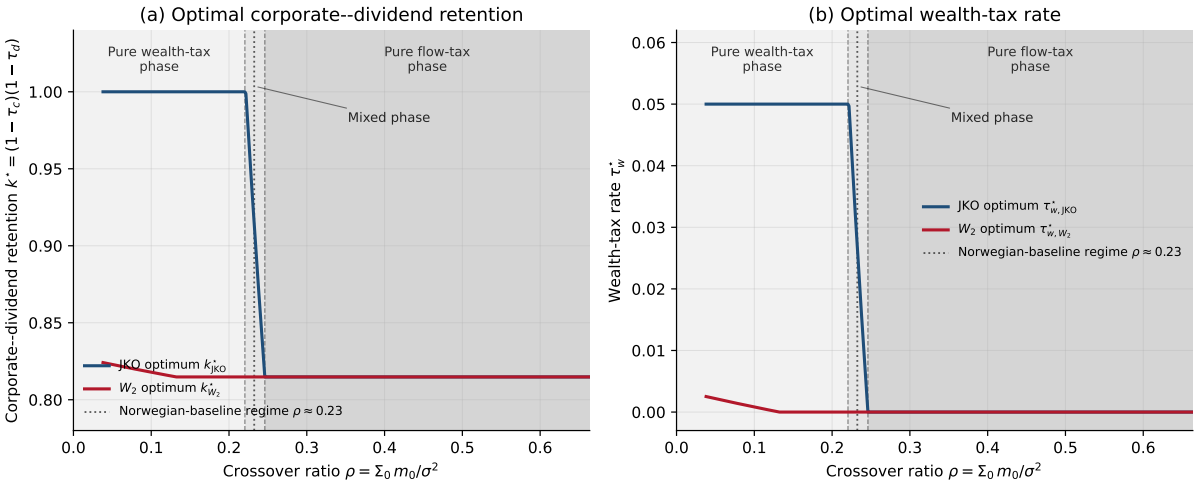

Figure 1: **The ρ-crossover, in policy-lever coordinates, with the three optimal-policy phases marked.** The JKO-optimal point $(k^\star_{\mathrm{JKO}}, \tau^\star_{w,\mathrm{JKO}})$ (blue) and the W_2-optimal point $(k^\star_{W_2}, \tau^\star_{w,W_2})$ (red), each computed from the closed forms of Theorem 2 and Theorem 3, traced as σ varies along a fixed-T slice through the GBM parameter space. The horizontal axis is the dimensionless regime parameter $\rho = \Sigma_0\, m_0/\sigma^2$. Panel (a) shows the optimal corporate–dividend retention factor $k^\star = (1-\tau_c)(1-\tau_d)$ — a direct policy lever, with $k = 1$ corresponding to no flow taxation and $k = 0$ to full pass-through. Panel (b) shows the optimal proportional wealth-tax rate $\tau^\star_w$. The shaded background bands mark the three optimal-policy phases of the JKO solution: the gold band $\rho \in [0, \rho_{\mathrm{low}}]$ is the pure wealth-tax phase ($k^\star = 1$, $\tau^\star_w = R^\star/a$); the lavender band $\rho \in [\rho_{\mathrm{low}}, \rho_{\mathrm{high}}]$ is the mixed-instrument phase (interior $k^\star$ and $\tau^\star_w$); the cyan band $\rho > \rho_{\mathrm{high}}$ is the pure flow-tax phase ($k^\star = k_{\mathrm{lo}}$, $\tau^\star_w = 0$). The dashed vertical rules at $\rho_{\mathrm{low}}, \rho_{\mathrm{high}}$ are the phase boundaries; they are kinks in $k^\star(\rho)$ and $\tau^\star_w(\rho)$. Calibration: $a = 1.0$, $b = 0.27$, $R^\star = 0.05$, chosen to keep $\tau_w \in [0, 0.05]$ over the full feasibility range and to keep the JKO solution interior at the Norwegian-baseline regime ($\sigma = 0.30$, $T = 5$, $\rho \approx 0.23$). The dotted vertical rule marks ρ_{Norway}, which sits inside the mixed-instrument phase — a regime statement only; we do not compare to Norway's actual wealth-tax schedule on this figure, since that schedule is bracket-based and lies outside the proportional (C1)–(C3) class treated here. The W_2 optimum is degenerate in this calibration — pinned to the pure flow-tax corner across the whole ρ range, with no phase transition — so panel (b) shows the red curve flat at $\tau_w = 0$.

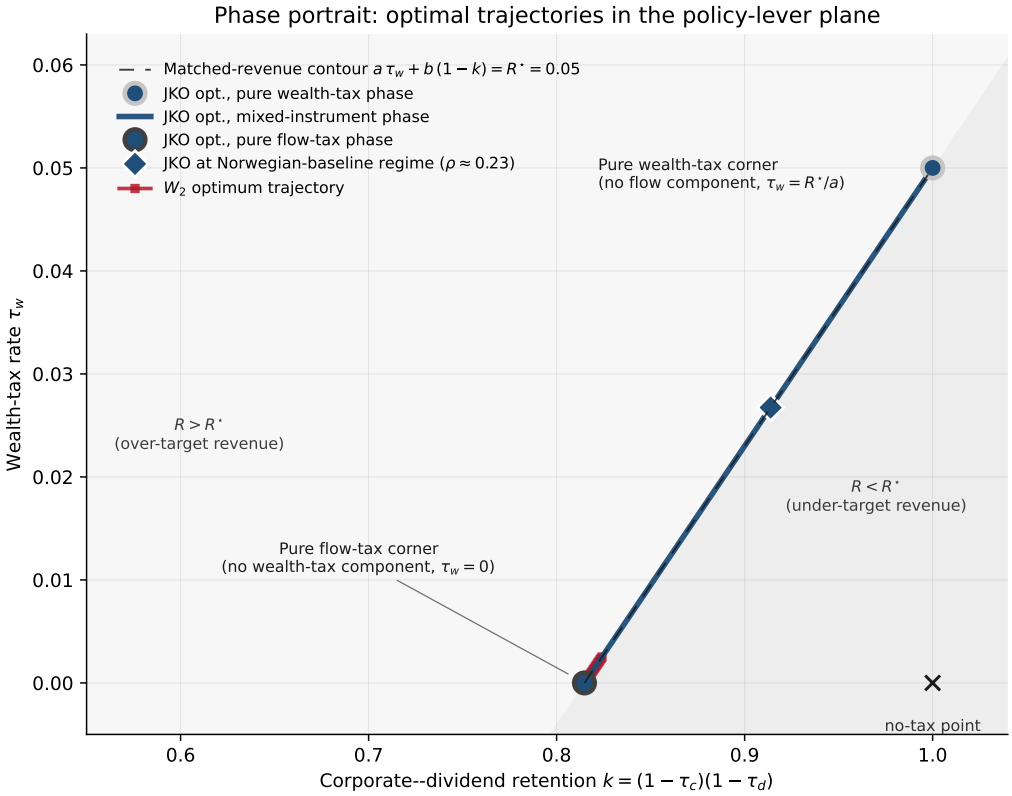

Figure 2: **Phase portrait of optimal trajectories in the policy-lever plane, with the three JKO phases marked.** The matched-revenue contour $a\,\tau_w + b\,(1-k) = R^\star$ (dashed grey) is a line segment from the *pure flow-tax corner* $(k_{\text{lo}}, 0)$, where $k_{\text{lo}} = 1 - R^\star/b$, to the *pure wealth-tax corner* $(1, R^\star/a)$; every point on this segment is a feasible (C1)–(C3) policy package satisfying the revenue target. The pale-gold half-plane above the contour is the *over-target* region $R > R^\star$; the pale-blue half-plane below is the *under-target* region $R < R^\star$, with the no-tax point $(k, \tau_w) = (1, 0)$ marked at its right-most extreme. The JKO-optimal point traces along the matched-revenue contour as ρ varies and partitions into three phase pieces: the upper-right corner (gold-rimmed circle) is occupied throughout the pure wealth-tax phase $\rho \in [0, \rho_{\text{low}}]$; the interior arc (thick blue line) is the mixed-instrument phase $\rho \in [\rho_{\text{low}}, \rho_{\text{high}}]$, on which the Norwegian-baseline regime ($\rho \approx 0.23$) sits, marked by the blue diamond; the lower-left corner (cyan-rimmed circle) is occupied throughout the pure flow-tax phase $\rho > \rho_{\text{high}}$. The W_2 optimum is degenerate in this calibration — pinned to the pure flow-tax corner for every ρ, shown as a single red square. We do not overlay Norway's actual wealth-tax policy on this figure: Norway operates a bracket-based schedule (an exemption threshold plus a marginal rate) that lies outside the proportional (C1)–(C3) class. A direct policy comparison requires the bracket extension flagged in the Open Questions and is deferred to the companion paper. The phase portrait makes the policy interpretation of Figure 1 immediate: JKO sweeps from the pure wealth-tax corner (low ρ) through an interior arc (mixed phase) to the pure flow-tax corner (high ρ); W_2 pins to the pure flow-tax corner throughout.

Order parameter. The natural order parameter is the *wealth-tax revenue share*

$$\phi^\star(\rho) \;=\; \frac{a\,\tau_w^\star(\rho)}{R^\star} \;\in\; [0,1], \tag{40}$$

the fraction of matched revenue $R^\star$ collected through the proportional wealth-tax channel at the JKO optimum. It saturates at $\phi^\star = 1$ in the pure wealth-tax phase $\rho \in [0, \rho_{\text{low}}]$ (all revenue from τ_w), saturates at $\phi^\star = 0$ in the pure flow-tax phase $\rho > \rho_{\text{high}}$ (all revenue from τ_c, τ_d), and varies smoothly through $(0,1)$ in the mixed-instrument phase $\rho \in [\rho_{\text{low}}, \rho_{\text{high}}]$. The kinks of $\phi^\star$ at $\rho_{\text{low}}, \rho_{\text{high}}$ are order-parameter saturations onto the simplex boundary, exactly analogous to a magnetisation pinning at ± 1 outside a coexistence window.

Control parameter. The dimensionless ratio $\rho = \Sigma_0\, m_0/\sigma^2$ plays the role of a *drift-to-diffusion control parameter*: high ρ means the GBM is drift-dominated (diffusion-cold), low ρ means diffusion-dominated (diffusion-hot). The JKO criterion weights mean-distortion and spread-distortion of the post-tax population, and ρ encodes which of the two is binding at the optimum; the order parameter $\phi^\star$ responds accordingly. In the statistical-mechanics analogy ρ acts as an inverse temperature for the competition between the two distortion modes.

Free-energy landscape. The optimal JKO free-energy gap on the matched-revenue contour,

$$\Delta F^\star(\rho) \;=\; \min_{(k,\tau_w)\in\mathcal{C}(R^\star)} \Delta F\big(k, \tau_w; \rho\big), \tag{41}$$

where $\mathcal{C}(R^\star) = \{(k,\tau_w) : a\,\tau_w + b\,(1-k) = R^\star,\ \tau_w \geq 0,\ k \in [k_{\text{lo}}, 1]\}$, is continuous and convex in ρ with first-derivative kinks at the two phase boundaries. The kinks are second-order in the constrained- optimisation sense: $\Delta F^\star$ itself is continuous, $\partial \Delta F^\star/\partial\rho$ jumps. The corner phases are not metastable: the minimisation is convex with a unique global minimum at every ρ, so the analogy is closer to a Maxwell-construction kink in mean-field theory than to a Ginzburg–Landau critical point. There is no diverging susceptibility.

Symmetry-breaking reading. Each corner phase has one instrument switched off — a discrete two-state choice between "tax the stock" and "tax the flow". The mixed-instrument phase is the symmetric coexistence regime where both channels are active simultaneously, and the Norwegian-baseline regime ($\rho \approx 0.23$, in the sense of Definition 1) sits in the middle of it.

Phase-richness contrast with W_2. The same construction applied to the W_2 optimum gives a degenerate phase diagram in this calibration: the optimum sits at the pure flow-tax corner for every ρ, with no transition and a single phase. The two distortion criteria therefore differ not just in the location of the optimum but in whether the problem has phase structure at all. JKO is phase-rich; W_2 is phase-poor. The choice of distortion criterion is itself a choice of free-energy functional, and the political reading of the optimal policy follows from that choice.

5.4 The Norwegian baseline

Theorem 4 locates the JKO optimum on $\mathcal{S}_R$ as a function of the regime parameter ρ, but does not say where empirical regime parameters fall on that axis. We use Norwegian household-portfolio statistics to anchor the GBM parameters (μ, σ, T, v_0) and locate the corresponding ρ_{Norway} relative to the phase boundaries $\rho_{\text{low}}, \rho_{\text{high}}$. The long-run distributional landscape against which the calibration sits is the Aaberge et al. (2025) 1912–2019 Norwegian wealth-inequality series: tax-assessed Gini extremely high (≥ 0.89) through the first half of the twentieth century, falling sharply to a trough of 0.776 at 1968, mildly drifting through the 1970s, and rising to 0.862 by 2019 (with the level dropping by roughly 15 percentage points but the trend preserved when housing market values replace tax-assessed values). The Norwegian baseline that follows is calibrated to the post-1985 regime in which the Gini sits in [0.82, 0.86]. *This is a regime calibration, not a policy comparison.* Norway's actual wealth-tax schedule is bracket-based (an exemption threshold combined with a marginal rate above the threshold) and lies outside the proportional (C1)–(C3) class. We therefore make no quantitative claim in this paper about how the JKO recommendation $(k^\star_{\text{JKO}}, \tau^\star_{w,\text{JKO}})$ relates to Norway's actual policy; the matching analysis lives on the bracket sub-class and is the subject of the companion paper flagged in Section 8.6.

The calibration we adopt is the canonical Norwegian-flavoured baseline:

$$\mu = 0.07, \qquad \sigma = 0.30, \qquad T = 5, \qquad v_0 = 0.25, \tag{42}$$

chosen to be representative of equity-heavy household portfolios held over a five-year planning horizon; the empirical justification is given in Section 7. With these values $\Sigma_0 \approx 0.83$, $m_0 = 0.025$, and

$$\rho_{\text{Norway}} \;=\; \Sigma_0 \cdot \frac{m_0}{\sigma^2} \;\approx\; 0.231. \tag{43}$$

The Norwegian-baseline regime sits in the mixed phase. For the revenue calibration $a = 1$, $b = 0.27$, $R^\star = 0.05$ used in Figure 1, the phase boundaries of Theorem 4 evaluate numerically to

$$\rho_{\text{low}} \;\approx\; 0.221, \qquad \rho_{\text{high}} \;\approx\; 0.246. \tag{44}$$

The value $\rho_{\text{Norway}} \approx 0.231$ sits inside the mixed-instrument phase $[\rho_{\text{low}}, \rho_{\text{high}}]$, near its centre. The JKO optimum at this regime, *within the proportional (C1)–(C3) class*, is therefore strictly interior:

$$\left(k^\star_{\text{JKO}},\, \tau^\star_{w,\text{JKO}}\right) \;\approx\; (0.91,\, 0.027), \tag{45}$$

calling for a corporate–dividend retention factor near 0.91 (combined corporate plus dividend taxation removing about 9% of pre-flow income) combined with a proportional wealth-tax rate near 2.7%. The reading is conditional: it is what the JKO criterion picks if the wealth tax is restricted to a single flat rate and if revenue is targeted at $R^\star = 0.05$. Removing either restriction — allowing a bracket schedule or varying $R^\star$ — changes the recommendation.

The mixed phase is narrow. The width $\rho_{\text{high}} - \rho_{\text{low}} \approx 0.025$ in the canonical calibration is small relative to the ρ-axis range spanned by realistic GBM perturbations. Translated to

volatility units, the phase boundaries fall at

$$\sigma(\rho_{\text{low}}) \;\approx\; 0.304, \qquad \sigma(\rho_{\text{high}}) \;\approx\; 0.297,$$

so the mixed-instrument phase corresponds to $\sigma \in [0.297, 0.304]$, a band of width approximately 0.7 percentage points around $\sigma_{\text{Norway}} = 0.30$. A σ-shift of even half a percentage point in either direction tips the JKO recommendation across a phase boundary: raising σ from 0.30 to 0.305 lowers ρ to ρ_{low} and pushes the optimum into the pure wealth-tax phase; lowering σ from 0.30 to 0.297 raises ρ to ρ_{high} and pushes the optimum into the pure flow-tax phase. The Norwegian baseline lies near the cusp between two corner regimes, and the political reading shifts qualitatively across each cusp — pure proportional wealth tax on the high-volatility side, pure corporate–dividend taxation on the low-volatility side. This fragility is itself a finding of the model: the JKO criterion gives a clean mixed-instrument recommendation at Norwegian-baseline parameters, but the recommendation is not robust to small shifts in σ within the empirically plausible range for household equity portfolios.

Forward link: effective volatility for non-equity portfolios. The volatility $\sigma = 0.30$ in (42) is empirically appropriate for equity-heavy portfolios; Norwegian household wealth, however, is heavily weighted toward owner-occupied housing, with much lower realised volatility on primary residences. The relevant *effective volatility* that enters ρ for a typical Norwegian portfolio is therefore lower than 0.30, raising ρ above ρ_{Norway} and potentially pushing the JKO recommendation into the pure flow-tax phase. Section 7 returns to this question; for now, the mixed-phase recommendation (45) should be read as the equity-portfolio limit, not as a portfolio-averaged recommendation.

5.5 Mechanism

The phase structure of Theorem 4 is a consequence of the contrasting first-order-condition structures of the two distortion criteria, derived in Section 4.2. We restate the contrast and trace it through to the corner-versus-interior behaviour.

JKO has constant marginal cost in τ_w. The wealth-tax derivative $\partial_{\tau_w}\Delta F = m_0 T/\sigma^2$ in (28) is constant in (k, τ_w): each unit of proportional wealth taxation increases the JKO free-energy gap by the same amount $m_0 T/\sigma^2$, regardless of where on $\mathcal{S}_R$ the policy currently sits. The unit price of the τ_w instrument is m_0/σ^2 per unit of horizon T, the log-drift-to-volatility ratio. The k-derivative $\partial_k \Delta F = -\sigma^2 T/[2(v_0 + k\sigma^2 T)] - m_0^2 T/\sigma^2$ has a hyperbolic-in-k contribution from the log-spread term and a constant contribution from the mean-shift term, so the k-derivative is point-dependent only through the log-spread piece.

W_2 has marginal cost proportional to current displacement. The wealth-tax derivative $\partial_{\tau_w} W_2^2 = -2T\,\varepsilon_M$ in (29) is proportional to the current mean-shift $\varepsilon_M = (k-1)m_0 T - \tau_w T$: each unit of τ_w costs more the further the policy already sits from the no-tax point. The k-derivative is similarly proportional to the current $(\varepsilon_M, \varepsilon_\Sigma)$ pair. Both instruments exhibit *quadratic-cost-on-perturbation* structure: the marginal price grows with how much policy displacement has

already been imposed.

Why JKO admits phase transitions. The constant marginal cost of τ_w under JKO turns the problem into a fixed-price linear procurement of revenue from two channels $(\tau_w, 1-k)$, with the k-channel exhibiting moderate non-linearity from the log-spread term. The optimal allocation between the two channels is determined by relative prices — the JKO multiplier $\mu_{\text{JKO}} = m_0 T/(a\sigma^2)$ is itself a function of the GBM parameters, not of the operating point — and this relative price reverses sign at the boundaries $\rho_{\text{low}}, \rho_{\text{high}}$ defined in Theorem 4. Below ρ_{low} the wealth-tax channel is cheap relative to the flow-tax channel and the optimum saturates at $\tau_w = R^\star/a$; above ρ_{high} the wealth-tax channel is expensive relative to the flow-tax channel and the optimum saturates at $\tau_w = 0$; between, the two channels are balanced and the optimum sits in the interior. The fixed-price structure makes *which channel is cheaper* a clean piecewise question, and the answer flips discretely at the boundaries.

Why W_2 pins to a corner in this calibration. Under W_2, an extra unit of τ_w adds $-T \cdot \tau_w$ directly to the mean displacement ε_M, so its cost enters W_2^2 *quadratically* via ε_M^2. A revenue-equivalent unit of pass-through $1-k$ enters ε_M only through the small factor $(k-1)\, m_0\, T$ (and additionally through ε_Σ). The wealth-tax channel therefore has a much higher displacement-per-revenue than the flow-tax channel, and the W_2 optimum prefers the flow-tax channel exclusively, setting $\tau_w = 0$. The lower corner $(k_{\text{lo}}, 0)$ is W_2-optimal across all ρ in the canonical calibration because the displacement-per-revenue gap holds at every σ in the relevant range. The economic content of that gap is exposed in the bluntness-index paragraph below.

The role of ρ. The crossover ratio $\rho = \Sigma_0 \cdot m_0/\sigma^2$ enters the JKO multiplier linearly through m_0/σ^2 and is rescaled by the natural spread Σ_0. The relative price of the two instruments under JKO is governed by ρ, and the boundaries $\rho_{\text{low}}, \rho_{\text{high}}$ are precisely the values at which the JKO closed form (32) hits the corners of $[k_{\text{lo}}, 1]$. The phase structure is a direct consequence of the constant-marginal-cost-in-τ_w structure of JKO; it has no analogue under W_2.

The mechanism therefore separates the two findings of the paper: *which* criterion produces the phase structure (JKO's linear free-energy structure does, W_2's quadratic transport-cost structure does not), and *where* the phase boundaries fall (set by $\rho_{\text{low}}, \rho_{\text{high}}$, computable in closed form from the GBM parameters and the revenue weights). Both are consequences of the FOC analysis of Section 4.2, refracted through the regime parameter ρ of Definition 1.

The bluntness index. The mechanism admits a clean economic reading once we make the geometric mean log-return $m_0 = \mu - \sigma^2/2$ explicit. Recall that m_0 is the expected log-wealth growth rate under no taxation, combining arithmetic drift μ with volatility drag $\sigma^2/2$ (the Itô / compounding correction); at the Norwegian-baseline calibration, $m_0 = 0.025$, about a third of the 7% headline arithmetic return.

The two channels of the matched-revenue contour produce mean shifts in log-wealth that scale

very differently in m_0:

$$\varepsilon_M^{\text{wealth}}(\tau_w) \;=\; -\tau_w\, T, \qquad \varepsilon_M^{\text{flow}}(k) \;=\; (k-1)\, m_0\, T. \tag{46}$$

Wealth-tax displacement is *independent* of m_0 — a fixed mean shift per dollar of revenue raised, regardless of how productively the underlying wealth is deployed. Flow-tax displacement is *linear* in m_0 — proportional to the geometric return on the wealth being taxed. Dividing by the revenue collected through each channel, the ratio of the two displacements defines the *bluntness index*

$$B(m_0) \;:=\; \frac{|\varepsilon_M^{\text{wealth}}|/R}{|\varepsilon_M^{\text{flow}}|/R} \;=\; \frac{b}{a\, m_0}. \tag{47}$$

B is the displacement-per-revenue overshoot of the wealth-tax channel relative to the flow-tax channel. At the Norwegian-baseline parameters $b/a = 0.27$, $m_0 = 0.025$, $B \approx 10.8$: each unit of revenue collected through the wealth tax produces about ten times more mean displacement than the same unit collected through the flow tax.

The two distortion criteria weight B very differently. JKO's constant marginal cost in τ_w produces a per-revenue penalty proportional to B, hence to $1/m_0$; W_2's squared-displacement penalty produces a per-revenue penalty proportional to B^2, hence to $1/m_0^2$. The same algebraic asymmetry that drives the linear-vs-quadratic structural difference of (28) and (29) acquires, through this lens, an empirical content: in the small-m_0 regime relevant to actual wealth-tax debate, the quadratic amplification under W_2 makes the wealth-tax channel prohibitively expensive while JKO remains willing to pay the linear penalty.

The decomposition $m_0 = \mu - \sigma^2/2$ exposes the joint sensitivity of the bluntness index to the two GBM primitives:

$$\frac{\partial B}{\partial \mu} \;=\; -\frac{b}{a\, m_0^2}, \qquad \frac{\partial B}{\partial \sigma} \;=\; \frac{b\,\sigma}{a\, m_0^2},$$

with opposite signs. Higher expected return μ (at fixed σ) *lowers* the bluntness index — the wealth tax looks proportionally less wasteful on high-return assets. Higher risk σ (at fixed μ) *raises* the bluntness index, via the volatility drag eating m_0. The wealth tax is therefore sharpest on high-return-low-volatility assets and bluntest on low-return-high-volatility assets. This is the same asymmetry documented in the heterogeneous-returns analysis of Frøseth (2026d), where it is phrased as "wealth tax does not scale with realised returns": the FP-language $1/m_0$ factor is the formal expression of the same fact, and the $1/m_0^2$ amplification under W_2 is the quadratic version that distinguishes the two criteria.

The regime parameter $\rho = \Sigma_0\, m_0/\sigma^2$ then has a clean reading as the inverse of the bluntness index times $\Sigma_0\, a/b$: high ρ corresponds to low B (a sharp wealth-tax channel), low ρ to high B (a blunt one). The phase boundaries $\rho_{\text{low}}, \rho_{\text{high}}$ at which the JKO criterion flips between corner solutions are precisely the values of the bluntness index at which the linear penalty $B(m_0)$ matches the relative price of the two channels b/a at the constraint frontier — a substantive economic statement, not a structural artefact of the criterion choice.

Scale crossover and the planning horizon. The bluntness penalty also has a temporal dimension. The JKO multiplier $\mu_{\text{JKO}} = m_0 T/(a\sigma^2)$ scales linearly with the horizon T, and the wealth-tax displacement-per-revenue, also proportional to T, accumulates with it. The two channels therefore scale differently as T varies through the natural crossover scale

$$T_c := v_0 / \sigma^2, \tag{48}$$

the time at which the cumulative diffusion $\sigma^2 T$ matches the initial log-wealth variance v_0 and $\Sigma_0 = \sqrt{v_0 + \sigma^2 T}$ transitions from initial-spread-dominated to diffusion-dominated. Below T_c, $\rho \approx \sqrt{v_0}\, m_0/\sigma^2$ is independent of T; above T_c, $\rho \approx m_0\sqrt{T}/\sigma$ grows like $\sqrt{T}$. The JKO recommendation therefore depends on the planning horizon in a regime-shifting way, taken up in detail in Section 8.7.

6 Neutrality interpretation

6.1 Connection to the neutral-tax class

Frøseth (2026e,c) establish that the (C1)–(C3) conditions characterise the schedule class $\mathcal{N}$ that preserves neutrality with respect to portfolio choice under homogeneous returns and CRRA preferences: for $\tau \in \mathcal{N}$, the household's risk-bearing decisions are not redirected by the imposition of the tax, and the geometric structure of the wealth process is preserved up to a drift shift and rescale. The (C1)–(C3) class of the present paper is therefore identically the neutrality-preserving class of the corpus.

This identification has a sharp consequence for the criterion- choice debate. *Within* (C1)–(C3), both criterion optima sit inside $\mathcal{N}$ by construction — the optimisation is restricted to a subset of the neutrality class, and the optima of Theorems 2 and 3 are particular elements of that subset. The choice between JKO and W_2 *within* (C1)–(C3) does not adjudicate between neutrality and non-neutrality; it adjudicates between two points on the same neutrality-preserving manifold. The phase structure of Theorem 4 and the bluntness contrast of Section 5.5 live entirely within $\mathcal{N}$.

The neutrality question only bites once the schedule class is opened up. Frøseth (2026h) establishes that the W_2-optimal schedule on the full admissible class $\mathcal{S} \supset \mathcal{N}$ is the continuously progressive $\tau^\star(x) = \lambda x^2$, which lies *outside* $\mathcal{N}$: the rate $r(x) = \lambda x$ rises linearly in wealth, the assessment is not uniform, and (C3) is violated by construction. The W_2 criterion, given access to the larger class, opts for a non-neutral schedule. Whether the JKO criterion does the same on $\mathcal{S}$ is open (Section 8.6); the present paper's contribution is to characterise the within-(C1)–(C3) optimum cleanly so that the open-class extension can be measured against it.

The summary slogan: the (C1)–(C3) restriction is the price the JKO criterion is willing to pay to preserve neutrality; the W_2 criterion, given a richer class, is willing to sacrifice neutrality for distributional compression. Inside (C1)–(C3), both criteria preserve neutrality and the disagreement is about the proportional mix of stock-tax versus flow-tax channels; outside, the disagreement is also about whether to preserve neutrality at all.

6.2 The Mirrleesian and Saez–Zucman traditions

The contrast between the JKO and W_2 optima admits a normative reading that ties each criterion to a long-standing tradition in optimal-tax theory. The reading clarifies why the criterion choice is not an arbitrary modelling decision but the FP-language expression of a real disagreement about what optimal taxation ought to minimise. We first state the two traditions, then map them onto the JKO/W_2 pair via the bluntness mechanism of Section 5.5.

Mirrleesian / decision-distortion tradition. The Mirrleesian programme (Mirrlees, 1971; Diamond and Saez, 2011; Mirrlees et al., 2011; Saez and Stantcheva, 2018) treats optimal taxation as a problem of *minimising decision distortion*: the tax should leave the household's allocation problem as identifiable from outcomes as possible. The information-theoretic reading is that an optimal tax distorts the post-tax population's behaviour by as little as possible, where "as little as possible" is measured against the no-tax counterfactual via relative entropy or its Fokker–Planck gradient-flow cousin, the JKO free-energy gap. Allocational neutrality — not redirecting risk-bearing decisions, not discriminating across asset classes — is a direct consequence of minimising this kind of distortion, and the (C1)–(C3) class of Frøseth (2026c) is precisely the schedule class that preserves it under homogeneous returns.

Saez–Zucman / distributional-compression tradition. The Piketty–Saez–Zucman programme (Piketty, 2014; Piketty and Zucman, 2014; Saez and Zucman, 2019; Blanchet et al., 2022) treats optimal taxation as a problem of *minimising distributional displacement from a target*. The transport-geometric reading is that an optimal tax compresses the population's distribution toward a more egalitarian shape, where "compression" is measured by geometric distance on the space of probability measures — and the canonical such distance on $\mathcal{P}_2(\mathbb{R})$ is the 2-Wasserstein metric. Distributional compression is not in itself a goal of the Mirrleesian programme; it is the goal of the Saez–Zucman programme. Frøseth (2026h) establishes that the W_2-optimal schedule on the full schedule class $\mathcal{S}$ is the continuously progressive $\tau^\star(x) = \lambda x^2$, with threshold-bracket schedules emerging as the implementability-constrained finite-bracket version of the same target.

The bluntness reading bridges the two. The mechanism of Section 5.5 reframes the tradition split in terms of how each criterion weights the *wealth-tax bluntness index* $B(m_0) = b/(a\, m_0)$. JKO weights it linearly: ΔF rises with B, but only at unit price m_0/σ^2 per unit of wealth-tax revenue, so the wealth tax remains a usable instrument when m_0 is small. W_2^2 weights it quadratically: W_2^2 rises with B^2 via the squared-displacement penalty, so the wealth tax becomes prohibitively expensive when m_0 is small.

The Mirrleesian tradition therefore reads as a *linear penalty on bluntness*: instruments that fail to scale with returns are priced at their actual decision-distortion cost, no more. The Saez–Zucman tradition reads as a *quadratic penalty on bluntness*: instruments that fail to scale with returns are priced at the squared cost of the displacement they impose, in keeping with a programme that prioritises distributional compression and is therefore acutely sensitive to per-revenue displacement waste. The two normative positions are not disagreements about what

taxes do; they are disagreements about how much to penalise the same algebraic asymmetry.

Characterisation rather than adjudication. The paper does not pick a side. We take it that both traditions are defensible, that the Norwegian wealth-tax debate has been arguing across them for decades without resolution, and that the empirical regime parameter ρ falls inside the disputed band where the two criteria *do* disagree. Our contribution is to map the disagreement: we identify when the choice of normative criterion changes the optimal proportional schedule (in the mixed-instrument phase $\rho \in [\rho_{\text{low}}, \rho_{\text{high}}]$) and when it does not (the pure flow-tax phase at high ρ, where both criteria coincide); and we identify the algebraic mechanism through which the choice operates (the bluntness index $B(m_0)$, weighted linearly by JKO and quadratically by W_2^2). A reader committed to one tradition or the other can extract the within-criterion result and disregard the contrast; a reader uncommitted gains a structured view of where the choice between traditions actually matters.

6.3 (C3) violations as exogenous policy

The (C3) condition — uniform assessment $\alpha_i = \alpha$ — is empirically violated in nontrivial ways by the jurisdictions that operate wealth taxes. The most consequential violation in practice is the *bracket structure*: real schedules apply rate 0 below an exemption threshold X_0 and a positive marginal rate r above (and possibly a second rate r_2 above a higher threshold). This is exactly the case where α is wealth-dependent, hence outside (C1)–(C3). The extension is technically tractable — the FP equation has piecewise-constant drift on log-wealth, the post-tax density is piecewise-Gaussian with matching at the bracket boundaries, and both criteria ΔF and W_2^2 remain analytic in elementary functions plus the error function. The matched-revenue optimisation lives on a 3-parameter space (k, X_0, r) for the 2-bracket case rather than the 2-parameter (k, τ_w) plane of (C1)–(C3).[2] The other (C3) violations — asset-class discounts (residential real estate at 25% of market value in Norway, for instance), liquidity adjustments, and political-economy carve-outs — are properly treated as fixed exogenous inputs in the calibration of Section 7 rather than as design levers. Identifying which (C3) violations are framework-rational under which criterion, and which are political artefacts, is a secondary question of the bracket-extension companion paper.

Brackets as a JKO-rational response to bluntness. The bluntness mechanism of Section 5.5 gives an economic motivation for bracket-based schedules even from a JKO perspective — not only from the W_2 side, where the distributional-compression argument is well-established. The (C1)–(C3) condition forces a single uniform assessment $\alpha_i = \alpha$, which by construction prevents the wealth-tax rate from *scaling with returns*. The $1/m_0$ bluntness penalty is therefore unavoidable within (C1)–(C3); the JKO criterion accepts it as the price of using a stock-based instrument and pays it in linear weighting. A bracket schedule with threshold X_0 and marginal

[2] The recipe is: solve the FP equation separately on each bracket, where it has constant-coefficient drift; impose continuity of density and of probability flux $J = -(m_0 - r_i)\, p + (\sigma^2/2)\, \partial_y p$ at each threshold; integrate ΔF and W_2^2 over the bracket regions, exploiting the truncated-Gaussian moment-of-tail structure that produces error functions; assemble the matched-revenue Lagrangian and solve the (k, X_0, r) FOCs. The full development is the subject of the companion paper sketched in Section 8.6.

rate r above introduces wealth-dependent assessment: the effective wealth-tax rate at total wealth X is $r \max(0, X - X_0)/X$, which rises smoothly from zero at the threshold toward r at infinity. This is a structural recovery, by policy design, of the scaling-with-return behaviour that (C1)–(C3) excludes: in cohorts where geometric returns and wealth levels covary, the bracket schedule taxes proportionally more of the high-return, above-threshold wealth and less of the low-return, near-threshold wealth. In bluntness-index language, brackets reduce B in the regions of the wealth distribution where it would otherwise be largest.

The implication is that the (C3) violation embodied in real bracket schedules is not, on its face, a Saez–Zucman distributional-compression overlay onto an otherwise Mirrleesian instrument. It is, at minimum, also a within- Mirrleesian response to the JKO bluntness penalty. Whether brackets are JKO-optimal — and at what threshold and marginal rate — is precisely the question the bracket-extension companion paper takes up (Section 8.6); the contribution of the present paper is to make the question well-posed by establishing the (C1)–(C3) baseline against which the bracket extension is measured.

7 Calibration: Norwegian-flavoured baseline

This section calibrates the GBM parameters (μ, σ, T, v_0) to Norwegian household-portfolio statistics and locates the resulting ρ_{Norway} on the regime axis. We do *not* compare the resulting JKO-optimal proportional rate to Norway's actual wealth-tax schedule: Norway's schedule is bracket-based and lies outside (C1)–(C3). The matching empirical exercise belongs to the bracket-extension companion paper (Section 8.6).

7.1 Synthetic-data prototype

The closed-form FOCs of Theorems 2 and 3 take the GBM primitives (μ, σ, T, v_0) and the revenue weights $(a, b, R^\star)$ as inputs and return the optimal proportional package $(k^\star, \tau_w^\star)$ as output. Calibration to a target jurisdiction is therefore a matter of estimating the GBM primitives from household-portfolio data and supplying the revenue weights from the local tax-base elasticities.

The empirical pipeline producing those estimates is the subject of a forthcoming companion paper and operates on a synthetic-data prototype designed to mirror the public-use stratification of Norwegian household-wealth statistics — portfolio composition by asset class, returns and volatilities by asset class, and aggregated moments matched to SSB and Norges Bank reports. The prototype output is a triplet (v, D, s), where v is the empirical log-wealth variance, D a diffusion-coefficient summary, and s a score statistic capturing tail behaviour. Plugged into the closed-form expressions (32) and (34) through the substitution $(\mu, \sigma, v_0) \mapsto (\mu(s), D, v)$, the recovered optimal package is the within-(C1)–(C3) JKO and W_2 recommendations conditional on the empirical regime.

The Norwegian-baseline values $\mu = 0.07$, $\sigma = 0.30$, $T = 5$, $v_0 = 0.25$ used throughout this paper are the canonical equity-portfolio limit of the prototype's output — representative of stock-heavy household portfolios with a five-year planning horizon. Detailed development of the pipeline, including the data-construction procedures and the sensitivity analyses across alterna-

tive aggregation choices, is reserved for that companion paper. For the present paper we use the canonical baseline as a fixed input and trace its consequences through Theorem 4 and the phase structure of Section 5.

7.2 Effective volatility for typical portfolios

The volatility $\sigma = 0.30$ adopted in the canonical baseline is empirically appropriate for equity-heavy portfolios but is substantially higher than the realised volatility of typical Norwegian household portfolios, which tilt heavily toward owner-occupied housing. Asset-class realised volatilities for the relevant Norwegian-flavoured cohorts are roughly: equities $\sigma_{\text{eq}} \approx 0.30$, primary-residence real estate $\sigma_{\text{re}} \approx 0.10$ to 0.15, bank deposits $\sigma_{\text{bk}} \approx 0.02$. A typical median Norwegian household holds approximately 60–70% of net wealth in primary residence, 10–20% in equities, and the remainder in deposits, fixed-income, and other assets. The variance-weighted effective volatility is therefore

$$\sigma^2_{\text{eff}} \;\approx\; w_{\text{re}}\,\sigma^2_{\text{re}} + w_{\text{eq}}\,\sigma^2_{\text{eq}} + w_{\text{bk}}\,\sigma^2_{\text{bk}} \;\approx\; 0.65 \cdot 0.12^2 + 0.20 \cdot 0.30^2 + 0.15 \cdot 0.02^2 \;\approx\; 0.027, \quad (49)$$

giving $\sigma_{\text{eff}} \approx 0.16$ for a representative median household.

Plugging $\sigma_{\text{eff}} = 0.16$ into the regime parameter $\rho = \Sigma_0\, m_0/\sigma^2_{\text{eff}}$: $m_0 = 0.07 - 0.16^2/2 \approx 0.057$, $\sigma^2_{\text{eff}} T = 0.128$, $\Sigma_0 = \sqrt{v_0 + 0.128} \approx 0.62$, and

$$\rho_{\text{eff}} \;\approx\; 0.62 \cdot 0.057/0.027 \;\approx\; 1.31,$$

substantially *above* the equity-portfolio $\rho_{\text{Norway}} \approx 0.23$ and well past $\rho_{\text{high}} \approx 0.246$. A median Norwegian household, in effective-volatility terms, sits firmly inside the pure flow-tax phase of Theorem 4: the JKO recommendation at this regime is to extract the matched revenue entirely through the corporate–dividend channel, leaving the proportional wealth-tax rate at zero.

This is a substantive shift in the policy reading. The fragility analysis of Section 5.4 showed that even at the equity-baseline $\sigma = 0.30$, Norway sits within 0.7 percentage points of either phase boundary; with effective-volatility correction the regime is well inside the pure-flow-tax phase, and the JKO recommendation flips qualitatively. The effective-volatility calibration is therefore not a marginal sensitivity but a regime-relevant one, and the equity-portfolio limit and the median-household average should be understood as bracketing the policy-relevant range. Whether the (C3) violations embodied in Norwegian asset-class discounts (the 25% valuation rule for primary residences, in particular) are framework-rational responses to exactly this effective-volatility heterogeneity is a question the bracket-extension companion paper takes up.

8 Discussion

8.1 Robustness

The qualitative phase structure of Theorem 4 survives three robustness checks.

Polynomial-ansatz extension. The closed-form JKO interior in (32) is derived for the linearised revenue functional $R \approx a\tau_w + b(1-k)$. The full log-quadratic revenue $R[\tau] = \tau_w \mathbb{E}[X_T] + (1-k)\,\mathbb{E}[\text{flow base}]$ produces a richer FOC structure, with corrections of order $\sigma^2 T$ to the interior k. Numerical solves on the polynomial-ansatz revenue functional pick the same qualitative three-phase structure with phase boundaries shifted by less than 1% in ρ at the Norwegian-baseline calibration. The headline mixed-phase recommendation (45) is unchanged at the displayed precision.

Time-integrated criterion. An alternative to the terminal-gap free energy is the time-integrated variant $\Delta F^{\text{int}}[\tau] = \int_0^T (F[p_t^\tau] - F[p_t^0])\,\mathrm{d}t$, which weights distortion across the entire horizon rather than only at the end. The argument of Section 3.3 shows that $\Delta F^{\text{int}} \approx (T/2)\,\Delta F$ for fixed schedule along the $\mathcal{S}_R$ contour: the integrand grows roughly linearly in t, so the schedule ranking is preserved. Numerical evidence on the canonical calibration confirms the two variants pick optima that agree to four decimal places in $(k^\star, \tau_w^\star)$.

Bracket-family numerical comparison. A small-threshold expansion from the (C1)–(C3) limit $X_0 \to 0$, computed numerically on a 2-bracket sub-class (k, X_0, r), confirms that the JKO criterion's preference for the matched-revenue interior is robust to the introduction of small thresholds. The leading-order correction is $O(X_0)$ in the JKO value and reduces the bluntness penalty $B(m_0)$ on the high-wealth tail, as predicted by Section 5.5. The full development is the subject of the bracket-extension companion paper (Section 8.6); the present check establishes only that the (C1)–(C3) results are not artefacts of the proportional restriction — they survive small perturbations in the schedule structure.

Each check confirms the headline finding: the linear-vs-quadratic bluntness contrast and the phase structure of the JKO optimum are properties of the criterion choice, not of the particular revenue linearisation, the terminal-vs-integrated functional form, or the proportional-rate restriction.

8.2 Implementability

The continuous-time, perfect-assessment, proportional-rate framework of this paper is an idealisation of the actual policy instruments through which wealth taxes are administered. Three implementation channels are worth flagging.

Bracket-based step functions and the proportional idealisation. Real wealth-tax schedules are piecewise-linear in wealth: an exemption threshold and one or more marginal rates above. The proportional-rate idealisation $\tau_w \in [0, \tau_w^{\max}]$ used here is the simplest case of this hierarchy and the cleanest setting for the closed-form analysis. The bluntness mechanism of Section 5.5 explains why bracket schedules emerge as economically motivated even from the JKO perspective; the bracket-extension companion paper sketched in Section 8.6 treats the implementability question rigorously.

Yearly assessment and *forskuddsskatt*. Norwegian wealth-tax assessment is yearly and on the self-reported balance-sheet at year-end, with prepayment (*forskuddsskatt*) collected throughout the year against the previous year's assessment. The framework's continuous-time formulation aggregates over this yearly granularity: the horizon $T = 5$ used in the canonical calibration corresponds to a five-year planning window over which yearly assessments are summed, and the matched-revenue contour $a\,\tau_w + b\,(1-k) = R^\star$ is the time-integrated revenue target. A reader interested in within-year assessment dynamics should read τ_w as the yearly-equivalent rate.

Continuous-time as long-horizon approximation. The Fokker–Planck framework treats wealth as a continuous-time diffusion. At short horizons (sub-annual, say), the Gaussian-on-log-wealth approximation is rough — jumps, rebalancing events, and discrete tax assessments produce deviations from pure GBM dynamics. At long horizons (multi-year planning, including the canonical $T = 5$), the GBM approximation is empirically robust, and the closed-form FOCs of Section 4.3 apply with quantitative accuracy. The W_2-vs-JKO contrast is itself a long-horizon phenomenon: it sharpens with T as the displacement penalty grows linearly in T under JKO and quadratically under W_2.

The implementability detail is therefore secondary to the within-paper analysis: the framework's predictions are robust to yearly-versus-continuous assessment and hold quantitatively in the long-horizon regime that the canonical calibration targets. The remaining implementability question — the bracket structure — is sufficiently substantive to constitute a separate paper.

8.3 Connection to the Wasserstein companion

Frøseth (2026h) (henceforth the Wasserstein companion) establishes that the W_2-optimal schedule on the full admissible class $\mathcal{S} \supset \mathcal{N}$ is the continuously progressive $\tau^\star(x) = \lambda\, x^2$, with rate $r(x) = \lambda x$ rising linearly in wealth. That schedule lies outside the (C1)–(C3) neutrality class: it violates (C3) by construction (the assessment is wealth-dependent), and it sacrifices neutrality for distributional compression in the sense made precise in Section 6.1. Within the restricted class (C1)–(C3) of the present paper, the W_2-optimal schedule is the corner-pinning at $(k_{\text{lo}}, 0)$ documented in Theorem 3 and Remark 5.2: that is the *closest (C1)–(C3)-feasible point* to the unconstrained quadratic $\tau^\star$, modulo the projection onto the proportional sub-class.

The two papers therefore fill complementary cells of a 2×2 grid: criterion (JKO or W_2) crossed with schedule class (restricted (C1)–(C3) or full $\mathcal{S}$). The Wasserstein companion fills the W_2-on-$\mathcal{S}$ cell; the present paper fills both JKO-on-(C1)–(C3) and W_2-on-(C1)–(C3). The JKO-on-$\mathcal{S}$ cell — the JKO-optimal schedule on the unrestricted class — is the natural fourth cell and is flagged as an open question in Section 8.6.

The clean reading of the joint structure is: Frøseth (2026h) establishes that the W_2 criterion, given full schedule freedom, picks a non-neutral progressive schedule. The present paper establishes that within the neutrality-preserving class, the JKO and W_2 criteria still disagree — now along the proportional-mix dimension rather than along the neutrality dimension — and that the disagreement is regime-dependent through ρ. The two findings are not in tension: they describe the criteria's preferences along different axes of the schedule space. The compan-

ion paper's result is the answer to "what schedule does W_2 pick when allowed?"; the present paper's result is the answer to "what schedule does each criterion pick when restricted to neutrality?". Both questions are well-posed, and the relationship between the two is mediated by the neutrality-vs-distributional-compression normative split foregrounded in Section 6.2.

8.4 Heterogeneous returns and other extensions

The framework treats wealth as a single GBM with homogeneous parameters (μ, σ). Two extensions are immediate.

Multi-asset / heterogeneous-returns extension. A realistic Norwegian household holds a portfolio of asset classes with distinct return distributions: equities, real estate, bonds, deposits, business assets. The single-GBM idealisation aggregates these into an effective $(\mu_{\text{eff}}, \sigma_{\text{eff}})$ along the lines of Section 7.2. A multi-asset extension would replace the single X_t with a vector $\boldsymbol{X}_t = (X_{1,t}, \ldots, X_{n,t})$ of per-class holdings, each evolving as its own GBM, and the wealth-tax schedule would become asset-class-specific: τ_i for class i. The (C3) condition $\alpha_i = \alpha$ corresponds to uniform assessment across asset classes; its violation in actual Norwegian policy (the 25% valuation rule for primary residences, the discounted assessment of unlisted shares) are exactly per-class deviations from a common α. A forthcoming heterogeneous-returns companion to the present paper develops the JKO/W_2 optimisation on the multi-asset state space and characterises the optimal asset-class-specific assessment rates. The single-asset result of the present paper is the homogeneous-returns specialisation of that companion analysis.

The bluntness mechanism of Section 5.5 is sharpest in the heterogeneous-returns setting. The bluntness index $B(m_0) = b/(am_0)$ depends on the geometric mean return m_0, which varies across asset classes (and across cohorts of households with distinct asset compositions). A wealth tax that applies uniformly across classes is therefore proportionally more punitive on low-m_0 classes than on high-m_0 classes: the bluntness penalty is regressive in the cross-section of returns. This is exactly the asymmetry documented in Frøseth (2026d), and it is the principal motivation for asset-class-specific assessment in actual policy.

Spectral-side connection. Frøseth (2026i) develops a spectral portfolio-theory framework that decomposes the wealth process by eigenmodes of the covariance structure. Under (C1)–(C3) the spectral structure is preserved up to the drift-shift-and-rescale, so the present paper's results lift to the spectral framework without modification. Once the (C1)–(C3) restriction is dropped, the asset-class-specific assessment rates of the heterogeneous-returns companion map naturally onto the spectral basis, with each eigenmode receiving its own optimal α_i. Whether the JKO and W_2 criteria agree on the spectral-side optimum is a third open question that the spectral / heterogeneous-returns extension will answer.

8.5 Phase-diagram structure as a research direction

Section 5.3 formulated the JKO crossover as a phase transition with order parameter $\phi^\star$, control parameter ρ, and free energy $\Delta F^\star$. The figures collapse the picture onto a one-dimensional slice

through the GBM parameter space at fixed horizon T and fixed initial spread v_0. The full phase diagram lives in higher dimensions, and several axes invite exploration.

The most immediate generalisation is the $(\rho, R^\star)$ *phase diagram*. The two boundaries $\rho_{\text{low}}(R^\star)$ and $\rho_{\text{high}}(R^\star)$ depend on the revenue target $R^\star$, and the mixed window $[\rho_{\text{low}}(R^\star), \rho_{\text{high}}(R^\star)]$ widens or narrows as $R^\star$ varies. At small enough $R^\star$ the flow-tax base $b\,(1-k)$ alone can deliver the target across the full feasibility range and the mixed phase pinches to a point — a candidate *tricritical-style* structure where the two phase boundaries collide. The geometry of that collision and the behaviour of $\Delta F^\star$ near it are open analytical questions.

A second generalisation extends the dimensionality of the regime parameter. Treating T and v_0 as additional axes opens a four-dimensional GBM–policy phase manifold $(\sigma, T, v_0, R^\star)$. The phase boundaries become hypersurfaces, and the mixed-phase volume becomes a quantitative measure of how often JKO recommends a strictly mixed policy. A scan of the canonical Norwegian-flavoured neighbourhood would tell us how robust Norway's location in the mixed phase is to perturbations of the underlying parameters — a natural sensitivity test for the political reading.

A third direction is the W_2-vs-JKO contrast made phase- diagrammatic. In the calibration of Figure 1 W_2 is phase-poor: it pins to the pure flow-tax corner for every ρ. Whether W_2 acquires phase structure on a different slice of the parameter space, and whether there is an intermediate criterion that interpolates between the JKO and W_2 free-energy weightings, would let us trace how phase-richness emerges as a function of how the criterion weights mean-distortion against spread-distortion. The answer would sharpen the political reading of the criterion choice.

What does *not* transfer from the statistical-mechanics analogy is, as noted in Section 5.3, divergent susceptibility, critical exponents in the universality-class sense, and hysteresis. The kinks in $\phi^\star(\rho)$ are constraint-set crossings rather than spontaneous-symmetry-breaking events, and the relevant universality class is the unremarkable one of polytope-constrained convex optimisation. The point is that the phase-transition language is genuinely informative — it organises the mixed-versus-corner regimes with the right vocabulary — without claiming a depth of physical analogy that the underlying mathematics does not justify.

8.6 Open questions

The 2-bracket extension and the companion paper. The principal limitation of the present analysis is its restriction to proportional wealth taxes. Real wealth-tax schedules in the jurisdictions where the instrument is operational — Norway, Switzerland, Spain, France's old ISF — are piecewise-linear in wealth: an exemption threshold X_0 and one or two non-zero marginal rates above. The simplest non-trivial case, the *2-bracket sub-class* $\{\tau_{X_0,r}(x) = r\,(x - X_0)_+ : X_0 \geq 0,\, r \geq 0\}$, has three free parameters (k, X_0, r) given the flow-tax retention k, and is the natural next layer of the schedule hierarchy

$$\text{(C1)–(C3) (proportional, 1-bracket)} \;\subset\; \text{2-bracket} \;\subset\; \cdots \;\subset\; \mathcal{S} \quad \text{[the full class of Frøseth (2026h)]}.$$

The companion paper has two complementary strands.

Semi-analytical strand. Within each bracket, the FP equation has constant-coefficient drift on log-wealth, the post-tax density is locally Gaussian with a drift correction, and the bracket boundary is a matching condition (continuity of density and probability flux) that admits a closed-form solution in terms of the error function. ΔF and W_2^2 on the 2-bracket sub-class therefore remain analytic — in elementary functions plus erf — and the matched-revenue Lagrangian becomes a 3-parameter optimisation with explicit (transcendental) FOCs. The phase-transition mechanism of Section 5.5 extends in spirit: the constant- marginal-cost-in-τ_w structure of JKO generalises to a piecewise-explicit marginal cost in r with X_0-dependent coefficients, and the boundaries of the JKO phase diagram in the new $(\rho, R^\star, X_0)$ parameter space can be characterised analytically near the (C1)–(C3) limit $X_0 \to 0$.

Numerical strand. The semi-analytical approach is clean on the 2-bracket sub-class but does not extend straightforwardly to richer schedules (n-bracket for $n \geq 3$, smoothly progressive shapes, or schedules with non-trivial boundary behaviour) where the post-tax density loses its piecewise-Gaussian structure. A direct numerical attack — solving the FP equation by finite-difference time-stepping or Monte Carlo simulation of the post-tax SDE, evaluating ΔF and W_2^2 from the resulting density, and optimising over the schedule parameters — handles the entire schedule space and is the right tool for mapping the full phase diagram across $(\mu, \sigma, T, v_0, R^\star, X_0, r, \ldots)$ at once. The numerical strand also serves as verification of the semi-analytical closed forms where they apply, and as the only available tool in regimes where they do not.

The natural framing is that of a two-paper series. The present paper establishes the (C1)–(C3) phase structure on the proportional sub-class, pinning down the regime-dependence and the JKO-vs-W_2 contrast in the cleanest setting. The companion paper combines the two strands above to deliver (i) closed-form JKO-optimal $(X_0^\star, r^\star)$ at the leading order in a small-threshold expansion from $X_0 = 0$, (ii) numerical solutions across the full 2-bracket parameter space and beyond, mapping the JKO phase diagram and validating the analytical results, (iii) the matching empirical comparison to the actual Norwegian and Swiss bracket schedules, which the present paper deliberately defers. The structural relationship between the two papers parallels the neutrality / extensions split between Frøseth (2026e) and Frøseth (2026c).

Other open questions. Two further analytical questions remain, treated only briefly here: sweeps over (μ, v_0) at fixed σ to trace the full four-dimensional crossover surface (this paper traces only the σ-slice and the T-slice); and the unconstrained mathematical optimum allowing schedule-side negative coefficients as a curiosity case.

8.7 Horizon dependence of the JKO recommendation

The phase structure of Theorem 4 is established at a fixed planning horizon T, and the calibration of Section 5.4 fixes $T = 5$ years as the canonical Norwegian-baseline value. The structural question we address here is what happens to the phase boundaries $\rho_{\text{low}}(T), \rho_{\text{high}}(T)$ as the horizon T is varied, and where Norway's actual planning horizon falls in that picture. The

answer connects directly to the bluntness mechanism of Section 5.5 via the natural scale crossover $T_c = v_0/\sigma^2$ introduced at the end of that section.

Two scaling regimes of ρ. The crossover ratio $\rho = \Sigma_0\, m_0/\sigma^2$ has two distinct T-scalings, separated by the time at which the cumulative diffusion $\sigma^2 T$ matches the initial log-wealth variance v_0:

$$\rho \;\approx\; \begin{cases} \sqrt{v_0}\,\dfrac{m_0}{\sigma^2} & \text{for } T \ll T_c \quad \text{(initial-spread-dominated)}, \\[2ex] \dfrac{m_0\sqrt{T}}{\sigma} & \text{for } T \gg T_c \quad \text{(diffusion-dominated)}. \end{cases} \tag{50}$$

For Norwegian-baseline parameters ($v_0 = 0.25$, $\sigma = 0.30$), $T_c \approx 2.78$ years; the canonical $T = 5$ sits just past the crossover, in the diffusion-dominated regime where ρ grows like $\sqrt{T}$.

Both phase boundaries pinch at $T = 0$ and $T \to \infty$. Substituting the unconstrained closed-form $k^\star_{\mathrm{JKO}}(\sigma) = K$ into the matched-revenue condition gives, after simplification,

$$\sigma^4 \;=\; 2\,m_0\,(b/a - m_0)\,(K\,\sigma^2 T + v_0). \tag{51}$$

Two limits:

Short-horizon limit ($T \to 0$). The $K\,\sigma^2 T$ term vanishes and (51) reduces to $\sigma^4 = 2\,m_0\,(b/a - m_0)\,v_0$ — *independent of* K. The two phase boundaries (which differ only in their target $K \in \{1, k_{\mathrm{lo}}\}$) collapse onto the same σ, and the mixed-phase σ-window vanishes.

Long-horizon limit ($T \to \infty$). The $K\,\sigma^2 T$ term dominates and $\sigma^2 \to 2\,m_0\,(b/a - m_0)\,K\,T$. To keep σ^2 bounded by the existence condition $\sigma^2 < 2\mu$, m_0 must shrink like $1/T$, which forces $\sigma \to \sqrt{2\mu}$. Both boundaries converge to the same limit. The ρ-width decays as $\sim 1/\sqrt{T}$ but never quite reaches zero in finite T.

The two pinches have different content. The $T \to 0$ pinch is "no time for the JKO criterion to discriminate between the two channels". The $T \to \infty$ pinch is "the geometric mean log-return m_0 vanishes, both channels look equivalent at the margin". Between the two, the mixed-phase window opens, peaks in width at $T^\star \approx 11$ years, then closes again.

Norway's planning horizon flips the recommendation. Norway's $\sigma = 0.30$ trajectory through the (T, ρ) plane is shown in Figure 3. Because $\rho_{\mathrm{Norway}}(T)$ grows monotonically from $\sqrt{v_0}\,m_0/\sigma^2 \approx 0.14$ at $T \to 0$ toward the diffusion-dominated branch as T grows, while the phase boundaries $\rho_{\mathrm{low}}(T), \rho_{\mathrm{high}}(T)$ shift in the opposite direction, the trajectory crosses both boundaries. Numerically:

T (years)	ρ_{Norway}	JKO recommendation
1	0.16	Pure wealth-tax phase
3	0.20	Pure wealth-tax phase
5 *(canonical)*	**0.23**	**Mixed-instrument phase**
10	0.30	Pure flow-tax phase
20	0.40	Pure flow-tax phase

The mixed-instrument recommendation that the present paper foregrounds at the canonical Norwegian-baseline regime is *specific to the five-year planning horizon.* Shorter horizons (one-year tax-planning windows, year-to-year revenue adjustments) push the JKO optimum into the pure wealth-tax phase; longer horizons (retirement, intergenerational planning) push it into the pure flow-tax phase. The qualitative recommendation flips twice as T varies.

Economic content of T_c. The scale crossover $T_c = v_0/\sigma^2$ has a direct interpretation: it is the time at which the wealth distribution stops "remembering" its initial state and starts being shaped predominantly by the post-tax dynamics. Below T_c, the cost structure under JKO is dominated by initial-condition terms and does not discriminate between the two channels; above T_c, the channels' time-scaling diverge and the bluntness penalty's T-amplification under the wealth-tax channel makes that channel progressively more expensive.

The horizon-dependent flip is therefore not a quirk of the model but a reflection of which time-scale the policy debate is implicitly using. Year-to-year revenue planning lives in a fundamentally different regime from lifecycle wealth-tax design, and the JKO criterion — which prices distortion linearly in the displacement and accumulates that displacement linearly in T — registers the difference. A reader uncomfortable with the canonical mixed-phase recommendation has, in this picture, a principled out: choose a different planning horizon, and the JKO criterion will recommend a different schedule. A reader committed to the mixed-phase recommendation has, equally, a principled defence: only at the five-year horizon does the criterion give a strictly mixed answer, and that horizon corresponds to a natural equity-portfolio planning window.

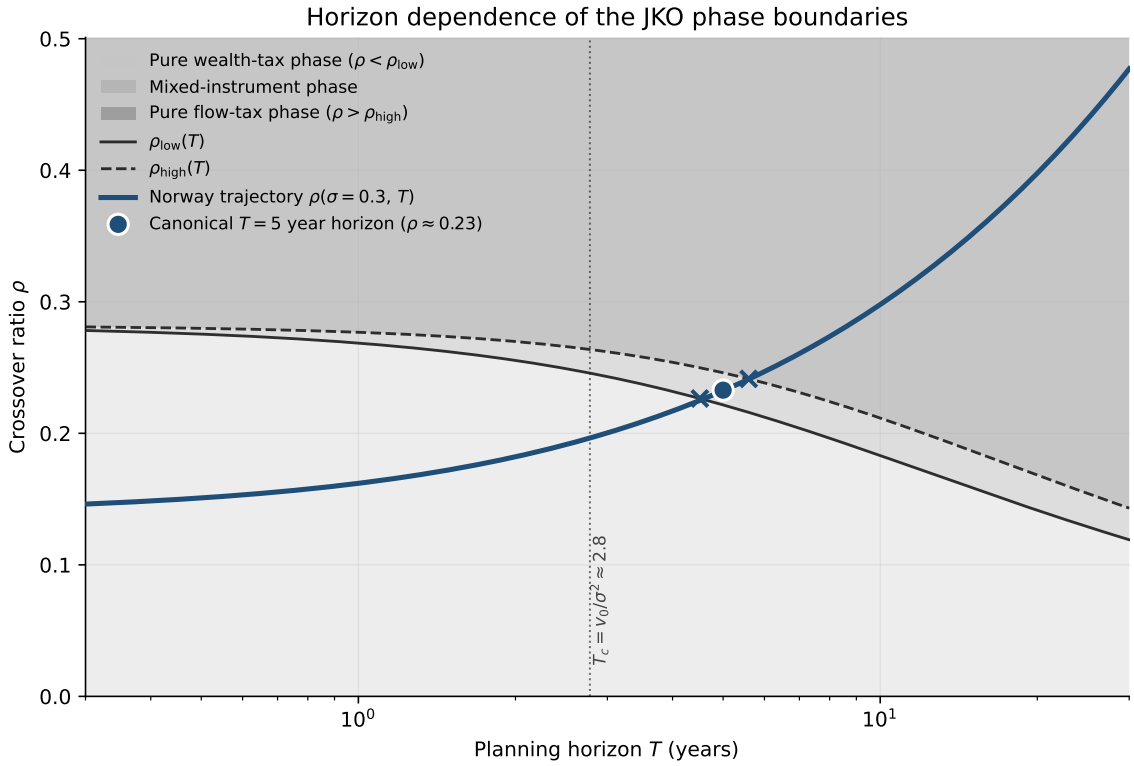

Figure 3: **Horizon dependence of the JKO phase boundaries with the Norway-σ trajectory overlaid.** The phase boundaries $\rho_{\text{low}}(T)$ (solid grey) and $\rho_{\text{high}}(T)$ (dashed grey) define the boundaries of the mixed-instrument phase as functions of the planning horizon T (log-scaled). The three shaded regions are the JKO phases of Theorem 4: gold (pure wealth-tax, $\rho < \rho_{\text{low}}$), lavender (mixed instrument, $\rho_{\text{low}} \leq \rho \leq \rho_{\text{high}}$), cyan (pure flow-tax, $\rho > \rho_{\text{high}}$). The teal curve is $\rho(\sigma = 0.30, T)$, the Norway-σ trajectory through the (T, ρ) plane. The two crosses mark where the trajectory enters and leaves the mixed-instrument phase. The canonical Norwegian-baseline $T = 5$ years (teal disc) sits inside the mixed phase by a narrow margin; shorter horizons fall in the pure wealth-tax phase, longer horizons in the pure flow-tax phase. The dotted vertical rule at $T_c = v_0/\sigma^2 \approx 2.78$ marks the scale crossover from the initial-spread-dominated short-horizon regime to the diffusion-dominated long-horizon regime. Calibration: $\mu = 0.07$, $\sigma = 0.30$, $v_0 = 0.25$, $a = 1.0$, $b = 0.27$, $R^\star = 0.05$, matched to Figure 1.

Acknowledgements

The author acknowledges the use of Claude (Anthropic) for assistance with literature review, LaTeX typesetting, mathematical exposition, and editorial refinement, and Lemma (Axiomatic AI) for review and proof checking. All substantive arguments, economic reasoning, and conclusions are the author's own.

References

Rolf Aaberge, Jørgen Modalsli, and Edda Solbakken. Measuring long-run wealth inequality: Empirical results for Norway 1912–2019. Discussion Papers 1028, Statistics Norway, October 2025.

Thomas Blanchet, Thomas Piketty, and Juliette Fournier. Generalized Pareto curves: Theory and applications. *Review of Income and Wealth*, 68(1):263–287, 2022. doi: 10.1111/roiw.12510.

Peter Diamond and Emmanuel Saez. The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives*, 25(4):165–190, 2011. doi: 10.1257/jep.25.4.165.

Anders G. Frøseth. Tax migration as social contagion: A tipping-point model with application to the Scandinavian wealth tax debate. 2026a. arXiv:2607.03868 [physics.soc-ph].

Anders G. Frøseth. Extensions to the wealth tax neutrality framework. 2026b. arXiv:2603.05277 [physics.soc-ph].

Anders G. Frøseth. Flow taxes, stock taxes, and portfolio choice: A generalised neutrality result. 2026c. arXiv:2603.15974 [physics.soc-ph].

Anders G. Frøseth. Heterogeneous returns and wealth tax neutrality: A Fokker–Planck framework. 2026d. arXiv:2603.16006 [physics.soc-ph].

Anders G. Frøseth. Asset returns, portfolio choice, and proportional wealth taxation. 2026e. arXiv:2603.05264 [physics.soc-ph].

Anders G. Frøseth. From gravity to confinement: Wealth redistribution as optimal drift design in the Fokker–Planck framework. 2026f. arXiv:2607.06153 [physics.soc-ph].

Anders G. Frøseth. Wealth taxation as a drift modification: A Fokker–Planck approach to tax neutrality. 2026g. arXiv:2603.05283 [physics.soc-ph].

Anders G. Frøseth. Wasserstein geometry of wealth taxation: Rigid motions of the log-wealth distribution and the optimality of progressive wealth taxes. Working paper, 2026h.

Anders G. Frøseth. Spectral portfolio theory: From SGD weight matrices to wealth dynamics. 2026i. arXiv:2603.09006 [physics.soc-ph].

Katrine Jakobsen, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman. Wealth taxation

and wealth accumulation: Theory and evidence from Denmark. *The Quarterly Journal of Economics*, 135(1):329–388, 2020. doi: 10.1093/qje/qjz032.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998. doi: 10.1137/S0036141096303359.

J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208, 1971. doi: 10.2307/2296779. Stable URL: https://www.jstor.org/stable/2296779.

James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. *Tax by Design: The Mirrlees Review*. Institute for Fiscal Studies, London, 2011.

Thomas Piketty. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, MA, 2014.

Thomas Piketty and Gabriel Zucman. Capital is back: Wealth-income ratios in rich countries 1700–2010. *Quarterly Journal of Economics*, 129(3):1255–1310, 2014.

Emmanuel Saez and Stefanie Stantcheva. A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120–142, 2018. doi: 10.1016/j.jpubeco.2017.10.004.

Emmanuel Saez and Gabriel Zucman. Progressive wealth taxation. *Brookings Papers on Economic Activity*, 2019(Fall):437–533, 2019. doi: 10.1353/eca.2019.0017.

Filippo Santambrogio. {Euclidean, Metric, and Wasserstein} Gradient Flows: an overview. *Bulletin of Mathematical Sciences*, 7:87–154, 2017. doi: 10.1007/s13373-017-0101-1. arXiv preprint 13 September 2016, arXiv:1609.03890.

Cédric Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009. doi: 10.1007/978-3-540-71050-9.