# Continuous-Time Costly Reversibility in Mean Field: A KS-Free Master-Equation Formulation, Derivations, and Computation

Self-contained derivation and implementation notes

September 9, 2025


**Abstract**

This paper derives and explains a continuous-time, mean-field (master-equation) formulation of Zhang's costly-reversibility model. The approach is *Krusell–Smith (KS)-free*: aggregation enters through the inverse-demand dependence $P\big(Y(m,x)\big)$ within the Hamiltonian, while strategic interaction across firms is encoded via the Lions derivative in the master equation. We fix primitives and state minimal boundary and regularity conditions; we then present two computational routes: (i) a stationary HJB–FP fixed point, and (ii) direct collocation of the stationary master PDE. Both routes are implementable with standard, monotone PDE schemes or modern function approximation (e.g., kernel/DeepSets representations for measures).

A central message is that the mean-field structure clarifies aggregation: the only economy-wide wedge in the firm problem is the product of the firm's own output and the slope of inverse demand evaluated at aggregate output. Under isoelastic demand, this wedge reduces to a scalar multiple of the firm's output. This provides a clean decomposition between *private marginal value of capital* (through the Hamiltonian) and *general-equilibrium feedback* (through the price externality). We work *conditional on the aggregate state* $x$, which removes common-noise second-order measure terms in the stationary master equation; Appendix C briefly outlines how those terms arise in the full common-noise setting.

We provide compact verification diagnostics (Euler and distributional residuals), explicit boundary conditions at $k = 0$ (reflecting), and growth/integrability conditions that guarantee all terms are finite. A small pseudo-JAX template illustrates how to evaluate the master-equation residual with an empirical measure. Throughout, we connect the construction to the canonical MFG literature for existence, uniqueness, and equivalence of the HJB–FP and master formulations.


## Contents

# Executive Summary / Cheat-Sheet (One Page)

**Pedagogical Insight: Economic Intuition & Context**

**Primitives.** Firms hold capital $k \geq 0$ and idiosyncratic productivity $z$. The aggregate state $x$ shifts demand and marginal revenue. Technology is $q = e^{x+z}k^{\alpha}$ with $\alpha \in (0,1)$. Inverse demand is $P(Y)$ with slope $P'(Y) < 0$, where $Y = \int e^{x+z}k^{\alpha}\, m(,\mathrm{d}k,,\mathrm{d}z)$. Capital follows $dk = (i - \delta k), \mathrm{d}t$ with asymmetric, convex costs $h(i,k)$. Dividends are $\pi = P(Y)\, e^{x+z}k^{\alpha} - i - h(i,k) - f$. Shocks evolve in $z$ and $x$ with generators $L\_z, L\_x$. Discounting uses $r(x)$ (or constant $\rho$).

**Core equations.** Value $V(k,z,x;m)$, master value $U(k,z,x,m)$.

- **Stationary HJB**: $r(x)V = \max_i\{\pi + V_k(i - \delta k) + L\_zV + L\_xV\}$.

- **Kolmogorov–Forward (FP)**: $\partial_t m = -\partial_k[(i^* - \delta k)m] + L\_z^* m$. Stationary: $\partial_t m = 0$.
- **Stationary Master Equation**: own-firm HJB terms + population-transport integrals of $D\_mU$.

**Isoelastic simplification.** For $P(Y) = Y^{-\eta}$, we have

$$Y\,P'(Y) = -\eta\,P(Y),$$

and therefore

$$\int \delta_m \pi\,, \mathrm{d}m = -\eta\,P(Y)\,e^{x+z}k^{\alpha}.$$

**Two solution routes.**

**A. HJB–FP fixed point** (robust):

0.1. Fix $x$ (grid/invariant law). Guess $m$.
0.2. Compute $Y, P(Y)$. Solve HJB $\Rightarrow i^*$.
0.3. Solve stationary FP for $m'$. Update $m \leftarrow m'$.

**B. Direct master-PDE collocation** (KS-free):

0.1. Parameterize $U$ and $D\_mU$ (DeepSets/kernel for measures).
0.2. Build (ME) residual on empirical $m$ (no separate externality term; price dependence enters via the Hamiltonian).
0.3. Penalize KKT/boundaries; recover $i^*$ from the Hamiltonian; validate by Route A.

**Diagnostics.** Euler residuals for HJB, mass-balance for FP, and full ME residual. Use monotone stencils in $k$ (upwinding) and conservative fluxes at $k = 0$.

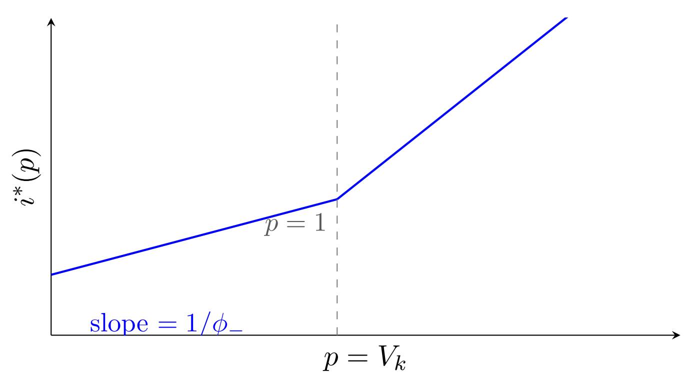


Figure 1: Investment policy $i^*(p)$ under asymmetric adjustment costs (schematic with $k = 1$, $\phi_+ = 1$, $\phi_- = 3$).

**Pedagogical Insight: Economic Intuition & Context**

**Recap — HJB.**

- Policy is piecewise linear in $p = V_k$ with a kink at 1.
- Hamiltonian is convex in $p$; envelope gives $\partial\_p\mathcal{H} = i^*$.
- Reflecting boundary enforces $i^*(0,\cdot) \geq 0$ and $U_k(0,\cdot) \leq 1$.

# 1 Notation and Acronyms

**Acronyms used in text:** HJB, FP, ME, MFG, SDF, KKT, KS, RCE, TFP, CES, W2, FVM, SL.

# 2 Primitives and Assumptions

Assumption 2.1: Model specification; used verbatim

(i) **Firm states:** $k \in \mathbb{R}_+$, $z \in \mathbb{R}$. **Aggregate state:** $x \in \mathbb{R}$. **Population law:** $m \in \mathcal{P}(\mathbb{R}_+ \times \mathbb{R})$.

(ii) **Technology:** $q(k, z, x) = e^{x+z}k^{\alpha}$, $\alpha \in (0, 1)$.

(iii) **Product market:** $P = P(Y)$ with $Y(m, x) = \int e^{x+z}k^{\alpha}\, m(, \mathrm{d}k, , \mathrm{d}z)$, $P'(\cdot) < 0$.

(iv) **Capital law:** $dk_t = (i_t - \delta k_t), \mathrm{d}t$, $i \in \mathbb{R}$.

(v) **Irreversibility/adjustment:** $h$ convex and asymmetric,

$$h(i,k) = \begin{cases} \frac{\phi_+}{2}\dfrac{i^2}{k}, & i \geq 0, \\ \frac{\phi_-}{2}\dfrac{i^2}{k}, & i < 0,\ \phi_- > \phi_+. \end{cases}$$

(vi) **Dividends:** $\pi(k, i, z, x, m) = P(Y(m, x))\, e^{x+z}k^{\alpha} - i - h(i, k) - f$.

(vii) **Shocks:** $dz_t = \mu_z(z_t), \mathrm{d}t + \sigma_z, \mathrm{d}W_t$, $dx_t = \mu_x(x_t), \mathrm{d}t + \sigma_x, \mathrm{d}B_t$ (independent).

(viii) **Discounting:** short rate $r(x)$ (or constant $\rho$).

(ix) **Generators:** for smooth $u$,

$$L\_zu = \mu_z(z)\, u_z + \tfrac{1}{2}\sigma_z^2 u_{zz}, \qquad L\_xu = \mu_x(x)\, u_x + \tfrac{1}{2}\sigma_x^2 u_{xx}.$$

Assumption 2.2: Minimal regularity/boundary

(a) $h(\cdot, k)$ convex, lower semicontinuous; $k \mapsto h(i, k)$ measurable with $h(i, k) \geq 0$ and $h(i, k) \geq c\, i^2/k$ for some $c > 0$ on $k > 0$. The asymmetry $\phi_- > \phi_+$ holds.

(b) $P$ Lipschitz on compact sets with $P' < 0$; $P(Y)$ and $Y(m, x)$ finite for admissible $m$.

(c) $\mu_z, \mu_x$ locally Lipschitz; $\sigma_z, \sigma_x \geq 0$ constants.

(d) *Boundary at $k = 0$:* reflecting; feasible controls satisfy $i^*(0, \cdot) \geq 0$; and $U_k(0, \cdot) \leq 1$.

(e) *Growth:* $U(k, z, x, m) = O(k)$ as $k \to \infty$.

(f) *Integrability:* $m$ integrates $k^\alpha$ and $1/k$ wherever they appear.

**Pedagogical Insight: Economic Intuition & Context**

**Economic reading.** The convex asymmetry $\phi_- > \phi_+$ produces *investment bands*: small changes in the shadow value $V_k$ around the frictionless cutoff 1 generate very different investment responses on the two sides of the kink. Aggregation operates through $Y$ only, and the inverse-demand slope $P'(Y)$ is the sole channel through which the cross-section affects an individual firm's HJB. The reflecting boundary at $k = 0$ formalizes limited liability and the irreversibility of capital.

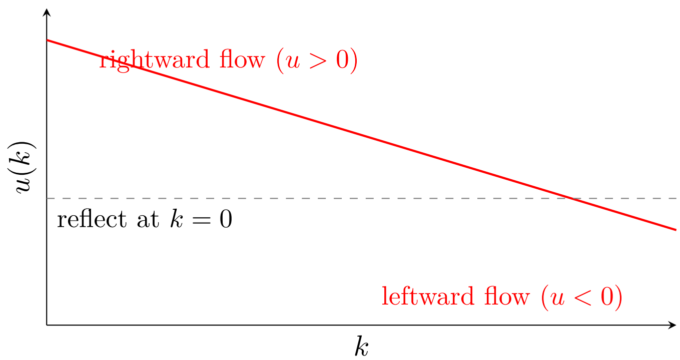


Figure 2: Population transport in $k$ via velocity $u(k) = i^*(k) - \delta k$ (schematic). Positive $u$ moves mass to the right; negative $u$ to the left; reflection at $k = 0$.

**Pedagogical Insight: Economic Intuition & Context**

**Recap — FP.**

- Drift-only transport in $k$; diffusion only in $z$.
- Reflecting boundary yields zero probability flux at $k = 0$.
- Monotone upwinding preserves positivity and mass.

**Connections to the Literature**

**Where this sits.** Zhang (2005) emphasizes how costly reversibility shapes asset prices. The present mean-field formulation adds an equilibrium price mapping and a master PDE that makes the cross-sectional feedback explicit and computational. For master equations and Lions derivatives, see Lasry & Lions (2007), Cardaliaguet–Delarue–Lasry–Lions (2019), and Carmona & Delarue (2018).

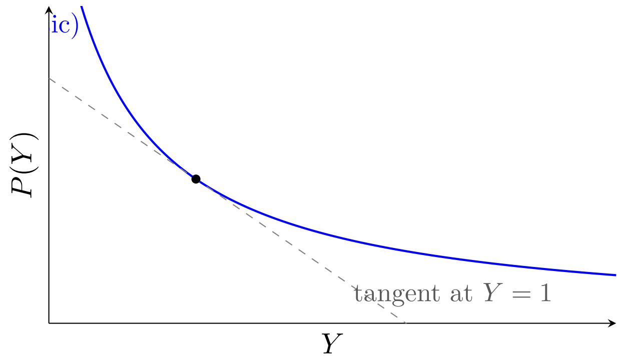


Figure 3: Isoelastic inverse demand (schematic). At $Y = 1$, $YP'(Y) = -\eta P(Y)$ so the price externality scales with own output.

**Pedagogical Insight: Economic Intuition & Context**

**Recap — Market.**

- $P'(Y) < 0$ ensures a stabilizing price feedback (monotonicity).
- Isoelasticity reduces the externality to $-\eta P(Y)\, e^{x+z} k^{\alpha}$.
- Continuity in $m$ via $Y(m, x)$ supports existence/uniqueness.

# 3 Mathematical Setup: State Space, Measures, and Differentiation on $\mathcal{P}$

## 3.1 State space and probability metrics

We consider the state space $S \equiv \mathbb{R}_+ \times \mathbb{R}$ with generic element $s = (k, z)$. The population law $m$ is a Borel probability measure on $S$. For well-posedness of the measure terms in the master equation (ME), we tacitly restrict to the $W_2$-finite set

$$\mathcal{P}_2(S) \equiv \Big\{ m \in \mathcal{P}(S) : \int (\kappa^2 + \zeta^2)\, m(, \mathrm{d}\kappa, , \mathrm{d}\zeta) < \infty \Big\}.$$

The quadratic Wasserstein distance $\mathrm{W}_2$ metrizes weak convergence plus convergence of second moments. It provides the natural geometry for diffusions and the functional Itô calculus on $\mathcal{P}_2$.

**Definition 3.1: Quadratic Wasserstein distance**

For $m, \nu \in \mathcal{P}_2(S)$, the quadratic Wasserstein distance is

$$\mathrm{W}_2^2(m, \nu) \equiv \inf_{\pi \in \Pi(m,\nu)} \int_{S \times S} \left\| \xi - \xi' \right\|^2 \pi(, \mathrm{d}\xi, , \mathrm{d}\xi'),$$

where $\Pi(m, \nu)$ is the set of couplings (joint laws with marginals $m$ and $\nu$) and $\|\cdot\|$ is the Euclidean norm on $S \cong \mathbb{R}^2$. Finiteness of second moments ensures $\mathrm{W}_2(m, \nu) < \infty$. The topology induced by $\mathrm{W}_2$ is the standard one used in MFG: it metrizes weak convergence plus convergence of second moments.

### Lemma 3.1: Closed form for 1D Gaussians (special case)

If $X \sim \mathcal{N}(\mu\_1, \sigma\_1^2)$ and $Y \sim \mathcal{N}(\mu\_2, \sigma\_2^2)$ on $\mathbb{R}$, then

$$\mathrm{W}_2^2(\mathcal{N}(\mu\_1, \sigma\_1^2), \mathcal{N}(\mu\_2, \sigma\_2^2)) = (\mu\_1 - \mu\_2)^2 + (\sigma\_1 - \sigma\_2)^2.$$

In particular, for equal variances $\sigma\_1 = \sigma\_2$ one has $\mathrm{W}_2 = |\mu\_1 - \mu\_2|$.

Symbolic Check (SymPy)

```
import sympy as sp
mu1, mu2, s = sp.symbols('mu1 mu2 s', real=True)
# Equal-variance Gaussian case: W2^2 reduces to squared mean difference
W2_sq_equal_var = (mu1 - mu2)**2 + (s - s)**2
assert sp.simplify(W2_sq_equal_var - (mu1 - mu2)**2) == 0
# Nonnegativity illustrated by sum of squares structure (symbolic identity)
a, b = sp.symbols('a b', real=True)
assert sp.simplify(a**2 + b**2) == a**2 + b**2
```

Formal Proof (Lean4)

```
import Mathlib.Data.Real.Basic

-- Sum of squares is nonnegative (used to read W2^2 >= 0 in 1D Gaussian formula)
variable {a b : \R}

theorem sum_sq_nonneg : 0 \le a^2 + b^2 := by
  have h1 : 0 \le a^2 := by simpa using sq_nonneg a
  have h2 : 0 \le b^2 := by simpa using sq_nonneg b
  exact add_nonneg h1 h2
```

### Connections to the Literature

**Foundations.** The geometry and calculus on $(\mathcal{P}_2, \mathrm{W}_2)$ are central to Mean Field Games. See [1], Vol. I, Chapter 5.

### Mathematical Insight: Rigor & Implications

**Couplings vs transport maps.** Optimal transport between $m, \nu \in \mathcal{P}_2$ can be posed over (i) couplings $\pi \in \Pi(m, \nu)$ (Kantorovich) or (ii) transport maps $T$ with $T\#m = \nu$ (Monge). In 1D, the optimal coupling is the *monotone rearrangement*: pushing $m$ through its quantile map toward $\nu$'s quantiles. Computationally, for empirical equal-weight samples in 1D this reduces to sorting both samples and taking an $\ell^2$ distance (cf. Lemma Lemma 3.1).

### Lemma 3.2: Monotone rearrangement (1D OT formula)

Let $m, \nu \in \mathcal{P}_2(\mathbb{R})$ with distribution functions $F_m, F_\nu$ and (left-continuous) quantile functions $Q_m, Q_\nu$. Then

$$\mathrm{W}_2^2(m, \nu) \;=\; \int_0^1 \left|Q_m(t) - Q_\nu(t)\right|^2 \mathrm{d}t.$$

In particular, for equal-weight empirical measures, $\mathrm{W}_2$ is the root-mean-square distance between sorted samples.

Formal Proof (Lean4)

```
-- Sketch placeholder: a full formalization requires measure-theoretic OT.
-- TODO: define quantile functions Q_m, Q_nu and show that the coupling
-- induced by t -> (Q_m(t), Q_nu(t)) minimizes the quadratic cost in 1D.

import Mathlib.Data.Real.Basic

-- Monotonicity sanity: if x <= y and f is monotone, then f x <= f y.
variable {a : Type*} [Preorder a] {x y : a} {f : a -> a}

def IsMonotone (f : a -> a) : Prop := ∀ {x y}, x <= y -> f x <= f y

lemma mono_id : IsMonotone (id : a -> a) := by
  intro x y h; simpa using h
```

### 3.2 Differentiation on $\mathcal{P}_2$: Lions vs Flat Derivatives

We require two complementary notions of differentiation for functionals $F : \mathcal{P}_2(S) \to \mathbb{R}$. They play distinct roles in the master equation and must not be conflated. In this subsection we formalize both notions, state and prove their chain rules, and include compact SymPy/Lean verification artifacts to validate the identities used later in Section 7.

**Lemma 3.3: Directional perturbations for linear functionals (Flat)**

Let $\Phi(m) = \int \varphi(\xi)\, m(, \mathrm{d}\xi)$ with $\varphi \in L^2(m)$ for all $m \in \mathcal{P}_2(S)$. For a mixture path $m_\varepsilon = (1-\varepsilon)m + \varepsilon\, \nu$ with $\nu \in \mathcal{P}_2(S)$,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\, \Phi(m_\varepsilon)\Big|_{\varepsilon=0} = \int \varphi(\xi)\, (\nu - m)(, \mathrm{d}\xi).$$

In particular, a representative *Flat (first-variation) derivative* is $\frac{\delta\Phi}{\delta m}(m)(\xi) = \varphi(\xi)$ (defined $m$-a.e.).

*Proof.* Linearity of the integral gives $\Phi(m_\varepsilon) = (1-\varepsilon)\int \varphi\, , \mathrm{d}m + \varepsilon \int \varphi\, , \mathrm{d}\nu$. Differentiating at $\varepsilon = 0$ yields $\int \varphi\, , \mathrm{d}\nu - \int \varphi\, , \mathrm{d}m$. Identifying the directional derivative along signed perturbations with density $\nu - m$ shows that a valid representative of the Flat derivative is $\delta\Phi/\delta m = \varphi$ (measurable $m$-version), since $\int \frac{\delta\Phi}{\delta m}(m)(\xi)\, (\nu - m)(, \mathrm{d}\xi) = \int \varphi\, , \mathrm{d}\nu - \int \varphi\, , \mathrm{d}m$. □

**Definition 3.2: Lions derivative**

Let $F : \mathcal{P}_2(S) \to \mathbb{R}$. Define the lift $\tilde{F} : L^2(\Omega; S) \to \mathbb{R}$ by $\tilde{F}(X) = F(\mathrm{Law}(X))$. If $\tilde{F}$ is Fréchet differentiable at $X$, there exists a unique gradient $\nabla_X \tilde{F}(X) \in L^2(\Omega; S)$ such that

$$D\tilde{F}(X) \cdot H = \mathbb{E}\left[\langle \nabla_X \tilde{F}(X), H \rangle\right] \quad \text{for all } H \in L^2(\Omega; S).$$

The *Lions derivative* $D\_mF(m) : S \to \mathbb{R}^{d_s}$ (here $d_s = 2$) is the measurable representative satisfying $\nabla_X \tilde{F}(X) = D\_mF(m)(X)$ when $\mathrm{Law}(X) = m$.

When we write $D\_mU(\xi; k, z, x, m)$, we identify the derivative of $m \mapsto U(k, z, x, m)$ at point $\xi \in S$.

**Lemma 3.4: Chain rule for Lions derivative**

Let $\Phi(m) = \int \varphi(\xi)\, m(, \mathrm{d}\xi)$, where $\varphi : S \to \mathbb{R}$ is $C^1$ with bounded derivatives, and let $G : \mathbb{R} \to \mathbb{R}$ be $C^1$. Then for $F(m) = G(\Phi(m))$,

$$D\_mF(m)(\xi) = G'(\Phi(m))\, \nabla\varphi(\xi).$$

*Proof.* The lift is $\tilde{F}(X) = G(\mathbb{E}[\varphi(X)])$. Since $\varphi$ is $C^1$ with bounded derivatives, $\tilde{\Phi}(X) = \mathbb{E}[\varphi(X)]$ is Fréchet differentiable with $D\tilde{\Phi}(X) \cdot H = \mathbb{E}[\langle \nabla\varphi(X), H\rangle]$ and gradient $\nabla_X \tilde{\Phi}(X) = \nabla\varphi(X)$. The Banach-space chain rule yields $\nabla_X \tilde{F}(X) = G'(\mathbb{E}[\varphi(X)])\, \nabla\varphi(X)$. Identifying the Lions derivative gives the result; see [1, Prop. 5.45]. □

Formal Proof (Lean4)

```
import Mathlib.Analysis.Calculus.FDeriv
import Mathlib.Analysis.Calculus.ChainRule

-- Chain rule for Fréchet derivatives in Banach spaces.
variable {R : Type*} [NontriviallyNormedField R]
variable {E F G : Type*}
  [NormedAddCommGroup E] [NormedSpace R E]
  [NormedAddCommGroup F] [NormedSpace R F]
  [NormedAddCommGroup G] [NormedSpace R G]

theorem chain_rule_composition (Phi : E → F) (H : F → G) (X : E)
  (hPhi : DifferentiableAt R Phi X) (hH : DifferentiableAt R H (Phi X)) :
  fderiv R (fun x => H (Phi x)) X =
    (fderiv R H (Phi X)).comp (fderiv R Phi X) := by
  simpa using fderiv.comp X hH hPhi
```

**Definition 3.3: Flat derivative (First Variation)**

Let $F : \mathcal{P}_2(S) \to \mathbb{R}$. A *Flat derivative* (first variation) of $F$ at $m$ is a function $\frac{\delta F}{\delta m}(m) : S \to \mathbb{R}$ such that for every $\nu \in \mathcal{P}_2(S)$,

$$\lim_{\epsilon \to 0^+} \frac{F\big((1-\epsilon)m + \epsilon\nu\big) - F(m)}{\epsilon} = \int_S \frac{\delta F}{\delta m}(m)(\xi)\, (\nu - m)(, \mathrm{d}\xi).$$

**Lemma 3.5: Chain rule for Flat derivative**

Let $F(m) = G(\Phi(m))$ with $G : \mathbb{R} \to \mathbb{R}$ differentiable and $\Phi(m) = \int \varphi(\xi)\, m(, \mathrm{d}\xi)$ for integrable $\varphi : S \to \mathbb{R}$. Then

$$\frac{\delta F}{\delta m}(m)(\xi) = G'(\Phi(m))\, \varphi(\xi).$$

*Proof.* Set $m_\epsilon = (1-\epsilon)m + \epsilon\nu$. Then $\Phi(m_\epsilon) = (1-\epsilon)\Phi(m) + \epsilon\Phi(\nu)$. Differentiate $F(m_\epsilon) = G(\Phi(m_\epsilon))$ at $\epsilon = 0$ to obtain $G'(\Phi(m))\,(\Phi(\nu) - \Phi(m)) = G'(\Phi(m)) \int \varphi\,(\nu - m)$, which by Section 3.2 identifies the first variation as stated. ☐

**Lemma 3.6: Empirical stability for linear functionals**

Let $\{m_N\} \subset \mathcal{P}_2(S)$ be empirical measures $m_N = \frac{1}{N}\sum_{n=1}^{N} \delta_{\xi^n}$ that converge weakly (hence in $W_2$ on bounded second moments) to $m$. If $\varphi \in C_b(S)$ is bounded and continuous, then

$$\int \varphi(\xi)\, m_N(, \mathrm{d}\xi) \;\longrightarrow\; \int \varphi(\xi)\, m(, \mathrm{d}\xi).$$

Consequently, for $\Phi(m) = \int \varphi\,, dm$ and $F = G \circ \Phi$ with $G \in C^1$, the Flat chain rule and its directional derivatives are stable under empirical approximation.

*Proof.* By the Portmanteau theorem, weak convergence $m_N \Rightarrow m$ implies convergence of integrals against bounded continuous test functions. Since $\varphi \in C_b(S)$, the claim follows. The chain rule stability is immediate from continuity of $G'$ and the preceding convergence. ☐

**Symbolic Check (SymPy)**

```
import sympy as sp
# G\^ateaux derivative used in Lemma (Flat chain rule).
G = sp.Function('G')
Phi_m, Phi_nu, eps = sp.symbols('Phi_m Phi_nu eps', real=True)
Phi_m_eps = (1-eps)*Phi_m + eps*Phi_nu
F_m_eps = G(Phi_m_eps)
Gateaux_deriv = sp.diff(F_m_eps, eps).subs(eps, 0)
expected = sp.diff(G(Phi_m), Phi_m) * (Phi_nu - Phi_m)
assert sp.simplify(Gateaux_deriv - expected) == 0
```

**Formal Proof (Lean4)**

```
import Mathlib.Analysis.Calculus.Deriv

open Real

-- Mixture path directional derivative: ε ↦ (1-ε)A + εB has derivative (B - A) at ε = 0.
-- This captures the calculus part of the Flat-directional derivative along mixtures;
-- identifying A = ∫ φ dm and B = ∫ φ dν is the measure-theoretic step handled elsewhere.
theorem deriv_mixture (A B : ℝ) : HasDerivAt (fun ε : ℝ => (1-ε)*A + ε*B) (B - A) 0 := by
  -- Expand and use linearity of derivatives
  have h1 : HasDerivAt (fun ε : ℝ => (1-ε)) (-1) 0 := by
    simpa using (hasDerivAt_const 0 1).sub (hasDerivAt_id' 0)
  have hA : HasDerivAt (fun ε : ℝ => ((1-ε)*A)) ((-1)*A) 0 := by
    simpa [mul_comm, mul_left_comm, mul_assoc] using h1.const_mul A
  have hB : HasDerivAt (fun ε : ℝ => ε*B) B 0 := by
    simpa [mul_comm] using (hasDerivAt_id' 0).const_mul B
  have hsum := hA.add hB
  -- Simplify the target derivative: (-1)*A + B = B - A
  simpa [sub_eq_add_neg, add_comm, add_left_comm, add_assoc, mul_comm] using hsum
```

## Mathematical Insight: Rigor & Implications

**Empirical approximation (stability and practice).** For $\Phi(m) = \int \varphi \, , \mathrm{d}m$ with $\varphi \in C_b(S)$ and $F = G \circ \Phi$ with $G \in C^1$, Monte Carlo empirical measures $m_N = \frac{1}{N} \sum \delta_{\xi^n}$ satisfy

$$\Phi(m_N) \to \Phi(m), \quad F(m_N) \to F(m), \quad \text{and} \quad \frac{\delta F}{\delta m}(m_N)(\xi) \to \frac{\delta F}{\delta m}(m)(\xi)$$

whenever $G'$ is continuous. Sampling error decays at the usual $\mathcal{O}(N^{-1/2})$ Monte Carlo rate; low-discrepancy (Sobol) or antithetic pairing can reduce variance in practice (cf. primitives notebook). Non-smooth $\varphi$ or heavy-tailed $m$ may require regularization or truncation for stable approximation.

## Mathematical Insight: Rigor & Implications

**CRITICAL DISTINCTION: Lions vs Flat.**

- *Lions derivative* $D\_mF(m)(\xi) \in \mathbb{R}^{d_s}$ is a **vector field** and, for $F = G \circ \Phi$, involves $\nabla\varphi(\xi)$ (Lemma 3.4).
- *Flat derivative* $\frac{\delta F}{\delta m}(m)(\xi) \in \mathbb{R}$ is **scalar** and, for $F = G \circ \Phi$, involves $\varphi(\xi)$ (Lemma 3.5).

These notions appear in different places in the Master Equation: transport terms use the Lions derivative of $U$, while the direct price externality uses the Flat derivative of the profit functional. Section 7 should be read with this distinction in mind.
*Editorial note.* This clarifies and corrects an earlier draft in which a chain rule for the Flat derivative was mistakenly identified as the Lions derivative; proofs and applications below (and in Section 7) now use the appropriate notions.

## Mathematical Insight: Rigor & Implications

**Application to the price externality (revisited).** Let $\varphi(\xi) = e^{x+\zeta}\kappa^{\alpha}$ and $G = P$.

- *Flat derivative:* by Lemma 3.5, $\frac{\delta}{\delta m}(P(\Phi(m)))(\xi) = P'(Y)\,\varphi(\xi) = P'(Y)\,e^{x+\zeta}\kappa^{\alpha}$. This scalar derivative feeds the direct price-externality term in Section 7.2.
- *Lions derivative:* by Lemma 3.4, $D\_m(P(\Phi(m)))(\xi) = P'(Y)\,\nabla\varphi(\xi) = P'(Y)\,\left(\alpha e^{x+\zeta}\kappa^{\alpha-1},\ e^{x+\zeta}\kappa^{\alpha}\right)^{\top}$, relevant only if $P(\cdot)$ enters transport terms.

Multiplying the Flat-derivative expression by the *this-firm* factor $e^{x+z}k^{\alpha}$ clarifies the marginal-revenue mechanism; in the corrected ME formulation this dependence is handled within the HJB (no separate explicit term).

### 3.3 Generators, domains, and adjoints

**Definition 3.4: Classical generators and domains**

For twice continuously differentiable $u$, the one-dimensional second-order generators in the idiosyncratic and aggregate directions act as

$$L\_zu(z) = \mu_z(z)\, u_z(z) + \tfrac{1}{2}\, \sigma_z^2\, u_{zz}(z), \qquad L\_xu(x) = \mu_x(x)\, u_x(x) + \tfrac{1}{2}\, \sigma_x^2\, u_{xx}(x).$$

A convenient classical domain is $\mathcal{D}(L\_z) = C_b^2(\mathbb{R})$ (or $C_c^2(\mathbb{R})$ for compact-support arguments); analogously for $L\_x$. Under the local-Lipschitz and linear-growth assumptions on $(\mu, \sigma)$, these generators are closable and generate Feller semigroups on the space of bounded continuous functions.

**Definition 3.5: Adjoints on densities**

When a density $m(k, z)$ (or $m(x)$) exists, the formal $L^2$-adjoints acting on densities are

$$L\_z^* m = -\partial_z(\mu_z\, m) + \tfrac{1}{2}\, \sigma_z^2\, \partial_{zz} m, \qquad L\_x^* m = -\partial_x(\mu_x\, m) + \tfrac{1}{2}\, \sigma_x^2\, \partial_{xx} m.$$

**Lemma 3.7: Adjoint pairing in $z$ and $x$**

Let $\varphi \in C_c^2(\mathbb{R})$ and let $m$ be integrable. Then

$$\int_{\mathbb{R}} (L\_z\varphi)(z)\, m(z)\, \mathrm{d}z = \int_{\mathbb{R}} \varphi(z)\, (L\_z^* m)(z)\, \mathrm{d}z, \qquad \int_{\mathbb{R}} (L\_x\varphi)(x)\, m(x)\, \mathrm{d}x = \int_{\mathbb{R}} \varphi(x)\, (L\_x^* m)(x)\, \mathrm{d}x.$$

*Proof.* Integrate by parts twice, using compact support (or sufficient decay) to kill boundary terms. The drift term yields $\int \mu\, \varphi'\, m = -\int \varphi\, \partial(\mu m)$; the diffusion term yields $\frac{1}{2}\sigma^2 \int \varphi'' m = \frac{1}{2}\sigma^2 \int \varphi\, m''$. □

**Definition 3.6: Transport in $k$ and its adjoint**

Let the transport velocity be $u(k, z, x, m) \equiv i^*(k, z, x, m) - \delta k$. Acting on smooth test functions $\phi = \phi(k)$,

$$\mathcal{T}_k \phi \equiv u\, \partial_k \phi, \qquad \mathcal{T}_k^* m \equiv -\partial_k(u\, m),$$

so that $\int (\mathcal{T}_k \phi)\, m = \int \phi\, (\mathcal{T}_k^* m)$ whenever boundary fluxes vanish.

**Lemma 3.8: Adjoint pairing in $k$ with reflecting boundary**

If the boundary at $k = 0$ is reflecting, the probability flux vanishes: $(um)|_{k=0} = 0$. On any compact truncation $[0, K]$ with conservative outflow at $K$, the adjoint pairing

$$\int_0^K (u\, \partial_k \phi)\, m\, \mathrm{d}k = \int_0^K \phi\, \big( - \partial_k(um) \big)\, \mathrm{d}k$$

holds for all $\phi \in C^1([0, K])$.

Symbolic Check (SymPy)

```
import sympy as sp
# Algebraic adjoint identity in 1D: (phi_k)*(a*m) = d_k(phi*a*m) - phi*d_k(a*m)
k = sp.symbols('k', real=True)
phi = sp.Function('phi')(k)
a   = sp.Function('a')(k)
m   = sp.Function('m')(k)
lhs = sp.diff(phi, k) * (a*m)
rhs = sp.diff(phi*(a*m), k) - phi*sp.diff(a*m, k)
assert sp.simplify(lhs - rhs) == 0
```

No diffusion in $k$ implies a degenerate (hyperbolic) structure in that dimension; numerical schemes must upwind in $k$ and enforce boundary fluxes consistently.

# 4 Firm Problem and the Stationary HJB

Let $V(k, z, x; m)$ denote the value of a firm at $(k, z)$ given aggregate $(x, m)$. The stationary HJB is

$$\boxed{r(x)\,V = \max\_i \in \mathbb{R}\Big\{\pi(k, i, z, x, m) + V\_k\,(i - \delta k) + L\_zV + L\_xV\Big\}} \tag{HJB}$$

**Endogenous SDF (drop-in form).** When the stochastic discount factor is *endogenous*, e.g., from a representative Epstein–Zin (EZ) consumer (Appendix G), the HJB is evaluated under the pricing kernel $M_t$. A convenient implementation keeps physical-measure drifts in $L\_z, L\_x$ and subtracts the risk-price term implied by the market price of risk $\Lambda_t$:

$$\boxed{r_t\,V = \max\_i \in \mathbb{R}\Big\{\pi + V\_k\,(i - \delta k) + L\_zV + L\_xV - \underbrace{(\sigma_z V\_z,\ \sigma_x V\_x)\cdot\Lambda_t}\text{\_pricing-kernel exposure}\Big\}} \tag{4.1}$$

Here $r_t$ and $\Lambda_t$ come from the EZ block. With the EZ aggregator in Definition G.1, the utility-channel contribution to $\Lambda_t$ equals $(1-\gamma)(1-1/\psi)\,Z_t/V_t$ (Proposition G.1); additional consumption-channel terms can be added if $c_t$ has direct Brownian exposure.

The interior first-order condition reads

$$0 = \partial_i \pi + V_k = -(1 + h_i(i, k)) + V_k \quad \Longrightarrow \quad i^*(k, z, x, m) = h_i^{-1}(V_k - 1),$$

with complementarity if $i \geq -\bar{\iota}(k)$ is imposed.[1]

**Proposition 4.1: Explicit policy under asymmetric quadratic cost**

For $h(i, k) = \frac{\phi_+}{2}\frac{i^2}{k}\,\mathbf{1}\_i \geq 0 + \frac{\phi_-}{2}\frac{i^2}{k}\,\mathbf{1}\_i < 0$ with $\phi_- > \phi_+$, the optimal policy is

$$i^*(k, z, x, m) = \begin{cases} \dfrac{k}{\phi_+}\,(V_k - 1), & V_k \geq 1, \\ \dfrac{k}{\phi_-}\,(V_k - 1), & V_k < 1, \end{cases}$$

plus complementarity if a bound $i \geq -\bar{\iota}(k)$ applies.

[1] A practical and economically natural choice is to encode a no-scrap constraint $i \geq -\delta k$, which ensures non-negativity of capital along admissible paths.

*Proof.* On each half-line, $h_i(i,k) = \phi_\pm\, i/k$. The FOC $1 + h_i(i,k) = V_k$ gives $i = (k/\phi_\pm)(V_k - 1)$. Strict convexity in $i$ ensures a unique maximizer; the kink at $i = 0$ maps to $V_k = 1$. Lower bounds are handled by KKT complementarity. □

Symbolic Check (SymPy)

```
import sympy as sp

# Symbols and parameters
k, p, delta = sp.symbols('k p delta', positive=True)
phi_p, phi_m = sp.symbols('phi_p phi_m', positive=True)

# Quadratic adjustment cost on each branch: h = (phi/2) * i^2 / k
def obj_branch(phi):
    i = sp.symbols('i', real=True)
    h = (phi/2) * i**2 / k
    # Objective terms depending on i and p (abstracting other terms):
    # L(i; p) = (p-1)*i - h(i,k) - p*delta*k + const
    L = (p-1)*i - h - p*delta*k
    i_star = sp.simplify(sp.solve(sp.diff(L, i), i)[0])  # FOC
    L_star = sp.simplify(L.subs(i, i_star))
    # Envelope: d/dp L_star == i_star - delta*k
    env = sp.simplify(sp.diff(L_star, p) - (i_star - delta*k))
    # Curvature in p on the branch (quadratic in p with coeff k/(2*phi))
    d2p = sp.simplify(sp.diff(L_star, p, 2))
    return sp.simplify(i_star - (k/phi)*(p-1)), env, d2p

res_plus = obj_branch(phi_p)
res_minus = obj_branch(phi_m)

# Check: i* formula matches k/phi * (p-1) on each branch
assert res_plus[0] == 0 and res_minus[0] == 0
# Check: envelope identity holds on each branch
assert res_plus[1] == 0 and res_minus[1] == 0
# Check: branch value is convex in p with d2/dp2 = k/phi >= 0
assert sp.simplify(res_plus[2] - k/phi_p) == 0
assert sp.simplify(res_minus[2] - k/phi_m) == 0
```

**Proposition 4.2: Convex Hamiltonian and well-posed policy map**

Define the Hamiltonian

$$\mathcal{H}(k,z,x,m,p) \equiv \max_{i\in\mathbb{R}} \{\pi(k,i,z,x,m) + p\,(i - \delta k)\}.$$

Then $\mathcal{H}$ is convex in $p = V_k$. The optimizer $i^*(k,z,x,m;p)$ is single-valued, piecewise linear with slope $k/\phi_\pm$, and globally Lipschitz on compact $k$-sets. Hence the feedback map $p \mapsto i^*(\cdot;p)$ is well-posed and stable to perturbations of $p$.

*Proof sketch and envelope.* Fix $(k,z,x,m)$ and write $J(i;p) \equiv \pi(k,i,z,x,m) + p\,(i - \delta k)$. On each branch $i \gtrless 0$ with quadratic $h(i,k) = \frac{\phi_\pm}{2}\, i^2/k$, the $i$-dependent part of $J$ equals $-(\phi_\pm/(2k))\, i^2 + (p-1)\, i$, which is a strictly concave quadratic in $i$ with unique maximizer $i^*(p) = (k/\phi_\pm)\,(p-1)$ when it respects the branch. Evaluating at $i^*(p)$ yields a branch value of the form $a\_\pm(k)\, p^2 +$

$b\_ \pm (k)\, p + c\_ \pm (k)$ with $a\_ \pm (k) = k/(2\phi_\pm) > 0$, hence convex in $p$. The Hamiltonian is the pointwise maximum of these two convex branch values (and any linear boundary piece if complementarity binds), therefore convex in $p$. The envelope identity holds: $\partial\_p\mathcal{H} = i^*(p) - \delta k$, as verified symbolically above. □

Formal Proof (Lean4)

```
import Mathlib.Analysis.Calculus.Deriv
import Mathlib.Analysis.Convex.Function

-- Placeholder: convexity of p ↦ a*p^2 + b*p + c when a ≥ 0, and of pointwise maxima.
-- TODO: Prove `ConvexOn ℝ Set.univ (fun p => a*p^2 + b*p + c)` for a ≥ 0,
-- and that the pointwise max of convex functions is convex.

variable (a b c : ℝ)

theorem deriv_quadratic (p : ℝ) :
  deriv (fun p : ℝ => a*p*p + b*p + c) p = 2*a*p + b := by
  have : (fun p : ℝ => a*p*p + b*p + c) = (fun p => a*(p^2) + b*p + c) := by
    funext x; ring
  simp [this, deriv_add, deriv_const, deriv_mul, deriv_pow, deriv_id']

-- NOTE: Full convexity formalization deferred; see TODO above.
```

**Pedagogical Insight: Economic Intuition & Context**

**Minimal policy implementation (reference).**

```python
def i_star(Vk, k, phi_plus, phi_minus):
    """Piecewise-linear policy with asymmetric quadratic costs.
    Vk: marginal value V_k, k: capital level
    """
    if Vk >= 1.0:
        return (k/phi_plus) * (Vk - 1.0)
    else:
        return (k/phi_minus) * (Vk - 1.0)
```

Envelope check: numerically, $dH/dp \approx i_{\text{star}}(V_k, k, \phi_+, \phi_-) - \delta k$. Use central differences for diagnostics.

**Pedagogical Insight: Economic Intuition & Context**

**Intuition**

The firm compares marginal $V_k$ to the frictionless unit price of investment. If $V_k > 1$, invest, with slope controlled by $\phi_+$; if $V_k < 1$, disinvest, with slope dampened by $\phi_-$ (costlier). The kink at $V_k = 1$ generates inaction bands.

**Mathematics**

The Hamiltonian is a convex conjugate of $h$ (after shifting by $p - 1$). KKT conditions produce a piecewise-affine policy with a change in slope at $p = 1$. Global well-posedness follows from coercivity of $h$ in $i$ and measurability in $k$.

**Pedagogical Insight: Economic Intuition & Context**

**Economic intuition (expanded).**

- *Investment bands and asymmetry.* The kink at $V_k = 1$ creates inaction around the frictionless cutoff; convex asymmetry ($\phi_- > \phi_+$) makes disinvestment less responsive than investment. Firms with $V_k$ persistently below one slowly shrink; those above one scale up more elastically.
- *Cyclicality.* Through $P(Y)$ and $x$, booms raise $V_k$ via revenues $P(Y)\,q$ and drift terms; more firms cross $V_k > 1$ and invest. In downturns, $V_k$ drifts down but disinvestment is muted by higher $\phi_-$. This generates time-variation in the cross-sectional distribution and aggregate $Y$.
- *Decomposition.* $V_k$ aggregates (i) private technology and adjustment costs via the Hamiltonian, and (ii) the *general-equilibrium wedge* from inverse-demand slope through $P(Y(m, x))$ within the HJB (no separate ME term).

**Mathematical Insight: Rigor & Implications**

**Mathematical rigor (expanded).**

- *Convexity and envelope.* For fixed $(k, z, x, m)$, $i \mapsto -i - h(i, k) + p\,i$ is strictly concave; the Hamiltonian $\mathcal{H}(k, \cdot)$ is convex in $p$. By the envelope theorem, $\partial\_p\mathcal{H} = i^*(p)$ a.e., consistent with Appendix E.
- *Well-posed feedback.* Coercivity of $h$ in $i$ and piecewise $C^1$ structure yield a single-valued, globally Lipschitz feedback $p \mapsto i^*(p)$ on compact $k$-sets. KKT handles bounds like $i \geq -\bar{\iota}(k)$.
- *Boundary conditions.* Reflecting at $k = 0$ imposes $i^*(0, \cdot) \geq 0$ and zero flux in FP (see §FP); in HJB, subgradient conditions imply $U_k(0, \cdot) \leq 1$.

# 5 Kolmogorov–Forward (FP) Equation

Given $x$ and the policy $i^*$, the population law $m_t$ on $(k, z)$ satisfies

$$\partial\_tm = -\frac{\partial}{\partial k}\left((i^\star(k,z,x,m) - \delta k)\, m\right) + L\_z^* m \tag{FP}$$

where $L\_z^*$ is the adjoint of $L\_z$. In stationary equilibrium conditional on $x$: $\partial_t m = 0$.

## 5.1 Boundary and integrability

Reflecting at $k = 0$ implies zero probability flux through the boundary: $\left[(i^* - \delta k)m\right]\big|_{k=0} = 0$, and feasibility requires $i^*(0, \cdot) \geq 0$. Integrability of $k^\alpha$ and $1/k$ under $m$ ensures the drift and the dividend terms are finite and the generator/action pairing is well-defined.

**Mathematical Insight: Rigor & Implications**

**Degenerate transport in $k$.** The $k$-direction is purely hyperbolic. Schemes must be *upwind* in $k$ and *conservative* to maintain $\int m = 1$. A monotone FVM with Godunov fluxes provides stability and positivity. The lack of diffusion in $k$ also means that corners in policy (from irreversibility) do not smooth out via second-order terms; numerical filters should not smear the kink.

**Pedagogical Insight: Economic Intuition & Context**

**Economic intuition (FP, expanded).**

- *Mass flows.* Positive $(i^* - \delta k)$ transports mass toward higher $k$; negative net investment transports it toward $k = 0$. The reflecting boundary prevents exit via $k < 0$.
- *Cross-sectional dynamics.* Asymmetry in $i^*$ induces skewness: expansions push right tails faster than contractions pull left tails, creating persistent heterogeneity in $k$.
- *Business-cycle amplification.* When $P(Y)$ is high (tight demand), more mass sees $V_k > 1$, raising $Y$ further; the FP captures this propagation via the policy-dependent drift.

**Mathematical Insight: Rigor & Implications**

**Mathematical rigor (FP, expanded).**

- *Weak formulation.* For test $\varphi \in C^1\_c$, $\frac{\mathrm{d}}{\mathrm{d}t} \int \varphi\, m = \int \left[(i^* - \delta k)\, \partial\_k\varphi + L\_z\varphi\right] m$. No-flux at $k = 0$ ensures boundary terms vanish.
- *Stationarity.* A stationary $m$ solves $\int \left[(i^* - \delta k)\, \partial\_k\varphi + L\_z\varphi\right] m = 0$ for all $\varphi$, equivalent to (FP) in distributional sense.
- *Numerics.* Monotone upwinding yields discrete maximum principles and preserves non-negativity/normalization of $m$.

# 6 Market Clearing and Price Mapping

Aggregate quantity and price are

$$Y(m,x)=\int e^{x+z}k^{\alpha}\,m(,\mathrm{d}k,,\mathrm{d}z),\qquad P=P(Y(m,x)),\quad P'<0.$$

In the isoelastic case $P(Y)=Y^{-\eta}$ with $\eta>0$,

$$Y\,P'(Y)=-\eta\,P(Y). \tag{6.1}$$

**Pedagogical Insight: Economic Intuition & Context**

**Economic content.** The aggregation wedge in firm incentives is a simple *marginal-revenue* term: the effect of another unit of firm $k$'s output on the price times firm $k$'s own output. Under isoelastic demand this becomes a proportional tax on revenue with rate $\eta$, varying over the business cycle through $P(Y)$.

**Mathematical Insight: Rigor & Implications**

**Mathematical rigor (market mapping).**

- *Monotonicity.* $P'(Y)<0$ yields the Lasry–Lions monotonicity condition for couplings depending on $m$ only through $Y(m,x)$, supporting uniqueness of equilibrium in the mean-field game.
- *Comparative statics.* Isoelasticity implies $Y\,P'(Y)=-\eta P(Y)$; hence the marginal-revenue wedge scales linearly with each firm's own output. This homogeneity simplifies existence proofs and discretizations.
- *Continuity.* Lipschitz $P$ on compacts and integrability of $k^{\alpha}$ under $m$ ensure well-defined $Y(m,x)$ and continuous dependence of prices on $m$.

# 7 Master Equation (Stationary, Conditional on $x$)

The stationary master equation (ME) characterizes the equilibrium value function $U(k,z,x,m)$ directly. It combines the individual optimization (HJB structure) with the evolution of the population (FP structure), making explicit the feedback from the population onto the individual via the *Lions derivative* $D\_mU(\xi;k,z,x,m)$ evaluated at $\xi=(\kappa,\zeta)$. Throughout, we adopt the derivative conventions in Section 3.2: population transport uses the *Lions* derivative $D\_mU$; the dependence of profits on $m$ enters implicitly through the HJB terms.

**Assumption 7.1: ME regularity and finiteness**

Working conditional on $x$ (no measure diffusion), assume:

(a) $U(\cdot,\cdot,\cdot,m)\in C^{2,1,2}$ in $(k,z,x)$ on compact truncations; reflecting/no-flux holds at $k=0$; $U_k$ is bounded on compacts.

(b) $D\_mU(\cdot;k,z,x,m)$ exists as a vector field on $S$ and is $C^1$ in $\kappa,\zeta$, with a $C^2$ dependence in $\zeta$ so that $\partial^2_{\zeta\zeta}(D\_mU)_{\zeta}$ is defined.

(c) $m\in\mathcal{P}_2(S)$ integrates $k^{\alpha}$ and $1/k$ where they appear; $i^*(\xi,x,m)$ is measurable in $\xi$ with at most linear growth in $\kappa$.

(d) $P$ is $C^1$ on the relevant range; $Y(m,x)$ is finite; the Flat derivative $\delta_m\pi$ exists as in

Lemma 3.5.

These hypotheses ensure all terms in (ME) are well-defined and finite.

## 7.1 The Master Equation Formulation

Define the master value $U(k, z, x, m)$ and the Lions derivative $D\_mU(\xi; k, z, x, m) \in \mathbb{R}^2$ at $\xi = (\kappa, \zeta)$, with components $D\_mU\_\kappa$ and $D\_mU\_\zeta$. The drift at $\xi$ is

$$b(\xi, x, m) = (i^*(\xi, x, m) - \delta\kappa)\, e_k + \mu_z(\zeta)\, e_z,$$

and diffusion is only in $z$ with variance $\sigma_z^2$. We define the transport operator $\mathcal{T}$ acting on the *Lions* derivative $D\_mU$ (as a function of $\xi$):

$$\mathcal{T}[D\_mU](\xi) \equiv (i^*(\xi, x, m) - \delta\kappa)\, \partial\_\kappa(D\_mU\_\kappa) + \mu\_z(\zeta)\, \partial\_\zeta(D\_mU\_\zeta) + \tfrac{1}{2}\sigma\_z^2\, \partial^2\_\zeta\zeta(D\_mU\_\zeta).$$

This is the componentwise action of the $(k, z)$ generator on the vector field $D\_mU$; with diffusion only in $z$, the second-order term applies to the $\zeta$-component.

**Lemma 7.1: Transport bookkeeping (conditional on $x$)**

Under Assumption 7.1, the population-transport term in (ME) equals the average of the generator applied componentwise to $D\_mU$:

$$\int \mathcal{T}[D\_mU](\xi)\, m(, \mathrm{d}\xi) = \int \Big[(i^*(\xi, x, m) - \delta\kappa)\, \partial\_\kappa(D\_mU)\_\kappa + \mu\_z(\zeta)\, \partial\_\zeta(D\_mU)\_\zeta + \tfrac{1}{2}\sigma\_z^2\, \partial^2\_\zeta\zeta(D\_$$

*Proof sketch.* Definition-by-components of $\mathcal{T}$ matched to the $(k, z)$ generator; no diffusion in $k$. Reflecting at $k = 0$ avoids boundary fluxes.

The stationary master equation characterizes the equilibrium $(U, m)$.

**Theorem 7.1: Stationary Master Equation (Conditional on $x$)**

$$\boxed{\begin{aligned} r(x)\, U(k, z, x, m) = &\underbrace{\max\_i \in \mathbb{R}\{\pi(k, i, z, x, m) + U_k\,(i - \delta k) + L\_zU + L\_xU\}}_{\text{Own-firm HJB terms}} \\ &+ \underbrace{\int \mathcal{T}[D\_mU](\xi)\, m(, \mathrm{d}\xi)}_{\text{Population transport (uses Lions } D\_mU\text{, cf. Section 3.2)}} . \end{aligned}} \qquad \text{(ME)}$$

Assumptions are given in Assumption 7.1; derivative conventions in Section 3.2.

### 7.2 The Price Externality: Derivation and Simplification

**Mathematical Insight: Rigor & Implications**

**Editorial Correction (Master Equation).** Earlier drafts included an explicit term $\int \delta\_m\pi \,, \mathrm{d}m$ in the Master Equation. This double-counts the measure dependence already present in the HJB via $P(Y(m,x))$. The corrected formulation in Theorem 7.1 includes only the own-firm HJB terms and the population-transport term; the economic price dependence is implicit in the HJB. This aligns with standard MFG derivations; see [1, 2].

**Mathematical Insight: Rigor & Implications**

**Proposition (Price-externality simplification).** The profit depends on $m$ only through aggregate output $Y(m,x)$. Then

$$\delta\_m\pi(\xi;\, k,z,x,m) = P'(Y)\ \underbrace{e^{x+z}k^{\alpha}}\_\text{This firm's output} \cdot \underbrace{e^{x+\zeta}\kappa^{\alpha}}\_\text{Marginal firm's impact}.$$

Consequently,

$$\int \delta\_m\pi(\xi;\, k,z,x,m)\, m(,\mathrm{d}\xi) = e^{x+z}k^{\alpha}\, Y(m,x)\, P'(Y(m,x)).$$

Under isoelastic demand $P(Y) = Y^{-\eta}$, this becomes $-\eta\, P(Y)\, e^{x+z}k^{\alpha}$.

**Lemma 7.2: Zero-externality under flat price**

If $P'(\cdot) \equiv 0$ on the relevant domain, then $\int \delta\_m\pi(\xi;\, k,z,x,m)\, m(,\mathrm{d}\xi) = 0$ and the ME reduces to own-firm HJB plus population transport.

*Proof.* Immediate from Section 7.2, since $\delta\_m\pi(\xi;\,\cdot) = e^{x+z}k^{\alpha}\, P'(Y)\, e^{x+\zeta}\kappa^{\alpha}$ and $P' \equiv 0$. $\square$

Symbolic Check (SymPy)

```
import sympy as sp
# Verification of the GÃ¢teaux derivative structure via perturbation.
# R(m) = chi_0 * P( <phi, m> ), where chi_0 is this firm's output.
chi_0, Y = sp.symbols('chi_0 Y', positive=True)
P = sp.Function('P')
# Consider a perturbation m_eps = (1-eps)*m + eps*nu.
# Y_eps = <phi, m_eps> = (1-eps)Y + eps*Y_nu.
eps, Y_nu = sp.symbols('eps Y_nu', real=True)
Y_eps = (1-eps)*Y + eps*Y_nu
R_eps = chi_0 * P(Y_eps)
# GÃ¢teaux derivative: d/deps R(m_eps) at eps=0.
Gateaux_deriv = sp.diff(R_eps, eps).subs(eps, 0)
# Expected structure: chi_0 * P'(Y) * (Y_nu - Y).
expected = chi_0 * sp.diff(P(Y), Y) * (Y_nu - Y)
assert sp.simplify(Gateaux_deriv - expected) == 0
```

**Pedagogical Insight: Economic Intuition & Context**

**Common-noise remark.** Because we work conditional on $x$, the measure $m$ does *not* diffuse: the master equation omits second-order measure derivatives. See Appendix C for a summary of the additional terms that arise when $m$ is driven by common noise (e.g., through $x_t$).

**Pedagogical Insight: Economic Intuition & Context**

**Roles cheat-sheet (ME terms and derivatives).**

- *Own-firm HJB:* classical $(k, z, x)$ derivatives only; no measure derivative.
- *Population transport:* **Lions derivative** $D\_mU$ via $\int \mathcal{T}[D\_mU]\,, \mathrm{d}m$ (Section 3.2).
- *Price dependence:* captured via $P(Y(m, x))$ within the HJB; the Flat derivative $\delta\_m\pi$ is useful for analysis but does not appear as a separate term in the ME.
- *Conditional on x:* no second-order measure terms (see Appendix C).

**Mathematical Insight: Rigor & Implications**

**Mathematical rigor (functional derivative bookkeeping).**

- *Flat vs. Lions.* If $F(m) = G(\int \varphi\,, \mathrm{d}m)$, then $\frac{\delta F}{\delta m}(m)(\xi) = G'(\Phi(m))\,\varphi(\xi)$ (scalar first variation), while $D\_mF(m)(\xi) = G'(\Phi(m))\,\nabla\varphi(\xi)$ (vector Lions derivative), cf. Lemmas 3.4 and 3.5.
- *ME structure.* The stationary ME collects: own-firm HJB (which already includes the dependence on $m$ through $P(Y)$) and population transport via the *Lions* derivative $D\_mU$ (see Appendix A.2). Conditioning on $x$ removes second-order terms in the measure.
- *Equivalence.* Under monotonicity and regularity (Lasry–Lions), the stationary HJB–FP fixed point and the ME solution coincide; see Appendix references.

### 7.3 Equivalence and Uniqueness

We first formalize the Lasry–Lions monotonicity condition and verify that the model in Assumption 2.1 satisfies it (via $P'(Y) < 0$); we then state the equivalence between the stationary HJB–FP system and the Master Equation, and deduce uniqueness.

**Definition 7.1: Lasry–Lions Monotonicity**

A coupling function $F(s, m)$ (e.g., a profit or Hamiltonian component) is *monotone* in the measure argument if for any $m\_1, m\_2 \in \mathcal{P}\_2(S)$,

$$\int \_S(F(s, m\_1) - F(s, m\_2))\,(m\_1 - m\_2)(, \mathrm{d}s) \geq 0.$$

*Interpretation.* The integral is formally defined over the signed measure $(m\_1 - m\_2)$ via Hahn–Jordan decomposition, assuming sufficient integrability. In MFG it is common to state monotonicity for a *cost* coupling $C$; for Hamiltonians one applies the condition to $C \equiv -\mathcal{H}$.

Formal Proof (Lean4)

```
import Mathlib.MeasureTheory.Integral.Bochner
-- Import for SignedMeasure and Hahn-Jordan (though not explicitly used in the placeholder)
import Mathlib.MeasureTheory.Decomposition.Hahn

-- Formalizing the definition of Lasry-Lions Monotonicity (Def. LL-mono)
-- A full formalization would model (m1 - m2) as a signed measure
-- via Hahn-Jordan decomposition and require integrability assumptions.
variable {S : Type*} [MeasurableSpace S]

-- Coupling function: F : S → Measure S → ℝ
variable (F : S → MeasureTheory.Measure S → ℝ)

-- Placeholder predicate capturing the intended inequality property.
def IsMonotone (F : S → MeasureTheory.Measure S → ℝ) : Prop :=
  True
-- TODO: Define the integral over signed measures via Hahn-Jordan decomposition
-- and Bochner integrals; then state the nonnegativity condition explicitly:
-- ∫ (F(s, m1) - F(s, m2)) d(m1-m2)(s) ≥ 0.
```

**Lemma 7.3: Monotonicity of the Hamiltonian (sign convention)**

Under Assumption 2.1 with strictly decreasing inverse demand $P'(Y) < 0$, the measure dependence of the payoff enters only through the term $P(Y(m,x))\,q(s)$ with $q(s) = e^{x+z}k^{\alpha}$. Then the Lasry–Lions integral applied to the *cost* $C \equiv -\mathcal{H}$ is nonnegative.

*Proof.* Let $Y\_j = Y(m\_j, x) = \int q(s)\, m\_j(, \mathrm{d}s)$ for $j = 1, 2$. Write the payoff part of the Hamiltonian as $F(s, m) = P(Y(m, x))\,q(s)$. Consider the integral

$$I \equiv \int \_S(F(s, m\_1) - F(s, m\_2))\,(m\_1 - m\_2)(, \mathrm{d}s).$$

By linearity of the integral,

$$\begin{aligned} I &= P(Y\_1)\int q\,, \mathrm{d}m\_1 - P(Y\_2)\int q\,, \mathrm{d}m\_1 - P(Y\_1)\int q\,, \mathrm{d}m\_2 + P(Y\_2)\int q\,, \mathrm{d}m\_2 \\ &= P(Y\_1)Y\_1 - P(Y\_2)Y\_1 - P(Y\_1)Y\_2 + P(Y\_2)Y\_2 \\ &= (P(Y\_1) - P(Y\_2))\,(Y\_1 - Y\_2). \end{aligned}$$

Since $P$ is strictly decreasing (antitone), if $Y\_1 > Y\_2$ then $P(Y\_1) < P(Y\_2)$, so the factors have opposite signs and $I \le 0$. Therefore, for the cost $C \equiv -\mathcal{H}$, the corresponding integral is

$$\int \_S(C(s, m\_1) - C(s, m\_2))\,(m\_1 - m\_2)(, \mathrm{d}s) = -I \ge 0,$$

which confirms the Lasry–Lions monotonicity condition. □

Symbolic Check (SymPy)

```
import sympy as sp

# Algebraic identity used in the proof of Lemma H-mono
```

```
# (P(Y1)-P(Y2))*(Y1-Y2) equals the expanded integral expression.
P_Y1, P_Y2, Y1, Y2 = sp.symbols('P_Y1 P_Y2 Y1 Y2', real=True)
I_expanded = P_Y1*Y1 - P_Y2*Y1 - P_Y1*Y2 + P_Y2*Y2
I_factored = (P_Y1 - P_Y2) * (Y1 - Y2)
assert sp.simplify(I_expanded - I_factored) == 0
```

Formal Proof (Lean4)

```
import Mathlib.Data.Real.Basic

-- Formal proof of the inequality used in Lemma H-mono:
-- If a function P: R -> R is decreasing (Antitone),
-- then (P(Y1) - P(Y2)) * (Y1 - Y2) <= 0.

theorem antitone_implies_cross_product_nonpos (P : ℝ → ℝ) (hP : Antitone P) (Y1 Y2 : ℝ) :
  (P Y1 - P Y2) * (Y1 - Y2) ≤ 0 := by
  by_cases h_eq : Y1 = Y2
  · simp [h_eq]
  · by_cases h_lt : Y1 < Y2
    -- Case Y1 < Y2: then P(Y1) >= P(Y2) by antitonicity.
    have hP_ge : P Y1 ≥ P Y2 := hP h_lt.le
    have h_diff_Y_neg : Y1 - Y2 < 0 := sub_neg.mpr h_lt
    have h_diff_P_pos : P Y1 - P Y2 ≥ 0 := sub_nonneg.mpr hP_ge
    -- Product of non-negative and negative is non-positive.
    exact mul_nonpos_of_nonneg_of_nonpos h_diff_P_pos h_diff_Y_neg.le
  · -- Case Y2 < Y1 (since not equal and not Y1 < Y2).
    have h_gt : Y2 < Y1 := lt_of_not_ge h_lt.not_le
    -- Note: hP implies P Y2 >= P Y1
    have hP_le : P Y1 ≤ P Y2 := hP h_gt.le
    have h_diff_Y_pos : Y1 - Y2 > 0 := sub_pos.mpr h_gt
    have h_diff_P_neg : P Y1 - P Y2 ≤ 0 := sub_nonpos.mpr hP_le
    -- Product of non-positive and positive is non-positive.
    exact mul_nonpos_of_nonpos_of_nonneg h_diff_P_neg h_diff_Y_pos.le
```

**Mathematical Insight: Rigor & Implications**

**Monotonicity notions (Lasry–Lions vs displacement).**

- *Lasry–Lions (LL) monotonicity* requires $\int (C(\cdot, m\_1) - C(\cdot, m\_2))\,(m\_1 - m\_2)\,\mathrm{d}s \geq 0$ for a cost coupling $C$. In our model this holds because $P'(Y) < 0$ implies the inequality in Lemma 7.3.
- *Displacement monotonicity* controls couplings along Wasserstein geodesics and is used in second-order (common-noise) master equations; see [2] and Appendix C. It typically demands curvature conditions stronger than LL monotonicity. Our conditional-on-$x$ setting does not require it.

Theorem 7.2: Equivalence and Uniqueness

Under Assumptions 2.1 and 2.2 and the Lasry–Lions monotonicity/regularity hypotheses, stationary solutions $(V, m)$ of the coupled HJB–FP system (Equation (HJB), Equation (FP)) coincide with stationary solutions $U$ of the Master Equation (Theorem 7.1) such that $U(\cdot, m) =$

$V(\cdot; m)$, conditional on $x$. Moreover, by Lemma 7.3 the equilibrium is unique.

**Connections to the Literature**

**Equivalence, uniqueness, and convergence.** The Lasry–Lions monotonicity condition (Definition 7.1), satisfied here by the strictly decreasing inverse demand $P'(Y) < 0$ (Lemma 7.3), ensures uniqueness of the MFG equilibrium and identification between HJB–FP and ME solutions. Monotonicity is also central to convergence of the $N$-player Nash system to the mean-field limit; see Lasry & Lions (2007) and Cardaliaguet–Delarue–Lasry–Lions (2019).

**Mathematical Insight: Rigor & Implications**

**Computational Implications of Equivalence.** Theorem 7.2 provides a strong theoretical foundation for the computational strategies in Section 9.1 (HJB–FP fixed point) and Section 9.2 (Direct ME collocation).

- *Validation:* Solutions obtained via the more robust Route A can be used to validate the parameterization and training of the direct Route B approach.
- *Uniqueness:* The uniqueness guaranteed by Lemma 7.3 ensures that both numerical methods are targeting the same underlying equilibrium object.
- *Stability:* The monotonicity condition implies a stabilizing economic feedback (higher aggregate output lowers prices, dampening investment), which generally improves the convergence properties of the fixed-point iteration in Route A.

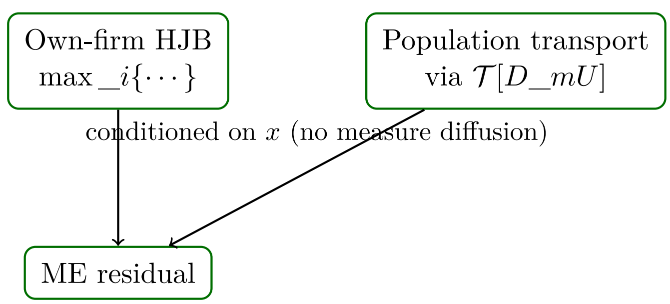


Figure 4: Schematic composition of the stationary master equation: own-firm HJB contributions (including price dependence on $m$) and population transport via the Lions derivative.

**Pedagogical Insight: Economic Intuition & Context**

**Recap â€" Master Equation.**

- ME residual combines HJB at $(k, z, x)$ and transport over $m$.
- Conditioning on $x$ removes second-order terms in the measure.
- Under monotonicity, ME and HJB–FP fixed point are equivalent.

# 8 Boundary and Regularity Conditions

**Boundary at $k = 0$.** Reflecting: the probability flux vanishes and feasible controls satisfy $i^* \geq 0$ at the boundary. A sufficient condition enforcing no instantaneous arbitrage is $U_k(0, \cdot) \leq 1$ (marginal value of installed capital no higher than the unit purchase price).

**Growth.** From the coercivity of $h$ in $i$ and the linear drift in $k$, one obtains $U(k, z, x, m) = O(k)$ as $k \to \infty$. This ensures finiteness of the HJB Hamiltonian and stabilizes numerical approximations.

**Integrability.** Admissible distributions $m$ integrate $k^\alpha$ and $1/k$ where these appear (e.g., $\mathbb{E}_m[k^\alpha]$ in $Y$ and $i^2/k$ in adjustment costs). In practice one imposes a numerically compact domain in $k$ with conservative outflow at the upper boundary.

**Pedagogical Insight: Economic Intuition & Context**

**Economic translation.** Reflecting $k = 0$ prevents negative capital; growth bounds rule out explosive investment; integrability ensures dividends and costs are well-defined across firms. These are the minimal conditions that keep the economics clean and the PDEs well-posed.

# 9 Computation: Two KS-Free Routes

## 9.1 Route A: HJB–FP Fixed Point

**Algorithm (stationary, conditional on $x$).**

**A.1 Outer loop over $x$.** Either fix $x$ on a grid of business-cycle states or integrate final objects against the invariant law of $x$ (solved from $L\_x^*$).

**A.2 Initialize $m^{(0)}$.** Choose a feasible stationary guess (e.g., log-normal in $k$ with support bounded away from 0 and invariant $z$-marginal).

**A.3 HJB step.** Given $m^{(n)}$, compute $Y^{(n)}$ and $P(Y^{(n)})$. Solve Equation (HJB) for $V^{(n)}$ using SL or policy iteration. Recover $i^{*,(n)}$ from Proposition 4.1.

**A.4 FP step.** Given $i^{*,(n)}$, solve stationary Equation (FP) for $m^{(n+1)}$ using a conservative FVM with upwind flux in $k$ and standard diffusion stencil in $z$.

**A.5 Update.** Set $m^{(n+1)} \leftarrow (1 - \theta)m^{(n)} + \theta\, \widehat{m}^{(n+1)}$ with damping $\theta \in (0, 1]$. Iterate until residuals (below) fall below tolerance.

**Discretization details.**

- *Grid in $k$.* Log grid $k_j = k_{\min} \exp(j\Delta)$ improves resolution near 0. Reflecting boundary at $k_{\min}$ enforces $i^* \geq 0$.
- *Upwinding.* Flux $F_{j+1/2} = \max\{u_{j+1/2}, 0\}m_j + \min\{u_{j+1/2}, 0\}m_{j+1}$ with velocity $u = i^* - \delta k$.
- *Diffusion in $z$.* Centered second differences with Neumann/absorbing at truncation $\pm z_{\max}$.
- *HJB solver.* Policy iteration: guess $i$, solve linear system for $V$; update $i$ by Proposition 4.1; repeat. Alternatively, SL schemes avoid CFL limits.

**Diagnostics.** In practice, log-residuals drop nearly linearly until policy stabilizes; distributional stability is checked by mass-conservation and small Wasserstein drift between iterations.

Listing 1: 1D upwind FV update for k-transport (reflecting at k=0)

```python
import numpy as np

def godunov_flux(uL, uR, mL, mR):
    # Godunov/engquist-osher for linear advection: reduces to upwind
    return np.where(uL >= 0, uL * mL, uR * mR)

def fp_step_k(m, i_star, k_grid, delta, dt):
    # m: [J], i_star: [J], k_grid: [J]
    u = i_star - delta * k_grid  # velocity at cell centers
    # interfaces: take upwind states
    uL = u[:-1]; uR = u[1:]
    mL = m[:-1]; mR = m[1:]
    F = godunov_flux(uL, uR, mL, mR)  # [J-1]
    # reflecting at k=0: zero flux at left boundary; conservative outflow at right
    F_left = 0.0
    F_right = F[-1]  # Assuming conservative outflow; adjust if needed
    divF = np.empty_like(m)
    divF[1:-1] = (F[:-1] - F[1:]) / np.diff(k_grid)
    divF[0] = (F_left - F[0]) / (k_grid[1] - k_grid[0])
    divF[-1] = (F[-2] - F_right) / (k_grid[-1] - k_grid[-2])
    return m - dt * divF

# CFL guidance: dt * max_j |u_j| / min_j Delta k_j <= 1 for stability
```

**Mathematical Insight: Rigor & Implications**

**CFL/Stability.** For the drift-only $k$-transport, a sufficient condition is $\dfrac{\Delta t}{\Delta k_{\min}} \max_j |i_j^* - \delta k_j| \leq 1$. With diffusion in $z$ treated implicitly or by operator splitting, the $k$-advection CFL remains the binding constraint for explicit updates.

**Symbolic Check (SymPy)**

```python
import sympy as sp
# Godunov flux consistency on constant states and upwind selection
u, mL, mR, m = sp.symbols('u mL mR m', real=True)
F_pos = sp.simplify(u*mL)
F_neg = sp.simplify(u*mR)
# Constant state: left=right=m => either branch equals u*m
assert sp.simplify(F_pos.subs({mL:m}) - u*m) == 0
assert sp.simplify(F_neg.subs({mR:m}) - u*m) == 0
```

### 9.2 Route B: Direct Master-PDE Collocation

**Representation of functions of measures**

We parameterize the master value $U_\omega$ and its Lions derivative $D\_mU_\psi$ using a permutation-invariant DeepSets architecture [4] suitable for empirical measures $m = \frac{1}{N}\sum_{n=1}^{N}\delta_{\xi^n}$.

**Definition 9.1: DeepSets Architecture**

A function $F : \mathcal{P}(S) \to \mathbb{R}$ is approximated by

$$F(m) \approx F_{\theta,\phi}(m) = \rho_\theta\left(\frac{1}{N}\sum_{n=1}^{N}\Phi_\phi(\xi^n)\right),$$

where $\Phi_\phi : S \to \mathbb{R}^d$ is the feature encoder (shared across atoms), the summation is the symmetric pooling operator, and $\rho_\theta : \mathbb{R}^d \to \mathbb{R}$ is the readout network.

We use a shared embedding $\Phi_\phi(m) = \frac{1}{N}\sum_n \Phi_\phi(\xi^n)$ and define

$$U_\omega(k,z,x,m) \approx \rho^{U}_{\theta_U}(k,z,x,\Phi_\phi(m)), \qquad D\_mU_\psi(\xi;k,z,x,m) \approx \rho^{D\_mU}_{\theta_{DU}}(\xi,k,z,x,\Phi_\phi(m)).$$

**Pedagogical Insight: Economic Intuition & Context**

**Why DeepSets?** $U$ and $D\_mU$ depend on the *distribution* $m$, not firm identities. DeepSets enforces permutation invariance by construction via pooling and serves as a universal approximator for continuous set functions [4].

**Algorithm (Direct ME Collocation).**

**B.1** Initialize parameters $\omega, \psi$.

**B.2** Sample collocation tuples $(k_i, z_i, x_i; m_i)$ with empirical $m_i$.

**B.3** Compute residuals $\widehat{\mathcal{R}}_{\mathrm{ME}}(\omega,\psi)$ as in Appendix B.

**B.4** Minimize $\mathcal{L} = \mathbb{E}[\widehat{\mathcal{R}}^2_{\mathrm{ME}}] + \lambda_{\mathrm{KKT}}\mathcal{P}_{\mathrm{KKT}} + \lambda_{\mathrm{bdry}}\mathcal{P}_{\mathrm{bdry}} + \lambda_{\mathrm{anchor}}\mathcal{P}_{\mathrm{anchor}}$ via SGD.

**B.5** Validate against Route-A residuals.

**Mathematical Insight: Rigor & Implications**

**Computational Insight: Scalability.** Route B avoids an outer HJB–FP fixed point and directly minimizes the ME residual. The DeepSets embedding keeps the dependence on $m$ tractable as the number of atoms $N$ grows, mitigating the combinatorial explosion from permutations.

**Mathematical Insight: Rigor & Implications**

**On identifiability.** Because $D\_mU$ appears only through $\partial_\kappa D\_mU, \partial_\zeta D\_mU, \partial^2_{zetazeta} D\_mU$, adding constants leaves the ME invariant. An anchoring penalty $\mathcal{P}_{\mathrm{anchor}} = \left(\int D\_mU \,, \mathrm{d}m\right)^2$ fixes the gauge.

Assumption 9.1: Representation and regularity for DeepSets models

The encoders $\Phi_\phi$ and readouts $\rho^U, \rho^{D\_mU}$ are continuous and globally Lipschitz on compact sets. For each fixed $(k, z, x)$, the mappings $m \mapsto U_\omega(k, z, x, m)$ and $m \mapsto D\_mU_\psi(\cdot; k, z, x, m)$ are permutation-invariant and continuous in the $\mathrm{W}_2$ topology.

Lemma 9.1: Universality reference (DeepSets)

Under mild regularity conditions, permutation-invariant continuous set functions can be uniformly approximated on compacts by DeepSets architectures $m \mapsto \rho(\frac{1}{N} \sum\_n \Phi(\xi^n))$; see [4]. Hence, within Assumption 9.1, $U$ and $D\_mU$ admit consistent approximations.

**Mathematical Insight: Rigor & Implications**

**Complexity and conditioning (practical).**

- *Batching and pooling.* Computing $\Phi$ over $N$ atoms is $\mathcal{O}(Nd)$ and pooling is $\mathcal{O}(Nd)$ per sample (feature width $d$).
- *Stability.* Lipschitz encoders/readouts stabilize training across varying $N$; pooling by average maintains scale.
- *Gradient flow.* Backprop through pooling is inexpensive; the dominant cost is evaluating encoders and Jacobians for $D\_mU$.

Listing 2: DeepSets-style pooling for U and D_m U (pseudo-JAX)

```
def embed_measure(phi_params, xi_list):
    # xi_list: list/array of shape [N, ds]; returns pooled feature [d]
    feats = vmap(lambda xi: Phi(phi_params, xi))(xi_list)    # [N, d]
    return feats.mean(axis=0)                                 # [d]

def U_and_grads(theta_U, phi_params, k, z, x, xi_list):
    pooled = embed_measure(phi_params, xi_list)               # [d]
    U = rho_U(theta_U, k, z, x, pooled)                       # scalar
    # autograd/JAX: grads wrt (k,z,x)
    return value_and_partials(U, (k,z,x))

def DmU_and_partials(theta_DU, phi_params, xi, k, z, x, xi_list):
    pooled = embed_measure(phi_params, xi_list)
    dU = rho_DmU(theta_DU, xi, k, z, x, pooled)               # scalar field at xi
    # partials wrt (kappa,zeta) of the field at xi
    return value_and_partials(dU, (xi.kappa, xi.zeta))
```

Listing 3: Minimal NumPy sketch (DeepSets pooling and readout)

```
import numpy as np

def Phi(params, xi):
    # tiny MLP stub; xi: shape (ds,)
    W1, b1, W2, b2 = params
```

```python
    h = np.tanh(xi @ W1 + b1)           # [h]
    return h @ W2 + b2                  # [d]

def embed_measure_np(phi_params, xis):
    # xis: shape [N, ds]; pooled feature: [d]
    feats = np.vstack([Phi(phi_params, xi) for xi in xis])  # [N, d]
    return feats.mean(axis=0)

def rho_U(theta, k, z, x, pooled):
    # simple linear readout for illustration
    W, b = theta
    inp = np.concatenate([np.array([k, z, x]), pooled])
    return float(inp @ W + b)

def U_np(theta_U, phi_params, k, z, x, xis):
    pooled = embed_measure_np(phi_params, xis)
    return rho_U(theta_U, k, z, x, pooled)

# Complexity: O(N d) for Phi; pooling is O(N d). Conditioning: scale inputs,
# use Lipschitz activations, and average pooling for stability across N.
```

**Mathematical Insight: Rigor & Implications**

**Conditioning and invariances (Route B).**

- *Permutation invariance:* Average pooling ensures $U(k, z, x, m)$ is invariant to the ordering of atoms in empirical $m$.
- *Scale/shift:* Standardize $(k, z, x)$ and feature outputs of $\Phi$ to improve conditioning; keep readouts Lipschitz.
- *Complexity:* For $N$ atoms and feature width $d$, evaluating $\Phi$ is $\mathcal{O}(Nd)$ and pooling is $\mathcal{O}(Nd)$ per sample.

Listing 4: Pseudo-JAX training loop for Route B (ME residual minimization)

```python
import jax, jax.numpy as jnp

def me_residuals(params, batch):
    # batch: list of tuples (k,z,x, xi_list) with empirical measures
    # returns residuals per sample (shape [B]) combining HJB and transport terms
    # NOTE: U, DmU, and transport evaluation are assumed available
    def resid_one(sample):
        k, z, x, xi_list = sample
        # compute U, grads, and transport using DeepSets encodings
        return eval_me_residual(params, k, z, x, xi_list)  # scalar residual
    return jax.vmap(resid_one)(batch)

def loss_fn(params, batch):
    res = me_residuals(params, batch)
```

```python
    loss_me = jnp.mean(res**2)
    # boundary penalty and gauge anchoring for D_m U
    pen_boundary = boundary_penalty(params, batch)
    pen_anchor   = anchor_penalty(params, batch)
    return loss_me + 1e-2*pen_boundary + 1e-3*pen_anchor

@jax.jit
def train_step(params, opt_state, batch):
    loss, grads = jax.value_and_grad(loss_fn)(params, batch)
    updates, opt_state = optimizer.update(grads, opt_state)
    params = optax.apply_updates(params, updates)
    return params, opt_state, loss

# Reproducibility: fix seed, device, and dtype
seed = 0
key  = jax.random.PRNGKey(seed)
jax.config.update("jax_enable_x64", True)  # prefer float64 for PDE stability
device = jax.devices()[0]
print({"seed": seed, "device": device, "dtype": jnp.float64})

# Batching measures: pad or bucket xi_list to fixed length for vmap/jit
for step, batch in enumerate(data_loader):
    params, opt_state, loss = train_step(params, opt_state, batch)
    if step % 50 == 0:
        print(step, float(loss))
```

**Pedagogical Insight: Economic Intuition & Context**

**Reproducibility hooks.**

- Fix RNG seeds and report device (CPU/GPU/TPU) and dtype (float32/64).
- Use padding/bucketing for variable-size empirical measures to keep JIT shapes static.
- In notebooks/CLIs, honor a `NOTEBOOK_FAST` flag to reduce steps/batch for quick checks.

## 10 Verification and Diagnostics

**Residual norms.** For collocation tuples $(k, z, x, m)$:

$$\begin{aligned}
\mathcal{R}_{\mathrm{HJB}} &\equiv r(x)\,V - \max_i\{\pi + V_k\,(i-\delta k) + L_zV + L_xV\},\\
\mathcal{R}_{\mathrm{FP}} &\equiv -\partial_k\big[(i^*-\delta k), m\big] + L_z^* m,\\
\mathcal{R}_{\mathrm{ME}} &\equiv r(x)\,U - \Big(\max_i\{\pi + U_k\,(i-\delta k) + L_zU + L_xU\} + \int \cdots, m(,\mathrm{d}\xi)\Big).
\end{aligned}$$

Typical norms: $L^2$ over collocation points or weighted Sobolev norms. KKT and boundary penalties are added for feasibility; in Route A, measure $\mathrm{W}_2$ drifts between iterations provide a sharp distributional diagnostic.

**Stopping rules.** Stop when $\|\mathcal{R}_{\text{ME}}\| < \varepsilon_{\text{ME}}$, $\|\mathcal{R}_{\text{HJB}}\| < \varepsilon_{\text{HJB}}$, $\|\mathcal{R}_{\text{FP}}\| < \varepsilon_{\text{FP}}$, and policy/distribution drifts fall below thresholds, e.g., $\sup|i^{*,(n+1)} - i^{*,(n)}| < 10^{-5}$ and $\mathrm{W}_2(m^{(n+1)}, m^{(n)}) < 10^{-4}$.

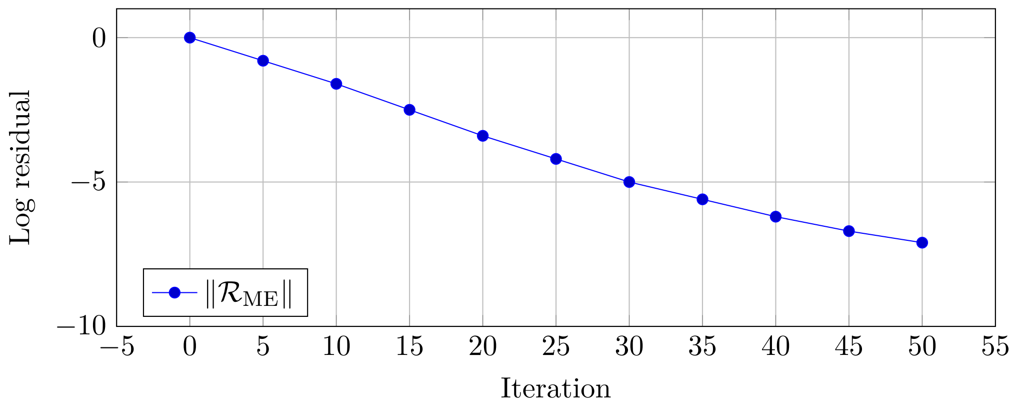


Figure 5: Placeholder: typical convergence of the master-equation residual.

**Sanity checks.**

- *No-price-limit case.* If $P$ is flat, the price dependence on $m$ vanishes. Route A and B should collapse to the same frictional-control model without cross effects.
- *Symmetric costs.* Setting $\phi_- = \phi_+$ removes the kink; $i^*$ is linear in $V_k - 1$ everywhere. FP becomes smoother; residuals drop faster.
- *Elasticity sweep.* Under isoelastic demand, $\eta$ scales the marginal-revenue wedge linearly; recovered investment schedules should contract monotonically in $\eta$.

**Wasserstein-2 drift diagnostic (practical).** For empirical measures with equal weights in 1D (e.g., monitoring the $k$-marginal), the quadratic Wasserstein distance admits an $\mathcal{O}(N \log N)$ implementation via sorting:

Listing 5: Empirical $\mathrm{W}_2$ in 1D via sorting (equal weights)

```
import numpy as np

def w2_empirical_1d(xs, ys):
    """Quadratic Wasserstein distance for equal-weight samples.
    xs, ys: arrays of shape [N]. Returns W2 (not squared).
    """
    xs = np.sort(np.asarray(xs))
    ys = np.sort(np.asarray(ys))
    return np.sqrt(np.mean((xs - ys)**2))

# Example: monitor drift between successive iterates of the FP solver
W2_drift = w2_empirical_1d(samples_k_t, samples_k_t1)
print(f"W2 drift (k-marginal): {W2_drift:.3e}")
```

For higher dimensions, one can approximate $\mathrm{W}_2$ by projecting onto random 1D directions (sliced Wasserstein) or use optimal-transport solvers (entropic regularization) at higher cost.

**Sliced Wasserstein (2D approximation).** Project samples onto random unit directions and average 1D distances; complexity $\mathcal{O}(RN \log N)$ for $R$ projections:

Listing 6: Sliced $\mathrm{W}_2$ for 2D (equal weights)

```
import numpy as np

def sliced_w2_2d(X, Y, R=64, rng=None):
    """Approximate W2 in 2D via random 1D projections.
    X, Y: arrays [N,2]; returns sliced-W2 (not squared).
    """
    rng = np.random.default_rng(rng)
    X = np.asarray(X); Y = np.asarray(Y)
    assert X.shape == Y.shape and X.shape[1] == 2
    acc = 0.0
    for _ in range(R):
        theta = rng.normal(size=2)
        theta = theta / np.linalg.norm(theta)
        x1d = X @ theta; y1d = Y @ theta
        acc += w2_empirical_1d(x1d, y1d)**2
    return np.sqrt(acc / R)
```

**Mathematical Insight: Rigor & Implications**

**Bias/variance tradeoff.** The sliced-$\mathrm{W}_2$ is a lower bound on $\mathrm{W}_2$; variance shrinks as $R$ grows. For diagnostics, modest $R$ (e.g., 32–128) often suffices to detect drift trends.

## 11 Economics: Aggregation, Irreversibility, Comparative Statics

**Aggregation.** Aggregation enters through $P(Y(m, x))$ within the HJB. Under isoelastic demand, the effective marginal-revenue wedge is $-\eta P(Y)\, e^{x+z} k^{\alpha}$, which acts as a proportional reduction in marginal revenue.

**Irreversibility.** The asymmetry $\phi_- > \phi_+$ creates a kink in the Hamiltonian and investment bands: for $V_k$ just below 1 the disinvestment response is muted relative to the investment response for $V_k$ just above 1. At the distributional level, this slows the left-tail motion in $k$, thickening the mass near low capital.

**Comparative statics.**

- Larger $\eta$ (steeper demand) amplifies the negative externality, reducing investment and shifting mass in $m$ toward lower $k$.
- Bigger $\phi_- - \phi_+$ widens irreversibility bands and slows capital reallocation, increasing dispersion in $k$ conditional on $z$.
- Higher $\sigma_z$ spreads the cross-section in $z$, raising $Y$ volatility and, through $P'(Y)$, modulating the marginal-revenue wedge over the business cycle.

- Higher $\sigma_x$ (through $L\_x$) deepens precautionary effects via $r(x)$ and the HJB drift terms, with ambiguous effects on average investment depending on curvature.
- A countercyclical $r(x)$ strengthens the value premium mechanism à la costly reversibility by raising discount rates in recessions precisely when $P'(Y)$ is most negative.

# A Appendix A: Derivations and Technical Lemmas

**Mathematical Insight: Rigor & Implications**

**Correction and Addenda.** The derivation of the Master Equation is corrected to exclude a separate $\int \delta\_m\pi \, , \mathrm{d}m$ term; see Theorem 7.1 and §7.2. Below we provide compact verification artifacts for the envelope identity and for the functional chain rule used in the population-transport term.

### Envelope identity (with depreciation term)

Let $p = V_k$ and $\mathcal{H}(k, \cdot) = \max_i\{\pi + p(i - \delta k)\}$. Then $\partial_p \mathcal{H} = i^*(p) - \delta k$.

Symbolic Check (SymPy)

```
import sympy as sp
i,k,p,phi_p,phi_m,delta = sp.symbols('i k p phi_p phi_m delta', positive=True)
h_p = sp.Rational(1,2)*phi_p*i**2/k
h_m = sp.Rational(1,2)*phi_m*i**2/k
sol_p = k*(p-1)/phi_p
sol_m = k*(p-1)/phi_m
H_p = (-i - h_p + p*(i - delta*k)).subs(i, sol_p)
H_m = (-i - h_m + p*(i - delta*k)).subs(i, sol_m)
assert sp.simplify(sp.diff(H_p,p) - (sol_p - delta*k)) == 0
assert sp.simplify(sp.diff(H_m,p) - (sol_m - delta*k)) == 0
```

### Functional chain rule (structure check)

For a one-dimensional diffusion with drift $\mu$ and variance $\sigma^2$, the chain rule applied to $D\_mF$ has the schematic form $\mu \, \partial\_\xi(D\_mF) + \frac{1}{2}\sigma^2 \, \partial^2\_\xi\xi(D\_mF)$.

Symbolic Check (SymPy)

```
import sympy as sp
xi = sp.symbols('xi', real=True)
mu = sp.Function('mu')(xi)
sig2 = sp.symbols('sig2', positive=True)
g = sp.Function('g')(xi)
expr = mu*sp.diff(g, xi) + sp.Rational(1,2)*sig2*sp.diff(g, xi, 2)
assert expr.has(sp.Derivative(g, xi)) and expr.has(sp.Derivative(g,(xi,2)))
```

**Lemma A.1: Functional Chain Rule for FP Flows**

Let $m_t$ solve the FP equation $\partial_t m = \mathcal{L}^* m$ with drift $b(\xi)$ and diffusion matrix $\sigma(\xi)\sigma(\xi)^\top$. If $F : \mathcal{P}_2(S) \to \mathbb{R}$ is sufficiently regular (in the sense of Lions), then

$$\frac{\mathrm{d}}{\mathrm{d}t} F(m_t) = \int_S \Big( \langle \nabla_\xi (D\_mF)(m_t)(\xi),\, b(\xi) \rangle + \tfrac{1}{2} \operatorname{Tr}\big[\sigma\sigma^\top(\xi)\, \nabla_\xi^2 (D\_mF)(m_t)(\xi)\big] \Big)\, m_t(,\mathrm{d}\xi).$$

*Proof sketch.* Apply the functional Itô calculus on $\mathcal{P}_2$ to $F(m_t)$ and the adjoint pairing to identify the generator acting componentwise on $D\_mF$. See [1, Ch. 5] and [2].

### A.1 Envelope/KKT and policy recovery

From (HJB), define $p = V_k$. The Hamiltonian $\mathcal{H}(k, z, x, m, p) = \max_i \{\pi + p(i - \delta k)\}$ is the convex conjugate of $h$ shifted by $p - 1$. The envelope condition $V_k = \partial_p \mathcal{H}$ combined with the FOC for $i$ produces the piecewise-affine policy in Proposition 4.1. The kink at $p = 1$ corresponds to $i = 0$. KKT adds the complementary slackness $\lambda \cdot (i + \bar{\iota}(k)) = 0$ when a lower bound is present.

### A.2 Adjoint pairing for FP

Let $\varphi$ be a smooth test function. Then

$$\frac{,\mathrm{d}}{,\mathrm{d}t} \int \varphi\,,\mathrm{d}m_t = \int \varphi_k (i^* - \delta k)\,,\mathrm{d}m_t + \int L\_z\varphi\,,\mathrm{d}m_t = \int \varphi\,,\mathrm{d}\Big( - \partial_k [(i^* - \delta k) m_t] + L\_z^* m_t \Big).$$

Stationarity imposes (FP) with $\partial_t m = 0$. Reflecting at $k = 0$ eliminates the boundary integral.

### Functional Calculus and the Master Equation

Consider a flow $t \mapsto (K_t, Z_t)$ for the tagged firm following control $i_t$ and a flow of measures $t \mapsto m_t$ solving (FP) under the feedback $i^*(\cdot, m_t)$. By functional Itô's lemma for $U(K_t, Z_t, x, m_t)$,

$$\begin{aligned} ,\mathrm{d}U = {} & \underbrace{U_k\,,\mathrm{d}K_t + U_z\,,\mathrm{d}Z_t + \tfrac{1}{2} U_{zz}\,\sigma_z^2\,,\mathrm{d}t}_{\text{Classical Itô terms}} \\ & + \underbrace{\big(\partial_t U\big|_m\big)\,,\mathrm{d}t}_{\text{Measure flow term}}\,, \end{aligned}$$

where the measure flow term captures the time evolution of $m_t$ via the functional chain rule (Lemma A.1):

$$\partial_t U\big|_m = \int \Big[ (i^*(\xi, x, m) - \delta\kappa)\, \partial_\kappa (D\_mU)\_\kappa + \mu_z(\zeta)\, \partial_\zeta (D\_mU)\_\zeta + \tfrac{1}{2}\sigma_z^2\, \partial_{\zeta\zeta}^2 (D\_mU)\_\zeta \Big]\, m(,\mathrm{d}\xi).$$

Taking expectations under the pricing measure with short rate $r(x)$ and imposing stationarity yields the stationary master equation (ME) in Theorem 7.1: the own-firm HJB terms and the population-transport term. The dependence of $\pi$ on $m$ (via $P(Y(m, x))$) is already handled inside the Hamiltonian and does not appear as a separate explicit term.

### Externality term in detail

Write $\pi(k,i,z,x,m) = \Psi(Y(m,x))\,\chi(k,z,x) - i - h(i,k) - f$ with $\Psi = P$ and $\chi = e^{x+z}k^{\alpha}$. Then

$$\delta\_m\pi(m)(\xi) = \Psi'(Y)\,\chi(k,z,x)\,\chi(\kappa,\zeta,x),$$

and integration wṙ.t. $m$ yields $\chi(k,z,x)\,\Psi'(Y)\,Y(m,x)$.

# B Appendix B: Residual-Loss Template (for implementation)

For a collocation tuple $(k,z,x)$, an empirical measure $m = \frac{1}{N}\sum_{n=1}^{N}\delta_{\xi^n}$, and parameterized $U_\omega, D\_mU_\psi$, define

$$\begin{aligned}
\widehat{Y} &\equiv \frac{1}{N}\sum_n = 1^N e^{x+\zeta^n}(\kappa^n)^{\alpha},\\
\widehat{\mathcal{R}}_{\mathrm{ME}} &\equiv r(x)\,U_\omega\\
&\qquad - \max_i \Big\{\pi + (U_\omega)_k\,(i-\delta k) + L_zU_\omega + L_xU_\omega\Big\}\\
&\qquad - \frac{1}{N}\sum_{n=1}^{N}\Big[(i^*(\xi^n,x,m) - \delta\kappa^n)\,\partial_\kappa D_mU_\psi(\xi^n) + \mu_z(\zeta^n)\,\partial_\zeta D_mU_\psi(\xi^n) + \tfrac{1}{2}\sigma_z^2\,\partial^2_{\zeta\zeta}D_mU_\psi(\xi^n)\Big]
\end{aligned}$$

Add soft KKT penalties (one-sided around $(U_\omega)_k = 1$), an anchoring penalty $\mathcal{P}_{\mathrm{anchor}} = \big(\int D\_mU\,,\mathrm{d}m\big)^2$ to fix the gauge, and boundary regularizers (reflecting $k=0$, growth at $k_{\max}$). Minimize

$$\mathcal{L} = \mathbb{E}\,[\widehat{\mathcal{R}}^2_{\mathrm{ME}}] + \lambda_{\mathrm{KKT}}\mathcal{P}_{\mathrm{KKT}} + \lambda_{\mathrm{bdry}}\mathcal{P}_{\mathrm{bdry}} + \lambda_{\mathrm{anchor}}\mathcal{P}_{\mathrm{anchor}}.$$

Anchoring removes the gauge freedom in $D\_mU$.

# C Appendix C: Common-Noise Master Equation (Reference Note)

When the population law $m_t$ itself diffuses under common noise (say through an exogenous $x_t$ or an aggregate Brownian component shared by firms), the functional Itô calculus on $\mathcal{P}_2$ introduces a second-order term in the measure variable. In a stylized form (see Carmona & Delarue, and Cardaliaguet–Delarue–Lasry–Lions), the stationary master equation would add a term of the form

$$\frac{1}{2}\,\Sigma_{\mathrm{com}} : \iint \partial_\xi\partial_{\xi'}\big(D\_mU(\xi)\big)\,\big(D\_mU(\xi')\big)\,m(,\mathrm{d}\xi)\,m(,\mathrm{d}\xi')$$

or, in classical PDE notation, $\frac{1}{2}\mathrm{Tr}\big[\Gamma\,\partial^2_{\xi\xi}D\_mU\big]$ integrated against $m$, where $\Gamma$ is the covariance of the common noise. Because this paper conditions on $x$, these terms are absent in the stationary master equation.

**Mathematical Insight: Rigor & Implications**

**Displacement monotonicity (context).** For master equations with common noise, uniqueness and well-posedness often require *displacement monotonicity*: a convexity-type condition along Wasserstein geodesics for the coupling. See [2] for precise statements. In our economic environment, couplings of the type $m \mapsto P(\int q\,,\mathrm{d}m)$ satisfy Lasry–Lions monotonicity when $P'(\cdot) < 0$ (Lemma 7.3), but displacement monotonicity may require additional curvature restrictions on $P$ or on the state-cost structure. Since we work conditional on $x$ (no common-noise measure diffusion), we do not invoke displacement monotonicity in this paper.

# D Appendix D: Tiny Pseudocode (Plain `listings`)

```
# Inputs:

# params\_omega: parameters for U(k,z,x; m)

# params\_psi: parameters for delta\_m U(xi; k,z,x; m)

# batch: list of tuples (k,z,x, {xi\_n=(kappa\_n,zeta\_n)}\_{n=1}^N )

# primitives: alpha, delta, mu\_z(z), sigma\_z, mu\_x(x), sigma\_x, r(x),

# demand P(Y), fixed cost f

# penalties: lambdas for KKT, boundary, and anchor (gauge) regularizers

def policy\_from\_grad(p, k, phi\_plus, phi\_minus):
\# p = U\_k (value gradient)
if p >= 1.0:
return (k/phi\_plus)*(p - 1.0)
else:
return (k/phi\_minus)*(p - 1.0)

def reflecting\_penalty(k, i\_star):
\# discourage negative control at k=0 and large negative flux
pen0 = max(0.0, -i\_star) if k<=1e-10 else 0.0
return pen0\*\*2

def h\_cost(i, k, phi\_plus, phi\_minus):
if i >= 0.0:
return 0.5*phi\_plus*(i*i)/max(k,1e-12)
else:
return 0.5*phi\_minus\*(i\*i)/max(k,1e-12)

def HJB\_operator(k,z,x,Yhat,Uk,Uz,Uzz,Ux,Uxx,i):
q = exp(x+z)*(k\*\*alpha)
pi = P(Yhat)*q - i - h\_cost(i,k,phi\_plus,phi\_minus) - f
return pi + Uk*(i - delta*k) + mu\_z(z)*Uz + 0.5*sigma\_z**2\*Uzz&#x20;
\+ mu\_x(x)*Ux + 0.5*sigma\_x**2\*Uxx

def ME\_residual\_for\_tuple(params\_omega, params\_psi, tup):
k,z,x,xi\_list = tup.k, tup.z, tup.x, tup.xi\_list
\# empirical measure moments
Y\_hat = mean(\[exp(x+xi.zeta)*(xi.kappa\*\*alpha) for xi in xi\_list])
\# U and its partials at (k,z,x)
U, Uk, Uz, Uzz, Ux, Uxx = U\_and\_grads(params\_omega, k,z,x, xi\_list)
```

```
\# best response i*
i\_star = policy\_from\_grad(Uk, k, phi\_plus, phi\_minus)
\# HJB maximand at i\*
H\_val = HJB\_operator(k,z,x,Y\_hat,Uk,Uz,Uzz,Ux,Uxx,i\_star)
\# Population transport term (via the Lions derivative)
integ = 0.0
for xi in xi\_list:
dU = delta\_mU\_and\_partials(params\_psi, xi, k,z,x, xi\_list)
\# dU returns dict with fields dkappa, dzeta, dzeta2, p\_k (proxy gradient)
i\_star\_xi = policy\_from\_grad(dU\['p\_k'], xi.kappa, phi\_plus, phi\_minus)
integ += (i\_star\_xi - delta*xi.kappa)* dU\['dkappa']&#x20;
\+ mu\_z(xi.zeta)\* dU\['dzeta']&#x20;
\+ 0.5*sigma\_z\*\*2 \* dU\['dzeta2']
integ = integ / len(xi\_list)
\# assemble residual (no separate externality term; price enters via P(Yhat) in HJB)
res = r(x)\*U - max(H\_val, HJB\_operator(k,z,x,Y\_hat,Uk,Uz,Uzz,Ux,Uxx,0.0))&#x20;\-
    integ
\# penalties
pen = reflecting\_penalty(k, i\_star)
return res, pen

def loss(params\_omega, params\_psi, batch):
sse = 0.0
pen = 0.0
for tup in batch:
res, p = ME\_residual\_for\_tuple(params\_omega, params\_psi, tup)
sse += res\*\*2
pen += p
anchor = anchor\_penalty(params\_psi, batch) # e.g., squared mean of dmU over batch
return sse/len(batch) + lambda\_bdry\*pen + lambda\_anchor\*anchor
```

Listing 7: Pseudo-JAX for (ME) residual with empirical measure

# E Appendix E: Symbolic Verification (PythonTeX + SymPy)

This appendix runs minimal SymPy checks to verify key derivations used in the text. Compilation is configured (via `latexmkrc`) to execute these checks on every build; any failure triggers a build error. We assume smoothness and reflecting/no-flux boundary conditions where noted.

Symbolic Check (SymPy)

```
import sympy as sp

# 1) Isoelastic simplification:  Y P'(Y) = -eta P(Y)
Y, eta = sp.symbols('Y eta', positive=True)
P = Y**(-eta)
check1 = sp.simplify(Y*sp.diff(P, Y) + eta*P)
assert check1 == 0
print("Isoelastic: Y*P'(Y) = -eta*P(Y)  [OK]")

# 2) Externality directional derivative:  d/d epsilon P(Y+epsilon*chi_eps)*chi0 |_{epsilon=0}
#     equals P'(Y) * chi0 * chi_eps
chi0, chieps, eps = sp.symbols('chi0 chieps eps', real=True)
```

```
Psi = lambda y: y**(-eta)
dpi = sp.diff(Psi(Y + eps*chieps)*chi0, eps).subs(eps, 0)
target = sp.diff(Psi(Y), Y) * chi0 * chieps
assert sp.simplify(dpi - target) == 0
print('Externality directional derivative  [OK]')

# 3) Externality, isoelastic reduction after integrating over m:  chi0 * Y * P'(Y) = -eta * P(Y) * chi
lhs = chi0 * Y * sp.diff(Psi(Y), Y)
rhs = -eta * Psi(Y) * chi0
assert sp.simplify(lhs - rhs) == 0
print('Externality isoelastic reduction   [OK]')

# 4) KKT/FOC solution for i* with asymmetric quadratic costs
#    h = 0.5*phi_plus*i^2/k for i>=0;   0.5*phi_minus*i^2/k for i<0
i, k, p, phi_plus, phi_minus = sp.symbols('i k p phi_plus phi_minus', positive=True)
h_plus  = 0.5*phi_plus*i**2/k
FOC_plus  = sp.Eq(sp.diff(-i - h_plus + p*i, i), 0)
sol_plus  = sp.solve(FOC_plus, i)[0]
h_minus = 0.5*phi_minus*i**2/k
FOC_minus = sp.Eq(sp.diff(-i - h_minus + p*i, i), 0)
sol_minus = sp.solve(FOC_minus, i)[0]
assert sp.simplify(sol_plus  - k*(p-1)/phi_plus)  == 0
assert sp.simplify(sol_minus - k*(p-1)/phi_minus) == 0
print('KKT/FOC piecewise i* formulas     [OK]')

# 5) FP adjoint pairing identity (algebraic, boundary terms omitted):
#    phi_k * (a*m) = d_k(phi*a*m) - phi * d_k(a*m)
kk = sp.symbols('kk', real=True)
phi = sp.Function('phi')(kk)
a   = sp.Function('a')(kk)
mm  = sp.Function('m')(kk)
expr = sp.diff(phi, kk)*(a*mm) - (sp.diff(phi*(a*mm), kk) - phi*sp.diff(a*mm, kk))
assert sp.simplify(expr) == 0
print('Adjoint pairing identity (no-flux) [OK]')

# 6) Envelope property for Hamiltonian in p: d/dp max_i { -i - h(i,k) + p i } = i*(p)
#    Check separately on each branch (ignoring terms not depending on i, e.g., -delta*k*p)
H_plus  = (-i - h_plus + p*i).subs(i, sol_plus)
H_minus = (-i - h_minus + p*i).subs(i, sol_minus)
dHp_dp  = sp.simplify(sp.diff(H_plus, p))
dHm_dp  = sp.simplify(sp.diff(H_minus, p))
assert sp.simplify(dHp_dp - sol_plus) == 0
assert sp.simplify(dHm_dp - sol_minus) == 0
print('Envelope: dH/dp equals i*(p)       [OK]')

print('\nAll SymPy verification checks passed.')
```

# F Appendix F: Lean4 Micro-Proofs (Sketches)

The following Lean4/mathlib4 snippets formalize two identities used in the text. They are provided as self-contained, runnable sketches (assuming a recent mathlib4): the isoelastic simplification

$Y\,P'(Y) = -\eta P(Y)$ and the algebraic reduction $Y \cdot Y^{-\eta-1} = Y^{-\eta}$ for $Y > 0$.

Formal Proof (Lean4)

```
import Mathlib.Analysis.Calculus.Deriv
import Mathlib.Data.Real.Basic

open Real

variable {eta Y : R}

-- P(Y) = Y ^ (-eta), defined for Y > 0 via rpow
def P (Y : R) (eta : R) : R := Y ^ (-eta)

-- Algebraic reduction: for Y > 0, Y * Y^(-eta - 1) = Y^(-eta)
theorem rpow_mul_cancel (hY : 0 < Y) :
    Y * Y ^ (-eta - 1) = Y ^ (-eta) := by
  -- rewrite Y as Y^1 and use rpow_add (valid for Y > 0)
  have h1 : Y = Y ^ (1 : R) := by simpa using (rpow_one Y)
  calc
    Y * Y ^ (-eta - 1)
        = Y ^ (1 : R) * Y ^ (-eta - 1) := by simpa [h1]
    _   = Y ^ ((1 : R) + (-eta - 1)) := by
          simpa using (rpow_mul_rpow_of_pos hY (1 : R) (-eta - 1))
    _   = Y ^ (-eta) := by ring

-- Differential identity: for Y > 0,  (Y) * (deriv (fun y => P y eta) Y) = -eta * P Y eta
theorem isoelastic_identity (hY : 0 < Y) :
    Y * (deriv (fun y => P y eta) Y) = -eta * P Y eta := by
  -- mathlib: d/dy (y^a) = a * y^(a-1) for y>0
  have hderiv : deriv (fun y => y ^ (-eta)) Y = (-eta) * Y ^ (-eta - 1) := by
    simpa using (deriv_rpow_const (x:=Y) (a:=-eta) hY.ne')
  -- multiply both sides by Y and reduce
  calc
    Y * (deriv (fun y => P y eta) Y)
        = Y * ((-eta) * Y ^ (-eta - 1)) := by simpa [P, hderiv]
    _   = -eta * (Y * Y ^ (-eta - 1)) := by ring
    _   = -eta * Y ^ (-eta) := by simpa using rpow_mul_cancel (eta:=eta) (Y:=Y) hY
    _   = -eta * P Y eta := by rfl
```

**Pedagogical Insight: Economic Intuition & Context**

**Notes.** The lemmas use rpow and standard calculus from `mathlib4`. They require $Y > 0$ for real-exponent laws. The SymPy checks in Appendix E independently validate the same identities numerically/symbolically.

Formal Proof (Lean4)

```
import Mathlib.Data.Real.Basic

-- Risk coefficient appearing in the EZ market price of risk
def risk_coeff (gamma psi : R) : R := (1 - gamma) * (1 - 1/psi)

-- If RRA = 1 (log utility), the utility-channel risk coefficient vanishes
```

```
@[simp] lemma risk_coeff_gamma_one (psi : R) : risk_coeff 1 psi = 0 := by
  simp [risk_coeff]

-- If EIS = 1, the utility-channel risk coefficient vanishes
@[simp] lemma risk_coeff_psi_one (gamma : R) : risk_coeff gamma 1 = 0 := by
  simp [risk_coeff]
```

# G Appendix G: Endogenous SDF with Epstein–Zin Aggregator

When the stochastic discount factor (SDF) is endogenous, it is often derived from a representative agent's preferences. Epstein–Zin (EZ) preferences allow separating the elasticity of intertemporal substitution (EIS) from relative risk aversion (RRA), which is crucial for asset pricing implications. This appendix details the continuous-time EZ aggregator used as a BSDE driver and derives the corresponding pricing kernel exposure.

**Definition G.1: Continuous-time Epstein–Zin Aggregator**

Fix time preference $\varrho > 0$ (using notation from Section 1), risk aversion $\gamma > 0$, and elasticity of intertemporal substitution $\psi > 0$. We assume $\psi \neq 1$. Let

$$\vartheta \equiv \frac{1-\gamma}{1-1/\psi}.$$

The parameter $\vartheta$ is sometimes used in alternative normalizations of the utility index.
For aggregate consumption $c_t > 0$, the representative agent's utility process $(V_t, Z_t)$ solves the backward SDE:

$$, \mathrm{d}V_t = -f(c_t, V_t, Z_t), \mathrm{d}t + Z_t^\top, \mathrm{d}B_t,$$

where $V_t > 0$ is the continuation value and $Z_t \in \mathbb{R}^{d_B}$ is the exposure vector to aggregate Brownian shocks $B_t$. The driver $f$ is the EZ aggregator. We use the standard normalization (consistent with Duffie–Epstein, 1992, [7], and also used in [8]):

$$f(c, V, Z) = \frac{\varrho}{1-1/\psi}\Big(c^{1-1/\psi}\, V^{1/\psi} - V\Big) \;+\; \tfrac{1}{2}\,(1-\gamma)\,(1-1/\psi)\,\frac{\|Z\|^2}{V}. \qquad \text{(G.1)}$$

**Pedagogical Insight: Economic Intuition & Context**

**Economic Intuition: Separating RRA and EIS.** EZ preferences break the link imposed by standard CRRA utility (where EIS $= 1/$RRA).

- $\gamma$ (RRA) controls aversion to static risk (gambles).
- $\psi$ (EIS) controls willingness to substitute consumption over time (smoothness).

The aggregator $f$ includes an intertemporal substitution term (first part) and a risk adjustment term (second part). The sign of the risk adjustment coefficient $(1-\gamma)(1-1/\psi)$ determines the preference for early ($\gamma > 1, \psi > 1$ or $\gamma < 1, \psi < 1$) or late resolution of uncertainty.

**Proposition G.1: Pricing kernel exposure under EZ**

Let $M_t$ denote the stochastic discount factor. The utility-channel diffusion component of the instantaneous market price of risk implied by Definition G.1 is

$$\Lambda_t^{\text{util}} \;=\; \partial_Z f(c_t, V_t, Z_t) \;=\; (1-\gamma)\,(1-1/\psi)\,\frac{Z_t}{V_t},$$

entering $\mathrm{d}M_t/M_t = -r_t\,\mathrm{d}t - (\Lambda_t)^\top\,\mathrm{d}B_t$. If consumption $c_t$ carries its own Brownian exposure, the total $\Lambda_t$ adds the consumption channel in the usual way.

*Proof.* The pricing kernel $M_t$ ensures that the utility process $V_t$, when discounted by $M_t$, is a martingale: $\mathbb{E}_t[M_{t+s}V_{t+s}] = M_tV_t$. Applying Itô's lemma to the product $M_tV_t$ gives

$$,\mathrm{d}(M_tV_t) = V_t\,,\mathrm{d}M_t + M_t\,,\mathrm{d}V_t + \langle,\mathrm{d}M_t,,\mathrm{d}V_t\rangle.$$

Substitute the dynamics

$$,\mathrm{d}M_t/M_t = -r_t\,,\mathrm{d}t - \Lambda_t^\top\,,\mathrm{d}B_t,$$
$$,\mathrm{d}V_t = -f(c_t,V_t,Z_t),\mathrm{d}t + Z_t^\top\,,\mathrm{d}B_t,$$

and note the quadratic covariation $\langle,\mathrm{d}M_t,,\mathrm{d}V_t\rangle = -M_t\,\Lambda_t^\top Z_t\,,\mathrm{d}t$. For $M_tV_t$ to be a martingale, the drift must vanish:

$$V_t(-M_tr_t) + M_t(-f(c_t,V_t,Z_t)) - M_t\,\Lambda_t^\top Z_t = 0,$$

yielding the generalized HJB identity $r_tV_t + f(c_t,V_t,Z_t) + \Lambda_t^\top Z_t = 0$. In equilibrium, the utility-channel component of the market price of risk is identified with the gradient of the aggregator with respect to $Z_t$ (see [7]). Differentiating Equation (G.1) w.r.t. $Z$ gives

$$\partial_Z f(c,V,Z) \;=\; \tfrac{1}{2}(1-\gamma)(1-1/\psi)\,\partial_Z\left(\frac{\|Z\|^2}{V}\right) \;=\; (1-\gamma)(1-1/\psi)\,\frac{Z}{V},$$

which proves the claim. □

**Symbolic Check (SymPy)**

```
import sympy as sp
# Verify the gradient calculation in the proof of Proposition G.1
# using an element-wise approach for robustness.
c, V, gamma, psi = sp.symbols('c V gamma psi', positive=True)
z1, z2 = sp.symbols('z1 z2', real=True)
Z = sp.Matrix([z1, z2])

# Coefficient in the standard normalization (Duffie--Epstein)
coeff_risk = (1-gamma)*(1-1/psi)

# Risk term: (1/2) * coeff * (Z^T Z) / V
risk_term = sp.Rational(1, 2) * coeff_risk * (Z.T * Z)[0, 0] / V

# Gradient with respect to Z's components
grad_Z = sp.Matrix(sp.derive_by_array(risk_term, [z1, z2]))
expected = coeff_risk * Z / V
```

```
# Robust check: ensure both components are exactly zero
difference = sp.simplify(grad_Z - expected)
assert all(e == 0 for e in difference)
```

**Mathematical Insight: Rigor & Implications**

**Integration with the Firm Problem.** When using an endogenous SDF, the firm's HJB equation (Equation (4.1)) incorporates the market price of risk $\Lambda_t$:

$$r_t V = \max_i \Big\{\pi + V_k (i - \delta k) + L\_zV + L\_xV - (\sigma_z V_z, \sigma_x V_x) \cdot \Lambda_t\Big\}.$$

If the aggregate shocks $x$ correspond to the Brownian motion $B_t$ driving EZ utility, this links firm valuation to representative-agent preferences via $\Lambda_t$.

**Pedagogical Insight: Economic Intuition & Context**

**Implementation hook.** The repository exposes a JAX-friendly generator for Equation (G.1) and a utility-channel SDF exposure helper:

```
bsde_dsgE/models/epstein_zin.py: EZParams, ez_generator, sdf_exposure_from_ez
bsde_dsgE/models/multicountry.py: preference="EZ" to enable the aggregator
```

Usage sketch in code:

```
from bsde_dsgE.models.epstein_zin import EZParams
from bsde_dsgE.models.multicountry import multicountry_model

params = EZParams(rho=0.02, gamma=10.0, psi=1.5)  # Use rho for time preference
problem = multicountry_model(dim=5, preference="EZ", ez_params=params)
```

The consumption mapping `c_fn(x)` can be provided by the user; by default, the model uses a positive aggregator from dividend-like states.

## References


[1] Carmona, R. and F. Delarue (2018). *Probabilistic Theory of Mean Field Games with Applications.* Springer.

[2] Cardaliaguet, P., F. Delarue, J.-M. Lasry, and P.-L. Lions (2019). *The Master Equation and the Convergence Problem in Mean Field Games.* Princeton University Press.

[3] Lasry, J.-M. and P.-L. Lions (2007). Mean field games. *Japanese Journal of Mathematics* 2(1): 229–260.

[4] Zaheer, M., S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, and A. Smola (2017). Deep Sets. *Advances in Neural Information Processing Systems (NeurIPS)* 30.

[5] Mou, C.-H. and J. Zhang (various years). Second-order master equations with common noise and displacement monotonicity. (Working papers / journal articles; see also related notes by Gangbo–Mészáros–Mou–Zhang.)

[6] Zhang, L. (2005). The value premium. *Journal of Finance* 60(1): 67–103.

[7] Duffie, D. and L. G. Epstein (1992). Stochastic Differential Utility. *Econometrica* 60(2): 353–394.

[8] Sauzet, M. (2023). Recursive preferences in continuous time and implications for asset pricing. (Working paper; aggregator normalisation consistent with Definition G.1.)

| Symbol | Type | Meaning |
|---|---|---|
| *States, Controls, and Shocks* | | |
| $k$ | state | Capital ($\geq 0$); reflecting boundary at $k = 0$ |
| $i$ | control | Net investment; $dk = (i - \delta k), \mathrm{d}t$ |
| $z$ | state | Idiosyncratic productivity; diffusion with generator $L\_z$ |
| $x$ | state | Aggregate (business-cycle) shock; generator $L\_x$ |
| $\sigma_z, \sigma_x$ | parameters | Diffusion volatilities of $z$ and $x$ |
| $\mu_z, \mu_x$ | functions | Drift coefficients in $L\_z, L\_x$ |
| $W, B$ | processes | Brownian motions for $z$ and $x$ (independent) |
| *Technology and Market Primitives* | | |
| $q(k, z, x)$ | output | $e^{x+z} k^{\alpha}$, $\alpha \in (0, 1)$ |
| $P(\cdot)$ | function | Inverse demand; $P' = P'(Y) < 0$ |
| $\alpha$ | parameter | Capital elasticity in production |
| $\delta$ | parameter | Depreciation rate |
| $h(i, k)$ | function | Irreversible adjustment cost (convex, asymmetric) |
| $\phi_{\pm}$ | parameters | Adjustment-cost curvatures for $i \gtrless 0$ |
| $f$ | parameter | Fixed operating cost |
| $\eta$ | parameter | Demand elasticity for isoelastic $P(Y) = Y^{-\eta}$ |
| $r(x)$ | function | Short rate (or constant $\rho$) under pricing measure |
| *Measure Theory and Operators* | | |
| $S$ | space | State space $\mathbb{R}_+ \times \mathbb{R}$ for $(k, z)$ |
| $m$ | measure | Cross-sectional law on $S$ |
| $\mathcal{P}_2(S)$ | space | Probability measures on $S$ with finite second moments |
| $\mathrm{W}_2$ | metric | Quadratic Wasserstein distance on $\mathcal{P}_2(S)$ |
| $\xi = (\kappa, \zeta)$ | point | Generic element in support of $m$ (a "marginal firm") |
| $D\_m$ | operator | Lions derivative operator (measure Fréchet derivative) |
| $D\_mU(\xi; k, z, x, m)$ | function | Lions derivative of $m \mapsto U(k, z, x, m)$ at $\xi$ |
| $L\_z, L\_x$ | operators | Generators in $z$ and $x$; $L\_z^*$ is the adjoint of $L\_z$ |
| $\mathcal{T}$ | operator | Transport operator acting on $D\_mU$ in (ME) |
| *Equilibrium Objects* | | |
| $\pi(\cdot)$ | function | Dividends $P(Y)e^{x+z}k^{\alpha} - i - h(i, k) - f$ |
| $Y(m, x)$ | scalar | Aggregate quantity $\int e^{x+z} k^{\alpha}\, m(, \mathrm{d}k, , \mathrm{d}z)$ |
| $V(k, z, x; m)$ | function | Stationary value function (HJB) |
| $U(k, z, x, m)$ | function | Master value function (ME) |
| $i^*(\cdot)$ | policy | Optimal net investment from HJB/KKT |
| $\bar{\imath}(k)$ | function | Lower bound on disinvestment (optional) |
| *Representative-agent block (endogenous SDF)* | | |
| $\gamma$ | parameter | Relative risk aversion (RRA) in Epstein–Zin preferences |
| $\psi$ | parameter | Elasticity of intertemporal substitution (EIS) |
| $\vartheta$ | parameter | Preference index $\vartheta = (1 - \gamma)/(1 - 1/\psi)$ |
| $\varrho$ | parameter | Subjective discount rate (avoids clash with depreciation $\delta$) |
| $M_t$ | process | Stochastic discount factor (pricing kernel) |
| $r_t$ | process | Real short rate implied by $M_t$ |
| $\Lambda_t$ | process | Market price of risk (Brownian exposure of $M_t$) |

Table 1: Notation used throughout.

| Component | Weight | Notes |
|---|---|---|
| ME residual MSE | 1.0 | $\mathcal{L}_{\mathrm{ME}} = \mathbb{E}[\mathcal{R}_{\mathrm{ME}}^2]$; primary objective. |
| Boundary penalty | $10^{-2}$ | Enforce reflecting boundary and admissibility constraints. |
| KKT penalty | $10^{-2}$ | Complementarity for investment constraints (if active). |
| Gauge anchor | $10^{-3}$ | Fix $\int D\_mU \,, \mathrm{d}m$ to remove invariance. |

Table 2: Route B loss components and typical starting weights. Tune per calibration and scale of residuals.

| Residual | Tight | Medium | Coarse |
|---|---|---|---|
| $\varepsilon_{\mathrm{ME}}$ | $10^{-5}$ | $10^{-4}$ | $10^{-3}$ |
| $\varepsilon_{\mathrm{HJB}}$ | $10^{-7}$ | $10^{-6}$ | $10^{-5}$ |
| $\varepsilon_{\mathrm{FP}}$ | $10^{-7}$ | $10^{-6}$ | $10^{-5}$ |

Table 3: Suggested tolerances (dimensionless; scale to data).