The Dynamics of Distributional Indices

Matthieu Gomez*

First version: June 2024
This version: March 2026
[latest version]

Abstract

Changes in distributional indices such as the Gini coefficient or the Herfindahl-Hirschman Index are hard to interpret: the same movement can reflect different underlying forces. This paper decomposes the change in any smooth distributional index into three terms: drift (systematic growth), dispersion (idiosyncratic reshuffling), and demography (entry and exit). Standard axioms — scale invariance, the transfer principle, population invariance — each restrict one of these terms. The framework unifies existing decompositions for weighted averages, variance, and top shares, and extends them to broader classes of indices. Applied to U.S. local banking, the decomposition reveals that modest net changes in concentration mask much larger offsetting gross forces.

Empirical researchers routinely summarize distributions with a single index: a quantile, a top share, the Gini coefficient, a generalized entropy measure, or the Herfindahl-Hirschman Index. Such statistics are useful because they compress an infinite-dimensional object into one number and often serve as sufficient statistics in economic models.[1] Yet their very parsimony makes dynamic interpretation difficult. A banking market whose Herfindahl index rises by 200 points over a decade could be one where the dominant bank steadily gains deposit share, one where the leader actually loses share while turbulent reshuffling among smaller banks raises concentration through a Jensen effect, or one where small competitors simply exit through acquisition. Each scenario implies different competitive forces and different policy responses, yet all produce the same scalar change. More broadly, the same net movement in any distributional index may reflect systematically faster growth of initially high-weight units, greater reshuffling among survivors, or changes in population composition through entry and exit.

This paper derives the local law of motion of any smooth distributional index. The key observation is that there are only three primitive local perturbations of a distribution that matter dynamically: adding or removing mass at a point, giving incumbent mass at a point a common proportional increase, and subjecting incumbent mass at a point to a local mean-preserving spread. Each perturbation loads on the index's influence function or its derivatives, and aggregating these three responses along a path with growth, risk, and turnover yields a

*Columbia University, mg3901@columbia.edu.

[1]Examples include the Herfindahl index in simple models of oligopoly, price dispersion in New-Keynesian models, and social welfare criteria of the form $\sum_i U(C_{it})$, which are closely related to Atkinson indices.

continuous-time decomposition into *drift*, *dispersion*, and *demography*, together with a discrete-time accounting framework for panel data.

The conceptual payoff is that classical axioms from inequality measurement become exact restrictions on these three objects. Scale invariance is equivalent to the statement that common proportional growth has zero local effect. The Pigou-Dalton transfer principle is equivalent to the statement that local mean-preserving spreads weakly raise the index. Population invariance is equivalent to the statement that representative replication has no demographic effect. The standard axioms therefore determine which dynamic forces survive and how they are signed.

The framework also clarifies what distinguishes canonical index families. Ordered weighted averages such as top shares and the Gini coefficient load on survivor growth through *rank*. Generalized entropy indices load on survivor growth and risk through *income level*, and use the same level weights for drift and dispersion up to a scalar coefficient. The Herfindahl-Hirschman Index shares the survivor weights of a quadratic entropy index but, because it is not population invariant, attaches independent weight to turnover in the number of units. The decomposition therefore provides a common language for comparing indices that are often studied separately.

The main empirical application studies U.S. local banking concentration over three decades. A moderate rise in HHI masks much larger offsetting gross forces: dominant incumbents lose deposit share on average, reshuffling among survivors raises concentration, and the exit of small banks through acquisition tips the balance. The same industry yields fundamentally different decompositions at the national versus county level, illustrating both the sensitivity of non-population-invariant indices to turnover and the importance of decomposing scalar concentration changes. A secondary illustration applies the framework to wealth inequality in Europe.

Related literature. A large empirical literature documents secular trends in distributional indices—top income shares (Piketty and Saez, 2003), wealth concentration (Saez and Zucman, 2016), market power (Autor et al., 2020)—yet the same net change in an index can reflect very different underlying forces, and existing frameworks for distinguishing them are specific to particular statistics.

The decomposition unifies and extends accounting frameworks for weighted averages (Melitz and Polanec, 2015), variance (Campbell et al., 2019), and top shares (Gomez, 2022). The organizing device is the influence function (Hampel, 1974), applied to inequality measurement by Cowell and Victoria-Feser (1996). The influence function is also the key tool in the distributional decomposition literature initiated by DiNardo et al. (1996) and Firpo et al. (2009); see Fortin et al. (2011) for a survey. That literature decomposes the effect of changes in characteristics on distributional indices; the present paper decomposes the effect of individual-level dynamics over time.

The results give dynamic content to the axiomatic theory of inequality measurement (Atkinson, 1970; Bourguignon, 1979; Shorrocks, 1980, 1984): standard properties of inequality indices map directly onto restrictions on each component of the decomposition. The decomposition also shows that the mobility measure of Genicot and Ray (2023) systematically understates upward mobility when individual income is stochastic, because it includes a negative dispersion term that only panel data can separate from genuine differential growth (Remark 1).

The continuous-time formulation complements random growth models of inequality (Gabaix et al., 2016; Benhabib et al., 2011; Achdou et al., 2022; Gomez and Gouin-Bonenfant, 2024), and in particular the Pareto exponent theory of Beare and Toda (2022); those papers derive stationary distributions, while the present paper tracks how distributional indices respond to the same forces along a transition path.

The rest of the paper proceeds as follows. Section 1 introduces the influence function, derives the continuous-time law of motion, and presents the discrete-time accounting framework. Section 2 shows that standard axioms restrict these three components. Section 3 computes the decomposition for canonical index families. Section 4 applies the framework to market concentration and wealth inequality.

1 Theory

1.1 Influence Functions

A distributional index is a mapping from a class of measures to the real line. Formally, let $X \subseteq \mathbb{R}$ denote the support and let $\nu : \mathcal{F} \to \mathbb{R}$, where $\mathcal{F}$ is a class of finite measures on X such that $|\nu(G)| < \infty$ for all $G \in \mathcal{F}$.

Following Hampel (1968) and Hampel (1974), I define the influence function of ν at G by adding mass at a point.

Definition 1 (Influence Function)**.** For a distributional index ν and a measure $G \in \mathcal{F}$, the *influence function* is:[2]

$$IF_\nu(x; G) \equiv \lim_{\epsilon \downarrow 0} \frac{\nu(G + \epsilon\delta_x) - \nu(G)}{\epsilon},$$

where δ_x denotes a point mass at x.

In words, $IF_\nu(x; G)$ measures the influence of adding an infinitesimal proportion of observations at x on the value of ν. Because any smooth perturbation of a measure can be decomposed into additions and removals of point masses, the influence function linearizes arbitrary distributional changes.

1.2 Law of Motion

Consider an economy populated by a continuum of agents indexed by i. Let Ω_t denote the set of agents alive at time t, and let x_{it} denote the wealth of agent $i \in \Omega_t$. Denote by G_t the

[2]The standard contamination definition perturbs $(1-\epsilon)G + \epsilon\delta_x = G + \epsilon(\delta_x - G)$, keeping total mass fixed. Denoting its influence function by IF^c_ν, linearity of the Gateaux derivative gives $IF^c_\nu(x; G) = IF_\nu(x; G) - \int_X IF_\nu(z; G)\, dG(z)$. The two definitions coincide when ν is population invariant ($\int_X IF_\nu\, dG = 0$); see Section 2.1. The add-mass definition is the relevant one here because entry and exit change the size of the measure.

cross-sectional measure of wealth at time t. For a quantity y, we write $\Delta y \equiv y(t+\Delta t) - y(t)$.

Assumption 1. Over a short interval Δt, the wealth of surviving agents changes by

$$\frac{\Delta x_{it}}{x_{it}} = \mu_t(x_{it})\,\Delta t + \sigma_t(x_{it})\,\sqrt{\Delta t}\,\varepsilon_{it}, \tag{1}$$

where $\mu_t(x)$ is the expected growth rate at wealth x, $\sigma_t(x)$ is the volatility of the growth rate, and ε_{it} is a mean-zero, unit-variance shock independent across agents. In addition, agents exit at rate δ_t with wealth distribution Ψ_{Dt}, and new agents enter at rate $\eta_t + \delta_t$ with wealth distribution Ψ_{Bt}, so that the net population growth rate is η_t.

The cross-sectional distribution moves through three primitive channels, each loading on a different order of the influence function:

$$\begin{aligned} \text{entry or exit at } x: &\quad IF_\nu(x; G_t), \\ \text{systematic growth at } x: &\quad x\,\partial_x IF_\nu(x; G_t) \cdot \mu_t(x)\,\Delta t, \\ \text{idiosyncratic shock at } x: &\quad \tfrac{1}{2}x^2\,\partial_{xx} IF_\nu(x; G_t) \cdot \sigma_t(x)^2\,\Delta t. \end{aligned}$$

Figure 1 illustrates the key idea: once we know the effect of adding a point mass at x — which is precisely $IF_\nu(x; G_t)$ — we know the effect of everything else by finite differences. Shifting mass from x to $x(1+\mu\,\Delta t)$ is a removal at x and an addition at $x(1+\mu\,\Delta t)$, so its effect is $IF_\nu(x(1+\mu\,\Delta t)) - IF_\nu(x)$, which is proportional to the first derivative. Splitting mass symmetrically from x to $x(1 \pm \sigma\sqrt{\Delta t})$ is a removal at x and two half-additions at $x(1 \pm \sigma\sqrt{\Delta t})$, so its effect is $\frac{1}{2}IF_\nu(x(1+\sigma\sqrt{\Delta t})) + \frac{1}{2}IF_\nu(x(1-\sigma\sqrt{\Delta t})) - IF_\nu(x)$, which is proportional to the second derivative.

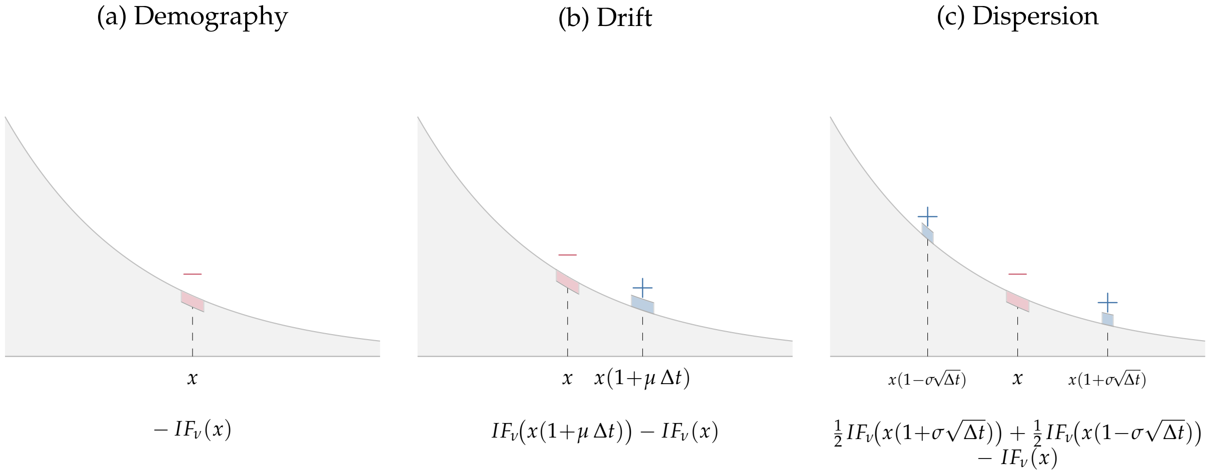


Figure 1: Three perturbations as combinations of adding and removing mass.

Notes: The influence function $IF_\nu(x)$ gives the effect of adding mass at x on any distributional index ν. Drift is the net effect of removing mass at x and adding it at $x(1+\mu\,\Delta t)$ (a local shift), which loads on the first derivative of IF_ν. Dispersion is the net effect of removing mass at x and splitting it symmetrically to $x(1 \pm \sigma\sqrt{\Delta t})$ (a local mean-preserving spread), which loads on the second derivative.

Proposition 1 (Local Law of Motion of a Distributional Index)**.** *Under Assumption 1, as* $\Delta t \to 0$,

$$\lim_{\Delta t \to 0} \frac{\Delta \nu(G_t)}{\Delta t} = \underbrace{\int_X x\, \partial_x IF_\nu(x; G_t)\, \mu_t(x)\, dG_t(x)}_{\textit{Drift}} + \underbrace{\frac{1}{2}\int_X x^2\, \partial_{xx} IF_\nu(x; G_t)\, \sigma_t(x)^2\, dG_t(x)}_{\textit{Dispersion}}$$
$$+ \underbrace{(\eta_t + \delta_t) \int_X IF_\nu(x; G_t)\, d\Psi_{Bt}(x) - \delta_t \int_X IF_\nu(x; G_t)\, d\Psi_{Dt}(x)}_{\textit{Demography}}. \tag{2}$$

Each term is a weighted average — of growth rates, volatilities, or entry/exit distributions — against a weighting function, or *kernel*, derived from the influence function. The *drift kernel* $x\, \partial_x IF_\nu(x; G_t)$ measures how the index responds to a proportional increase at x; the drift term is its covariance with growth rates $\mu_t(x)$. The *dispersion kernel* $\frac{1}{2} x^2\, \partial_{xx} IF_\nu(x; G_t)$ measures the curvature of the influence function; the dispersion term weights volatilities $\sigma_t(x)^2$ against it, so that mean-zero reshuffling raises the index whenever this kernel is positive. The *demographic kernel* is IF_ν itself, since entry and exit are additions and removals of mass. Section 2 shows that standard axioms restrict these three kernels, and Section 3 computes them for each canonical index family.

Figure 2 illustrates why the decomposition can matter: starting from the same distribution, the same increase in inequality can arise through three entirely different mechanisms — richer individuals growing faster (drift), similar individuals receiving different shocks (dispersion), or the poorest individual exiting (demography). A scalar change in the index cannot distinguish these forces; the decomposition can.

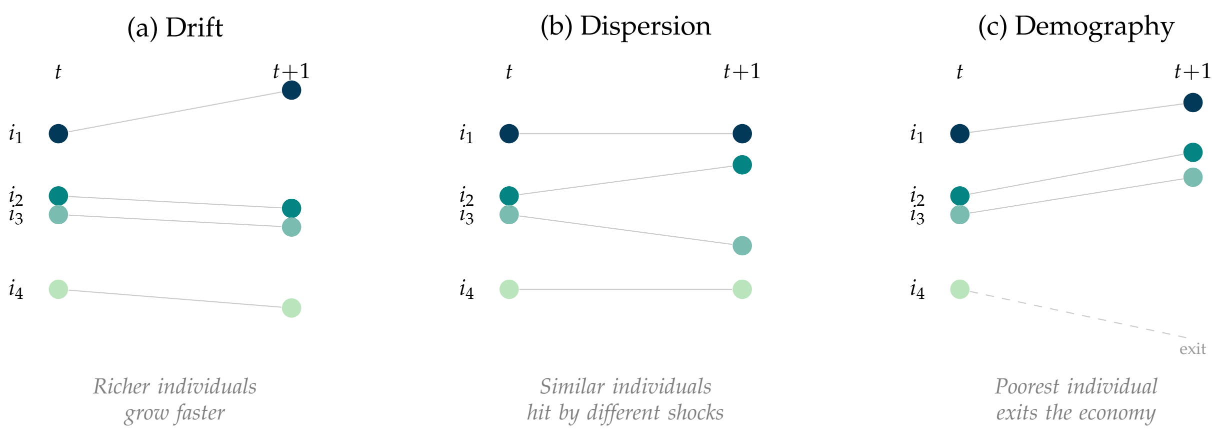


Figure 2: Three scenarios with similar increases in inequality.

Notes: Each dot represents an individual; height indicates income or wealth. All three panels start from the same distribution. Inequality rises in each case, but the mechanism differs. Panel (a): richer individuals grow faster, widening existing gaps (drift). Panel (b): individuals with similar initial positions receive different shocks, creating new dispersion (dispersion). Panel (c): the poorest individual exits and the distribution shifts upward (demography).

Economically, drift and dispersion capture two fundamentally distinct sources of inequality (Figure 3). Dispersion is the *origin* of inequality: even among agents with identical initial positions, idiosyncratic shocks create differences, so that inequality can rise without any systematic divergence. Drift determines whether existing inequality is perpetuated or eroded: it is positive when agents who are already richer grow systematically faster, and negative when growth is mean-reverting. The two forces are conceptually independent — an economy can have high dispersion but zero drift (pure reshuffling), or high drift but zero dispersion (deterministic divergence) — and their relative importance is an empirical question that requires panel data.

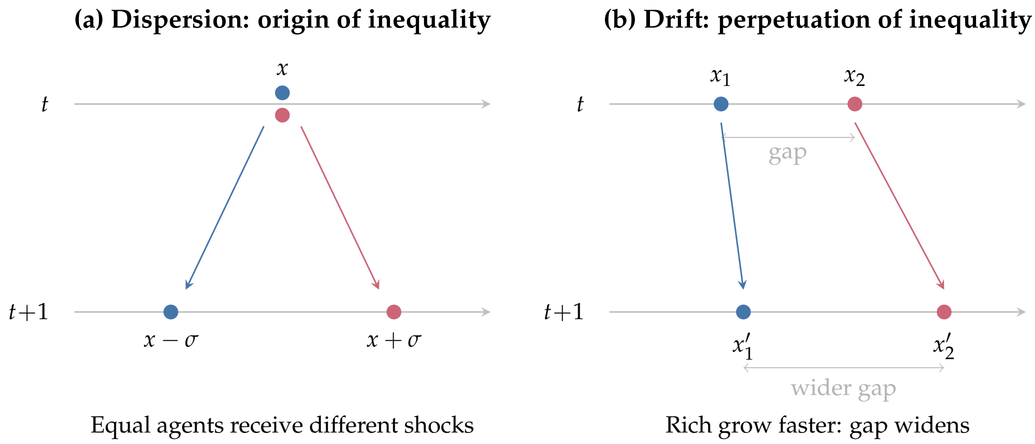


Figure 3: Two forces driving distributional change.

Notes: Panel (a) illustrates dispersion: two agents start with identical positions and receive different idiosyncratic shocks, creating inequality where none existed. Panel (b) illustrates drift: two agents start with different positions and the richer agent grows systematically faster, widening the initial gap.

The decomposition extends naturally to several settings discussed in the Appendix, including arbitrary continuous-time processes via a generator form, multiple variables, jump-diffusion processes, and weighted averages.

1.3 Discrete-Time Accounting

Proposition 1 identifies the three forces that move a distributional index: drift, dispersion, and demography. For panel data observed at discrete intervals, each force has a natural empirical counterpart that can be computed exactly. Let $S = \Omega_t \cap \Omega_{t+\Delta t}$ denote the set of survivors, and let $\widetilde{G}_t$ and $\widetilde{G}_{t+\Delta t}$ denote the distributions of survivor wealth at t and $t+\Delta t$.

Demography. The simplest term to compute is demography: it measures how the index changes when we add entrants or remove exiters. Comparing the full population with the survivor population at each endpoint gives

$$\text{Demography} = \left[\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t})\right] + \left[\nu(\widetilde{G}_t) - \nu(G_t)\right].$$

The first bracket is the effect of adding entrants at $t+\Delta t$; the second is the effect of removing exiters at t. Computing demography requires only evaluating the index on the full and survivor populations — no knowledge of the influence function is needed.

Drift. Among survivors, the change $\nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t)$ reflects both systematic growth and idiosyncratic reshuffling. To isolate the first-order effect of systematic growth, approximate using $\partial_x IF_\nu(x_{it}; \widetilde{G}_t)$:

$$\text{Drift} = \mathrm{E}\left[\partial_x IF_\nu(x_{it}; \widetilde{G}_t)\, \Delta x_{it} \;\middle|\; i \in S\right].$$

This values each survivor's observed change at the rate $\partial_x IF_\nu(x_{it}; \widetilde{G}_t)$, which measures how much a small income increase at x_{it} raises the index. Drift is positive when units with high $\partial_x IF_\nu$ grow faster. For each canonical index, this takes a simple closed form derived in Section 3.

Dispersion. The dispersion term is defined as the residual — the part of the actual survivor change not captured by the first-order approximation:

$$\text{Dispersion} = \nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t) - \text{Drift}.$$

For specific indices, the dispersion can also be computed in closed form (see Section 3), but the general definition is always this residual.

Proposition 2 (Discrete-Time Decomposition)**.** *The three terms above sum to the total change:*

$$\Delta\nu(G_t) = \textit{Drift} + \textit{Dispersion} + \textit{Demography}. \tag{3}$$

As $\Delta t \to 0$, each term is asymptotically equivalent to its counterpart in Proposition 1.

Remark 1 (Dispersion as a Distinct Economic Force)*.* When a distributional index rises, a natural interpretation is that growth has been systematically unequal. The decomposition shows this inference can be misleading: the dispersion term captures a separate force. In an economy with no differential growth ($\mu_t(x) = \mu$ for all x), the drift term has no effect on a scale-invariant index, yet the index still rises through the dispersion channel whenever incomes are stochastic. Separating drift from dispersion requires panel data. This distinction is relevant for the upward mobility measure of Genicot and Ray (2023), which equals the growth rate of Atkinson equivalent income; see Appendix A.4.

2 Properties of the Decomposition

Recall that each component of the decomposition is a weighted average against a kernel derived from the influence function: the demographic kernel $IF_\nu(x; G)$, the drift kernel $x\,\partial_x IF_\nu(x; G)$, and the dispersion kernel $\frac{1}{2}x^2\,\partial_{xx} IF_\nu(x; G)$. These are the zeroth, first, and second derivatives of the influence function. Each classical axiom restricts exactly one order of this hierarchy: population invariance restricts IF_ν itself (order 0), scale invariance restricts $\partial_x IF_\nu$ (order 1), and the transfer principle restricts $\partial_{xx} IF_\nu$ (order 2). Table 1 previews the results.

Table 1: Each classical axiom restricts one component of the decomposition.

Axiom	Component	Implication
Population invariance	Demography	Demography $= 0$ unless selection on entry and exit
Scale invariance	Drift	Drift $= 0$ unless differential growth across units
Transfer principle	Dispersion	Dispersion ≥ 0: reshuffling always raises the index

2.1 Population Invariance and Demography

The demography term compares the full population with the survivor population. The natural benchmark is *representative replication* — adding entrants drawn from the same distribution as incumbents — which should have no effect on an inequality measure, since doubling a population while preserving its shape does not change relative positions.

Definition 2 (Population Invariance)**.** ν is *population invariant* (Dalton, 1920) if $\nu(\lambda G) = \nu(G)$ for all $\lambda > 0$.

Proposition 3 (Population Invariance, Replication, and Demography)**.** *ν is population invariant if and only if $\int_X IF_\nu(x; G)\, dG(x) = 0$ for all G. Consequently, for a population-invariant index, representative entry and exit have zero demographic effect, and demography reduces to selective turnover: Demography $= (\eta_t + \delta_t) \int_X IF_\nu\, d\Psi_{Bt} - \delta_t \int_X IF_\nu\, d\Psi_{Dt}$.*

Proof. The first-order effect of representative replication is $\int_X IF_\nu\, dG$. This is zero for all G if and only if $\nu(\lambda G)$ is constant in λ. $\square$

The top share, the Gini, and generalized entropy are all population invariant: their demography term captures only *selective* entry and exit. The HHI is the leading non-population-invariant example: adding a representative firm mechanically lowers concentration through a $1/N$ dilution effect.

2.2 Scale Invariance and Drift

The drift term captures the first-order effect of systematic growth. In the discretized formulation, when every survivor grows at the same rate $\bar{\mu}$ (i.e., $\mu_t(x) = \bar{\mu}$ for all x), each income changes by $\Delta x_{it} = \bar{\mu}\, x_{it}\, \Delta t$. The question is whether this common scaling affects the index.

Definition 3 (Scale Invariance)**.** ν is *scale invariant* (Kolm, 1976) if $\nu(G_\lambda) = \nu(G)$ for all $\lambda > 0$, where G_λ is the distribution of λx when $x \sim G$.

Proposition 4 (Scale Invariance, Common Growth, and Drift)**.** *ν is scale invariant if and only if $\int_X x\, \partial_x IF_\nu(x; G)\, dG(x) = 0$ for all G. Consequently, for a scale-invariant index, common proportional growth has zero drift and, for any $\bar{\mu}_t$,*

$$\mathit{Drift} = \int_X x\, \partial_x IF_\nu(x; G_t) \cdot (\mu_t(x) - \bar{\mu}_t)\, dG_t(x).$$

Proof. Scaling all incomes by $1 + \epsilon$ changes each individual's income by ϵx. The local effect on ν is $\epsilon \int_X x\, \partial_x IF_\nu\, dG$, which is zero for all G if and only if $\nu(G_\lambda)$ is constant in λ. $\square$

For scale-invariant indices, only *differential* growth matters: an economy where everyone grows at 3% has zero drift, and the index moves only through the covariance between individual growth rates $\mu_t(x)$ and the drift kernel $x\,\partial_x IF_\nu(x; G_t)$.

2.3 Transfer Principle and Dispersion

The dispersion term captures the second-order effect of idiosyncratic shocks. In the discretized formulation, each survivor experiences a mean-zero shock $\sigma_t(x_{it})\sqrt{\Delta t}\,\varepsilon_{it}$. By a second-order (Jensen) effect, these shocks can raise or lower the index depending on the curvature of the influence function.

Definition 4 (Transfer Principle)**.** ν satisfies the *Pigou-Dalton transfer principle* (Dalton, 1920) if transferring $\tau > 0$ from a richer unit to a poorer one, without reversing their ranking, reduces the index.

Proposition 5 (Transfer Principle and Dispersion)**.** *ν satisfies the transfer principle if and only if $\partial_{xx} IF_\nu(x; G) \geq 0$ for all x and G. Consequently, Dispersion $= \frac{1}{2}\int_X x^2\,\partial_{xx} IF_\nu\,\sigma_t^2\,dG_t \geq 0$.*

Proof. A transfer from y to $x < y$ changes ν by approximately $\tau[\partial_x IF_\nu(x; G) - \partial_x IF_\nu(y; G)]$, which is negative for all $x < y$ if and only if $\partial_{xx} IF_\nu \geq 0$. □

The connection to Jensen's inequality is direct. The dispersion term is approximately the average convexity gap of the influence function:

$$\text{Dispersion} \approx \mathrm{E}\Big[IF_\nu(x_{it+\Delta t}; \widetilde{G}_t) - IF_\nu(x_{it}; \widetilde{G}_t) - \partial_x IF_\nu(x_{it}; \widetilde{G}_t)\,\Delta x_{it}\Big]\,.$$

For any function f, the quantity $f(y) - f(x) - f'(x)(y - x)$ is non-negative if and only if f is convex. When $IF_\nu(x; G)$ is convex ($\partial_{xx} IF_\nu(x; G) \geq 0$, i.e., the transfer principle holds), the dispersion is therefore non-negative: mean-zero shocks always raise the index.

3 Canonical Index Families

This section computes the decomposition for each canonical index family. The decomposition itself is always the same; what changes across indices is which parts of the distribution receive weight. Table 2 previews the closed-form drift and dispersion expressions; the rest of the section derives them. Throughout this section, G_t and $G_{t+\Delta t}$ denote the survivor distributions $\widetilde{G}_t$ and $\widetilde{G}_{t+\Delta t}$, $s_{it} = x_{it}/\mathrm{E}[x_{it}]$ denotes relative income among survivors, and R_{it} denotes CDF rank among survivors. The demography term is always $[\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t})] + [\nu(\widetilde{G}_t) - \nu(G_t)]$, as in Proposition 2.

The OWA and GE families are both population invariant, so ν depends only on the distribution of relative income $s_{it} = x_{it}/\mathrm{E}[x_{it}]$, while the HHI is the leading non-population-invariant case.

Table 2: Decomposition formulas for canonical indices.

Index		Decomposition of $\Delta\nu$	
Name	Formula	Drift	Dispersion
Top p share	$\frac{1}{p}\,\mathrm{E}\big[s_{it}\,\mathbf{1}_{R_{it}\geq 1-p}\big]$	$\frac{1}{p}\,\mathrm{E}\big[\Delta s_{it}\,\mathbf{1}_{R_{it}\geq 1-p}\big]$	$\frac{1}{p}\,\mathrm{E}\big[s_{it+\Delta t}\,\Delta\mathbf{1}_{R_{it}\geq 1-p}\big]$
Gini	$\mathrm{E}[(2R_{it}-1)\,s_{it}]$	$\mathrm{E}[(2R_{it}-1)\,\Delta s_{it}]$	$2\,\mathrm{E}[s_{it+\Delta t}\,\Delta R_{it}]$
OWA (general)	$\mathrm{E}[w(R_{it})\,s_{it}]$	$\mathrm{E}[w(R_{it})\,\Delta s_{it}]$	$\mathrm{E}[s_{it+\Delta t}\,\Delta w(R_{it})]$
HHI	$N\,\mathrm{E}\big[s_{it}^2\big]$	$2N\,\mathrm{E}[s_{it}\,\Delta s_{it}]$	$N\,\mathrm{E}\big[(\Delta s_{it})^2\big]$
GE(1)	$\mathrm{E}[s_{it}\,\ln s_{it}]$	$\mathrm{E}[\ln s_{it}\cdot\Delta s_{it}]$	$\mathrm{E}\left[s_{it+\Delta t}\,\ln\frac{s_{it+\Delta t}}{s_{it}}\right]$
GE(α)	$\frac{1}{\alpha(\alpha-1)}\,(\mathrm{E}[s_{it}^{\alpha}]-1)$	$\frac{1}{\alpha-1}\,\mathrm{E}\big[s_{it}^{\alpha-1}\,\Delta s_{it}\big]$	$\frac{1}{\alpha(\alpha-1)}\,\mathrm{E}\big[s_{it+\Delta t}^{\alpha}-s_{it}^{\alpha}-\alpha\,s_{it}^{\alpha-1}\Delta s_{it}\big]$

Notes: All averages are over survivors. For population-invariant indices, $s_{it}=x_{it}/\mathrm{E}[x_{it}]$ denotes relative income (with contemporaneous mean, so $\mathrm{E}[s_{it}]=1$ and $\mathrm{E}[\Delta s_{it}]=0$). R_{it} denotes CDF rank among survivors. For HHI, $s_{it}=x_{it}/\sum_j x_{jt}$ denotes market shares and N the number of survivors. The demography term is always $[\nu(G_{t+\Delta t})-\nu(\widetilde{G}_{t+\Delta t})]+[\nu(\widetilde{G}_t)-\nu(G_t)]$.

3.1 Ordered Weighted Averages

A natural first class is Ordered Weighted Averages (OWA), axiomatized by Mehran (1976) and Yaari (1988). They cover quantiles, top percentile shares, and the Gini coefficient.

Definition 5. Consider a weighting function $w:[0,1]\to\mathbb{R}^+$ such that $\int_0^1 w(r)\,dr=1$. The corresponding OWA is defined as:

$$\nu(G)=\mathrm{E}[w(R_{it})\,s_{it}]$$

where $s_{it}=x_{it}/\mathrm{E}[x_{it}]$ denotes relative income and $R_{it}=G_t(x_{it})$ denotes the CDF rank.

The top p share is $w(r)=\mathbf{1}_{r\geq 1-p}/p$. The standard Gini coefficient is the OWA with $w(r)=2r-1$.[3] OWA indices are scale invariant and population invariant; they satisfy the transfer principle if and only if $w(\cdot)$ is weakly increasing. Since $\mathrm{E}[\Delta s_{it}]=0$ (relative income has mean one at both dates), the discrete-time decomposition from Table 2 takes the form:

$$\nu(G_{t+\Delta t})-\nu(G_t)=\underbrace{\mathrm{E}[w(R_{it})\,\Delta s_{it}]}_{\text{Drift}}+\underbrace{\mathrm{E}[s_{it+\Delta t}\,\Delta w(R_{it})]}_{\text{Dispersion}}$$

The drift holds rank constant and weights income changes Δs_{it} by the rank weight $w(R_{it})$. The dispersion captures re-ranking: it weights changes in rank weights $\Delta w(R_{it})$ by final rela-

[3] The OWA with $w(r)=2r$ equals $1+\text{Gini}$. The constant is immaterial for the decomposition, which concerns changes.

tive income $s_{it+\Delta t}$. In continuous time, the drift kernel is $w(R)\,s$ and the dispersion kernel is $\frac{1}{2}w'(R)\,f(x)\,s^2$, where $f(x)$ is the density (see Appendix for derivation).

Top share. With $w(r) = \mathbf{1}_{r\geq 1-p}/p$, the drift weights only survivors initially above the threshold, and the dispersion captures individuals crossing in and out of the top group. This recovers the decomposition in Gomez (2022).

Gini coefficient. With $w(r) = 2r - 1$, the drift weight is $(2R_{it} - 1)$: individuals are weighted by rank, with those above the median contributing positively. The dispersion captures all rank changes, weighted by final income. Because the OWA is linear in w, and the Gini weight can be expressed as a weighted average of top-share weights, the Gini decomposition is the corresponding weighted average of top-share decompositions: $\nu_{\text{Gini}}(G) = \int_0^1 2p\,\nu_p(G)\,dp$, where ν_p denotes the top p share.

3.2 Generalized Entropy

The generalized entropy (GE) family provides the natural contrast to OWA indices. Characterized by Bourguignon (1979) and Shorrocks (1995), it is the unique class of additively decomposable distributional indices. It includes the Theil-L index ($\alpha = 0$), the Theil-T index ($\alpha = 1$), and is ordinally equivalent to the Atkinson index.[4]

Definition 6. The generalized entropy index with parameter α is:

$$\nu_\alpha(G) = \frac{\mathrm{E}[s_{it}^\alpha] - 1}{\alpha(\alpha - 1)}$$

GE indices are scale invariant, population invariant, and satisfy the transfer principle for all α, with influence function

$$IF_\nu(s; G) = \frac{\mathrm{E}^G[s^\alpha]}{\alpha(\alpha-1)}\left(\frac{s^\alpha}{\mathrm{E}^G[s^\alpha]} - 1 - \alpha\,(s-1)\right)$$

The discrete-time decomposition from Table 2 is:

$$\nu_\alpha(G_{t+\Delta t}) - \nu_\alpha(G_t) = \underbrace{\frac{1}{\alpha - 1}\,\mathrm{E}\Big[s_{it}^{\alpha-1}\,\Delta s_{it}\Big]}_{\text{Drift}} + \underbrace{\frac{1}{\alpha(\alpha-1)}\,\mathrm{E}\Big[s_{it+\Delta t}^\alpha - s_{it}^\alpha - \alpha\, s_{it}^{\alpha-1}\Delta s_{it}\Big]}_{\text{Dispersion}}$$

The drift captures the first-order effect of changes in relative income, weighted by $s_{it}^{\alpha-1}$, while the dispersion is the second-order (Taylor) residual of s^α. Since $\nu_2 = \text{Var}(s)/2$, the $\alpha = 2$ case reduces to half the variance decomposition of Campbell et al. (2019).

[4] The Atkinson index with parameter $\epsilon > 0$ is defined as

$$A_\epsilon(G) \equiv \begin{cases} 1 - \left(1 - \epsilon(1-\epsilon)\nu_{1-\epsilon}(G)\right)^{1/(1-\epsilon)} & \text{if } \epsilon > 0 \text{ and } \epsilon \neq 1 \\ 1 - e^{-\nu_{1-\epsilon}(G)} & \text{if } \epsilon = 1 \end{cases}$$

In contrast to OWA indices, both the drift and dispersion kernels weight by *income level* ($s^{\alpha-1}$) rather than rank. The parameter α controls which part of the distribution receives more weight: $\alpha < 1$ emphasizes the bottom, $\alpha = 1$ weights proportional to income, and $\alpha > 1$ emphasizes the top. In continuous time (see Appendix B.3), both kernels are proportional to s^α: the drift kernel is $\frac{1}{\alpha-1}s^\alpha$ and the dispersion kernel is $\frac{1}{2}s^\alpha$. This common-weighting property characterizes the GE family (see Appendix B.4).

4 Empirical Applications

This section illustrates the decomposition in two settings. The main application studies local banking concentration through the Herfindahl-Hirschman Index. A shorter secondary illustration studies wealth inequality through the Gini coefficient and top shares. In both settings, the same lesson emerges: small net changes in scalar indices often mask much larger offsetting gross forces.

4.1 Rising Concentration in the Banking Industry

The Herfindahl-Hirschman Index (HHI), denoted H, shows how the decomposition changes when population invariance fails. Letting $s_{it} = x_{it} / \sum_j x_{jt}$ denote market shares, $H = N\,\mathrm{E}[s_{it}^2]$. The HHI is scale invariant and satisfies the transfer principle, but unlike GE(2), it is *not* population invariant: adding firms drawn from the same distribution mechanically decreases concentration. The full decomposition, including entry and exit, is:

$$\Delta H = \underbrace{2N\,\mathrm{E}[s_{it}\,\Delta s_{it}]}_{\text{Drift}} + \underbrace{N\,\mathrm{E}\big[(\Delta s_{it})^2\big]}_{\text{Dispersion}} + \underbrace{\Big[H(G_{t+\Delta t}) - H(\widetilde{G}_{t+\Delta t})\Big] + \Big[H(\widetilde{G}_t) - H(G_t)\Big]}_{\text{Demography}}$$

The drift is positive when larger firms gain share, and the dispersion is always non-negative by the transfer principle. Unlike population-invariant indices, the demography term includes a mechanical dilution effect: even representative entry reduces HHI toward $1/N$ (see Appendix B.2).

Local Banking. I illustrate the HHI decomposition using U.S. local banking markets. I use the FDIC Summary of Deposits, which reports branch-level deposits for all FDIC-insured institutions annually from 1994 to 2024.[5] For each county-year pair $(t, t+1)$, I classify banks as survivors, entrants, and exiters, and compute the four terms of the HHI decomposition.

Local banking concentration has risen steadily from roughly 1,500 to 2,600 (county-level HHI, weighted by average deposits, 1994–2024). The central finding is that this moderate net increase masks much larger gross forces. *Drift* (cumulative $\approx -1{,}800$) is the dominant force *reducing* concentration: in the typical county, larger banks systematically lose deposit share to smaller survivors, a strong mean-reversion pattern consistent with smaller community banks

[5] Data from the FDIC BankFind Suite API (`https://api.fdic.gov/banks/sod`). Following the banking literature, I define markets as counties and compute each bank's local deposit share by summing branch deposits within each county.

and credit unions gaining local market share even as national consolidation proceeds. *Dispersion* (cumulative $\approx +1{,}900$) increases concentration, as predicted by the transfer principle; even mean-zero shocks to market shares raise HHI through a Jensen's inequality effect. *Demography* (cumulative $\approx +1{,}050$) also increases concentration, driven mainly by the exit of small banks through acquisition: when a bank is acquired, the acquirer's CERT persists while the target's CERT disappears from the data, concentrating deposits among fewer competitors.

Table 3 contrasts the cumulative decomposition at two levels of aggregation: the national aggregate (treating all banks as competing in a single market) and the deposit-weighted average across county-level markets. The two rows tell strikingly different stories. At the national level, HHI rose from 28 to 388, driven almost entirely by demography (+370): the number of FDIC-insured banks fell from roughly 13,000 to 4,500, primarily through mergers and acquisitions. Among surviving banks, drift is negative (-136) and dispersion modest (+126). At the county level, the increase is larger (+1,148) and the gross forces much larger still: drift ($-1{,}834$) and dispersion (+1,913) nearly cancel, with demography (+1,069) tipping the balance. The same industry yields fundamentally different decompositions depending on the market definition, illustrating both the sensitivity of the HHI to the number of competitors (non-population invariance) and the importance of decomposing scalar concentration changes into their primitive forces.

Table 3: HHI decomposition for U.S. banking, 1994–2024.

	ΔH	Drift	Dispersion	Demography
National aggregate	360	-136	126	370
Avg. county market	1148	-1834	1913	1069

Notes: All values on the HHI $\times$ 10,000 scale. "National aggregate" treats all banks as competing in a single market. "Avg. county market" is the deposit-weighted average of county-level decompositions. Source: FDIC Summary of Deposits.

The negative drift is the most striking finding. One might expect consolidation to produce a positive drift, with large acquirers gaining share among survivors. While acquisitions do contribute positively to drift, this is more than offset by organic mean reversion: in the typical county, the largest bank loses deposit share year over year. The rise in concentration therefore comes from dispersion and demography, not from a systematic tendency of dominant incumbents to grow.

Table 4 decomposes the cross-county variance of ΔH into the contributions of each force, weighted by average deposits. Since $\Delta H = \text{Drift} + \text{Dispersion} + \text{Demography}$ exactly, the variance decomposes as

$$\text{var}(\Delta H) = \text{cov}(\text{Drift}, \Delta H) + \text{cov}(\text{Dispersion}, \Delta H) + \text{cov}(\text{Demography}, \Delta H),$$

and the three covariance shares sum to 100%. Dispersion accounts for 80% and demography for 55%, while drift works in the opposite direction (-35%). Table 5 reports the weighted correlation matrix: drift and dispersion are strongly negatively correlated (-0.84), indicating that counties with more reshuffling also experience more mean reversion.

Table 4: Variance decomposition of ΔH across counties.

Share of var(ΔH) explained by		
Drift	Dispersion	Demography
−34.8%	80.1%	54.7%

Notes: 3,229 counties, weighted by average of 1994 and 2024 deposits. Since Drift + Dispersion + Demography = ΔH exactly, the three covariance shares sum to 100%.

Table 5: Weighted correlation matrix of HHI components across counties.

	Drift	Dispersion	Demography
Drift	1.00		
Dispersion	−0.84	1.00	
Demography	−0.55	0.42	1.00

Notes: 3,229 counties, weighted by average of 1994 and 2024 deposits.

The local banking application serves as the paper's main empirical example because it highlights the role of non-population invariance: unlike indices of income or wealth inequality, concentration measures are sensitive to the number of competitors, and the decomposition makes this sensitivity explicit through the demographic dilution term.

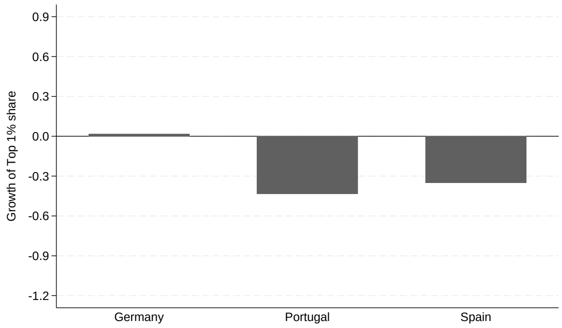
4.2 Changes in Inequality across European Countries

As a secondary illustration, I apply the framework to wealth inequality in the Household Finance and Consumption Survey (HFCS), a harmonized household wealth survey coordinated by the European Central Bank. The HFCS includes a panel component in which the same households are re-interviewed across waves. I use the panel linking waves 3 ($\sim$2017) and 4 ($\sim$2020–21) for three countries with large panel samples: Germany (2,430 panel households), Spain (4,317), and Portugal (2,716).[6]

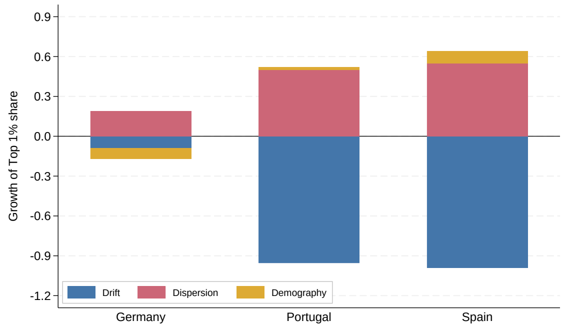
To apply the top-share decomposition, I restrict attention to households whose reference person is aged 20–80 in both waves. Households aging into the sample are treated as births, while those aging out are treated as deaths. In each wave, I normalize net wealth by the cross-sectional mean, so that the top p share equals $\mathrm{E}[x/\bar{x} \mid R \geq 1-p]/p$. I then apply the top-share decomposition of Gomez (2022) at each $p \in \{0.01, 0.02, \ldots, 0.99\}$.

The same broad pattern emerges as in concentration. Figure 4 shows the total change and its decomposition for three indices: the top 1% share, the top 10% share, and the Gini coefficient. In each case, drift is negative, indicating mean reversion in relative wealth, while dispersion is positive, indicating that reshuffling raises inequality. Demography is comparatively small over this short horizon. Net changes in all three indices are modest relative to the underlying drift and dispersion terms, again showing that scalar distributional changes can conceal large offsetting forces.

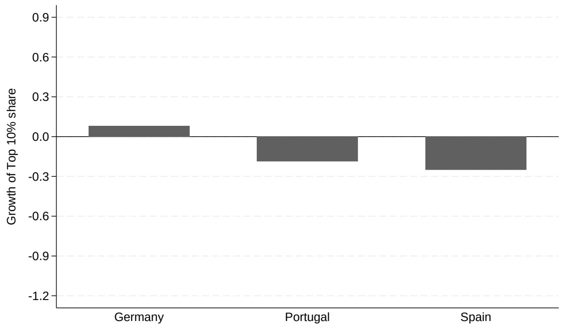
[6] The HFCS panel component is not available for France, Belgium, Finland, and several other countries. Among countries with panel data, I restrict to those with at least 2,000 panel households to ensure reliable percentile-level estimates. Net wealth is multiply imputed in five implicates, and all results average across implicates.



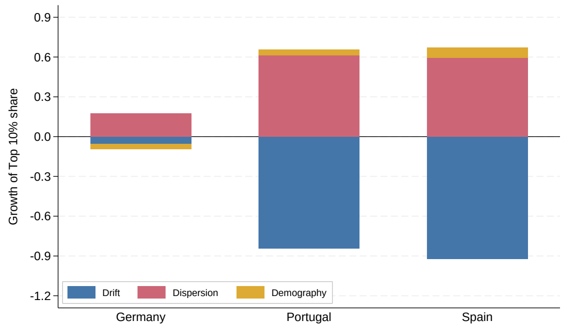
(a) Top 1%: total change



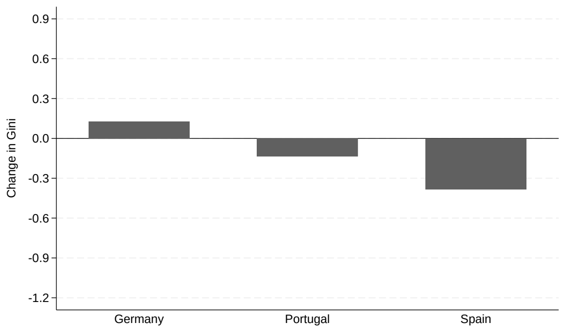
(b) Top 1%: decomposition



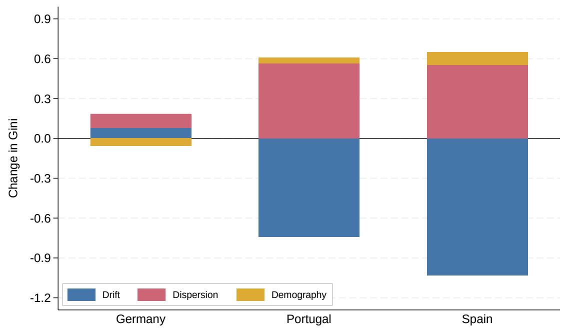
(c) Top 10%: total change



(d) Top 10%: decomposition



(e) Gini: total change



(f) Gini: decomposition

Figure 4: Wealth inequality decomposition by country.

Notes: HFCS wave 3 (∼2017) to wave 4 (∼2020–21). Left panels show the total change; right panels decompose the change into drift, dispersion, and demography. The Gini decomposition is obtained by integrating the top-share decomposition with weights $2p$. Net wealth is normalized by the cross-sectional mean in each wave. Results average across five multiply imputed datasets.

Table 6: Gini decomposition by country.

Country	Total	Drift	Dispersion	Demography
Germany	0.128	0.081	0.105	−0.057
Spain	−0.401	−1.082	0.580	0.101
Portugal	−0.133	−0.725	0.552	0.041

Notes: HFCS wave 3 to wave 4. Values are obtained by integrating the top-share decomposition with Gini weights $2p$ over $p \in (0, 1)$. The total equals the sum of the three components.

5 Conclusion

Distributional indices remain indispensable because they summarize complex distributions parsimoniously. But that same parsimony obscures the forces that move them over time. This paper provides a general decomposition of the change in any smooth distributional index into drift, dispersion, and demography, and shows that these three terms are governed by the index's influence function and its derivatives.

The decomposition yields two main payoffs. The first is conceptual. Classical axioms from inequality measurement have direct dynamic content: scale invariance removes the common-growth component of drift, the transfer principle signs the dispersion term, and population invariance removes the neutral-growth component of demography. The second is comparative. Different indices differ because they place different sensitivity weights on drift, dispersion, and demography. Ordered weighted averages emphasize rank, generalized entropy indices emphasize income level, and non-population-invariant indices such as the Herfindahl-Hirschman Index attach independent importance to turnover in the number of units.

The empirical applications illustrate why this matters. In local banking concentration, small net movements in HHI hide much larger offsetting gross forces, with negative drift and positive dispersion nearly canceling over long periods and demography determining the sign of the net change. The shorter wealth illustration shows that the same logic extends beyond concentration to more familiar inequality measures. More broadly, the framework preserves the tractability of scalar indices while recovering part of the dynamic information lost when an entire distribution is compressed into a single number. For empirical work, this suggests that the relevant dynamic sufficient statistics are often not one distributional index, but the three forces that move it.

A Appendix for Section 1

A.1 Proofs

Proof of Proposition 1. Assume that G_t has a smooth density $f_t(x)$. Writing $\partial_t G_t$ as a superposition of point masses and using linearity of the Gateaux derivative gives:

$$\frac{d}{dt}\nu(G_t) = \int_X IF_\nu(x; G_t)\,\partial_t f_t(x)\,dx \tag{4}$$

The Kolmogorov forward equation associated with the process (1) gives:

$$\partial_t f_t(x) = -\partial_x(\mu_t(x)\,x\,f_t(x)) + \frac{1}{2}\partial_{xx}(\sigma_t^2(x)\,x^2\,f_t(x)) + (\eta_t + \delta_t)\psi_{Bt}(x) - \delta_t\psi_{Dt}(x)$$

where ψ_{Bt} and ψ_{Dt} are the densities of Ψ_{Bt} and Ψ_{Dt}. Substituting into (4) and integrating the first two terms by parts (once and twice, respectively) yields the drift and dispersion terms. The demography term arises directly from the $(\eta_t + \delta_t)\psi_{Bt} - \delta_t\psi_{Dt}$ contribution. □

Proof of Proposition 2. Part 1: The identity. Since dispersion is defined as $\nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t) -$ Drift, the sum Drift + Dispersion $= \nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t)$ by construction. Adding demography:

$$\begin{aligned}\text{Drift} + \text{Dispersion} + \text{Demography} &= \left[\nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t)\right] + \left[\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t})\right] + \left[\nu(\widetilde{G}_t) - \nu(G_t)\right] \\ &= \nu(G_{t+\Delta t}) - \nu(G_t).\end{aligned}$$

Part 2: Convergence to Proposition 1. Under Assumption 1, $\Delta x_{it} = x_{it}\mu_t(x_{it})\Delta t + x_{it}\sigma_t(x_{it})\sqrt{\Delta t}\,\varepsilon_{it}$. A second-order Taylor expansion of $IF_\nu(\cdot\,; \widetilde{G}_t)$ around each survivor's initial income gives:

$$\nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t) = \mathrm{E}\left[\partial_x IF_\nu(x_{it}; \widetilde{G}_t)\,\Delta x_{it} + \tfrac{1}{2}\,\partial_{xx} IF_\nu(x_{it}; \widetilde{G}_t)\,(\Delta x_{it})^2 \;\middle|\; i \in S\right] + o(\Delta t).$$

The discrete drift is the first term. Since ε_{it} is mean-zero and independent of x_{it}:

$$\frac{\text{Drift}}{\Delta t} = \mathrm{E}\left[x_{it}\,\partial_x IF_\nu(x_{it}; \widetilde{G}_t)\,\mu_t(x_{it}) \;\middle|\; i \in S\right] + O(\sqrt{\Delta t}) \xrightarrow{\Delta t \to 0} \int_X x\,\partial_x IF_\nu\,\mu_t\,dG_t.$$

For the dispersion, using $\mathrm{E}[(\Delta x_{it})^2 \mid x_{it}] = x_{it}^2\sigma_t(x_{it})^2\Delta t + O(\Delta t^2)$:

$$\frac{\text{Dispersion}}{\Delta t} \xrightarrow{\Delta t \to 0} \frac{1}{2}\int_X x^2\,\partial_{xx} IF_\nu\,\sigma_t^2\,dG_t.$$

For demography, recall Demography $= [\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t})] + [\nu(\widetilde{G}_t) - \nu(G_t)]$. Over the interval Δt, a mass $\delta_t\Delta t$ exits with distribution Ψ_{Dt} and a mass $(\eta_t + \delta_t)\Delta t$ enters with distribution Ψ_{Bt}. The survivor distribution at t satisfies $\widetilde{G}_t \approx G_t - \delta_t\Delta t\,\Psi_{Dt}$, so by the definition of the influence function:

$$\nu(\widetilde{G}_t) - \nu(G_t) \approx -\delta_t\Delta t \int_X IF_\nu(x; G_t)\,d\Psi_{Dt}(x).$$

Similarly, $G_{t+\Delta t} \approx \widetilde{G}_{t+\Delta t} + (\eta_t + \delta_t)\Delta t\, \Psi_{Bt}$, so:

$$\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t}) \approx (\eta_t + \delta_t)\Delta t \int_X IF_\nu(x; G_t)\, d\Psi_{Bt}(x).$$

Combining:

$$\frac{\text{Demography}}{\Delta t} \xrightarrow{\Delta t \to 0} (\eta_t + \delta_t) \int_X IF_\nu\, d\Psi_{Bt} - \delta_t \int_X IF_\nu\, d\Psi_{Dt} \square$$

A.2 General Continuous-Time Processes and Jump-Diffusions

The diffusion specification in Assumption 1 is convenient but not essential. The same decomposition applies to time-inhomogeneous Markov survivor processes for which an extended generator exists and the local first moment is well defined.

Let X_t denote the state of a surviving unit. For smooth scalar test functions φ, define the extended generator by

$$(A_t\varphi)(x) \equiv \lim_{h \downarrow 0} \frac{1}{h} \mathrm{E}[\varphi(X_{t+h}) - \varphi(X_t) \mid X_t = x],$$

whenever the limit exists. Define the local drift coefficient by

$$b_t(x) \equiv \lim_{h \downarrow 0} \frac{1}{h} \mathrm{E}[X_{t+h} - X_t \mid X_t = x],$$

and let

$$(A_t^{\text{drift}}\varphi)(x) \equiv b_t(x)\varphi'(x), \qquad A_t^{\text{dispersion}} \equiv A_t - A_t^{\text{drift}}.$$

Thus A_t^{drift} captures the predictable first-order displacement of surviving units, while $A_t^{\text{dispersion}}$ is the residual mean-zero reshuffling / risk component after subtracting that first-order effect.

Two cases are useful.

Case 1: Diffusion. For the diffusion model in the main text,

$$\frac{dX_t}{X_t} = \mu_t(X_t)dt + \sigma_t(X_t)dW_t,$$

we have

$$(A_t^{\text{drift}}\varphi)(x) = x\mu_t(x)\varphi'(x),$$
$$(A_t^{\text{dispersion}}\varphi)(x) = \frac{1}{2}x^2\sigma_t(x)^2\varphi''(x).$$

Case 2: Compensated jump-diffusion. For a compensated jump-diffusion specification,

$$\frac{dX_t}{X_{t-}} = \mu_t(X_t)dt + \sigma_t(X_t)dW_t + \left(e^{\phi_t(X_t)U} - 1\right) dN_t - \mathrm{E}_U\left[e^{\phi_t(X_t)U} - 1\right] \lambda\, dt,$$

where N_t is a Poisson process with intensity λ, U is an i.i.d. bounded jump shock drawn at each jump, and the jump shocks are independent of W_t and N_t, we have

$$\begin{aligned}
\left(A_t^{\text{drift}}\varphi\right)(x) &= x\mu_t(x)\varphi'(x), \\
\left(A_t^{\text{dispersion}}\varphi\right)(x) &= \frac{1}{2}x^2\sigma_t(x)^2\varphi''(x) \\
&\quad + \lambda\, \mathrm{E}_U\left[\varphi(xe^{\phi_t(x)U}) - \varphi(x) - x\left(e^{\phi_t(x)U} - 1\right)\varphi'(x)\right].
\end{aligned}$$

Proposition 6 (Markov Continuous-Time Local Law of Motion)**.** *Suppose survivor dynamics are governed by the extended generator A_t defined above. Then for any smooth distributional index ν,*

$$\begin{aligned}
\frac{d}{dt}\nu(G_t) &= \int_X \left(A_t^{\text{drift}} IF_\nu(\cdot; G_t)\right)(x)\, dG_t(x) \\
&\quad + \int_X \left(A_t^{\text{dispersion}} IF_\nu(\cdot; G_t)\right)(x)\, dG_t(x) \\
&\quad + (\eta_t + \delta_t)\int_X IF_\nu(x; G_t)\, d\Psi_{Bt}(x) - \delta_t \int_X IF_\nu(x; G_t)\, d\Psi_{Dt}(x).
\end{aligned}$$

Proof. Apply the extended generator to the scalar test function $x \mapsto IF_\nu(x; G_t)$, integrate over the survivor distribution, and add the demographic mass-flow term. The decomposition into drift and dispersion follows from writing $A_t = A_t^{\text{drift}} + A_t^{\text{dispersion}}$. Because the paper works with the continuum cross-sectional law G_t rather than a random empirical measure, the resulting law of motion is deterministic. □

In the case of the compensated jump-diffusion process above, Proposition 6 becomes

$$\begin{aligned}
\frac{d}{dt}\nu(G_t) &= \underbrace{\int_X x\,\partial_x IF_\nu(x; G_t)\,\mu_t(x)\, dG_t(x)}_{\text{Drift}} + \underbrace{\frac{1}{2}\int_X x^2\,\partial_{xx} IF_\nu(x; G_t)\,\sigma_t(x)^2\, dG_t(x)}_{\text{Dispersion (diffusion)}} \\
&\quad + \underbrace{\lambda\int_X \mathrm{E}_U\left[IF_\nu(x\,e^{\phi_t(x)U}; G_t) - IF_\nu(x; G_t) - x\left(e^{\phi_t(x)U} - 1\right)\partial_x IF_\nu(x; G_t)\right] dG_t(x)}_{\text{Dispersion (jumps)}}.
\end{aligned}$$

The drift term is unchanged from the diffusion case, since the jump process is compensated. The jump dispersion term captures the nonlinear effect of discrete jumps beyond the first-order displacement already incorporated in the drift. When ν satisfies the transfer principle ($\partial_{xx} IF_\nu \geq 0$, i.e., IF_ν is convex), the jump dispersion term is non-negative by Jensen's inequality.

A.3 Multivariate Distribution

We now generalize to multiple variables as well as multiple sources of idiosyncratic risk (Brownian motions). For generality, we use the additive form for the multivariate process, since different variables may have different natural scales. Denote $\boldsymbol{x} = [x_1, \ldots, x_n]$ and G the multivariate distribution. We define ν as the distributional index. The influence function generalizes directly:

$$IF_\nu(\boldsymbol{x};G) \equiv \lim_{\epsilon\downarrow 0} \frac{\nu\left(G+\epsilon\,\delta_{\boldsymbol{x}}\right)-\nu(G)}{\epsilon}$$

where $\delta_{\boldsymbol{x}}$ denotes a point mass at $\boldsymbol{x}$.

Assume

$$d\boldsymbol{x}_t = \boldsymbol{\mu}_t(\boldsymbol{x}_t)dt + \boldsymbol{\Sigma}_t(\boldsymbol{x}_t)d\boldsymbol{W}_{it}$$

for a n-dimensional vector $\boldsymbol{\mu}_t$, a $n\times n$ matrix $\boldsymbol{\Sigma}_t$ and n-dimensional vector of standard Brownian Motion $\boldsymbol{W}_{it}$.

Proposition 7.

$$\frac{d}{dt}\nu(G_t) = \underbrace{\int_X (\nabla IF_\nu)'\boldsymbol{\mu}_t\,dG_t}_{\textit{Drift}} + \underbrace{\frac{1}{2}\int_X \mathrm{Tr}\left(\boldsymbol{\Sigma}_t'\left(\mathrm{Hess}\,IF_\nu\right)\boldsymbol{\Sigma}_t\right)dG_t}_{\textit{Dispersion}}$$

where ∇IF_ν denotes the Jacobian of IF_ν, Hess IF_ν *denotes its Hessian, and* Tr *denotes the trace operator.*

Proof. The multivariate Kolmogorov forward equation for the process $d\boldsymbol{x}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\Sigma}_t d\boldsymbol{W}_{it}$ is:

$$\partial_t g_t(\boldsymbol{x}) = -\sum_j \partial_{x_j}(\mu_{jt}\,g_t) + \frac{1}{2}\sum_{j,k}\partial_{x_j x_k}\left([\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t']_{jk}\,g_t\right)$$

Substituting and integrating by parts—once for the drift terms and twice for the diffusion terms—yields:

$$\frac{d}{dt}\nu(G_t) = \int_X \sum_j \partial_{x_j} IF_\nu \cdot \mu_{jt}\,dG_t + \frac{1}{2}\int_X \sum_{j,k}\partial_{x_j x_k} IF_\nu \cdot [\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t']_{jk}\,dG_t$$

which in matrix notation is $\int_X (\nabla IF_\nu)'\boldsymbol{\mu}_t\,dG_t + \frac{1}{2}\int_X \mathrm{Tr}(\boldsymbol{\Sigma}_t'(\mathrm{Hess}\,IF_\nu)\boldsymbol{\Sigma}_t)\,dG_t$. □

To fix ideas, consider the case of a two-dimensional vector, $\boldsymbol{x} = [x_1, x_2]$:

$$\begin{aligned} dx_{1it} &= \mu_{1t}(x_{1it}, x_{2it})dt + \sigma_{1t}(x_{1it}, x_{2it})dZ_{1it} \\ dx_{2it} &= \mu_{2t}(x_{1it}, x_{2it})dt + \sigma_{2t}(x_{1it}, x_{2it})dZ_{2it} \end{aligned}$$

with $\mathrm{corr}(dZ_{1it}, dZ_{2it}) = \rho_t(x_{1it}, x_{2it})$, which gives:

$$\frac{d}{dt}\nu(G_t) = \underbrace{\int_X \left(\partial_{x_1} IF_\nu \mu_{1t} + \partial_{x_2} IF_\nu \mu_{2t}\right) dG_t}_{\text{Drift}} + \underbrace{\frac{1}{2}\int_X \left(\partial^2_{x_1} IF_\nu \sigma^2_{1t} + \partial^2_{x_2} IF_\nu \sigma^2_{2t} + 2\partial_{x_1 x_2} IF_\nu \rho_t \sigma_{1t}\sigma_{2t}\right) dG_t}_{\text{Dispersion}}$$

A.4 Relation to upward mobility

The Atkinson equivalent income $a_\alpha(G) = (\mathrm{E}^G[x^{-\alpha}])^{-1/\alpha}$ for $\alpha > 0$ is a welfare-weighted average of incomes that declines (relative to the mean) as inequality rises. Its growth rate measures the combined effect of growth and changing inequality on social welfare, and is the basis of the upward mobility measure axiomatized by Genicot and Ray (2023). We show how this growth rate decomposes into drift and dispersion under the population dynamics in Assumption 1.

Proposition 8 (Decomposition of the Growth Rate of Atkinson Equivalent Income)**.** *Define* $E_{-\alpha}[f] \equiv \int_X x^{-\alpha} f(x)\, dG_t / \int_X x^{-\alpha}\, dG_t$*. Under Assumption 1, ignoring demography:*

$$\frac{d}{dt} \ln a_\alpha(G_t) = \underbrace{E_{-\alpha}[\mu_t(x)]}_{drift} - \underbrace{\frac{\alpha+1}{2} E_{-\alpha}[\sigma_t^2(x)]}_{dispersion}. \tag{5}$$

The drift term is a pro-poor weighted average of individual growth rates (equivalently, the instantaneous upward mobility measure M_α *of Genicot and Ray, 2023). The dispersion term is strictly positive whenever* $\sigma_t^2 > 0$*.*

The growth rate of the ratio $a_\alpha / \bar{y}$ *(a measure of equality, since* $a_\alpha \leq \bar{y}$ *with equality only under perfect equality) is:*

$$\frac{d}{dt} \ln \frac{a_\alpha}{\bar{y}} = \underbrace{(E_{-\alpha}[\mu_t(x)] - E_1[\mu_t(x)])}_{relative\ drift} - \underbrace{\frac{\alpha+1}{2} E_{-\alpha}[\sigma_t^2(x)]}_{dispersion}, \tag{6}$$

where $E_1[\mu_t(x)] = \int_X x\, \mu_t(x)\, dG_t / \int_X x\, dG_t$ *is the income-weighted (plutocratic) growth rate.*

The relative drift is positive when the pro-poor weighted growth rate exceeds the plutocratic growth rate — that is, when lower-income individuals are systematically growing faster. This is the force that most observers have in mind when discussing whether inequality is rising or falling. The dispersion term, by contrast, is always non-positive for the ratio $a_\alpha / \bar{y}$: idiosyncratic volatility mechanically raises inequality even in the absence of any systematic divergence. Separating the two requires panel data: the drift term depends on the covariance between individual income levels and individual growth rates, which is only observable when individuals are tracked over time.

The measure of Genicot and Ray (2023) over an interval $[s, t]$ is $\mu_\alpha = \frac{1}{t-s}[\ln a_\alpha(\mathbf{y}(t)) - \ln a_\alpha(\mathbf{y}(s))]$, which equals the time integral of (5). Their axioms — Growth Progressivity, reducibility, and additivity — characterize this measure under continuously differentiable income trajectories, where $\sigma_t \equiv 0$ and the dispersion term vanishes. In that deterministic setting, μ_α captures drift alone: the pro-poor weighted average growth rate, which is the economically relevant object. But when the formula is applied to stochastic income data — as in their empirical exercises using repeated cross-sections — it captures the combined effect of drift and dispersion. A key feature of their approach is panel independence: the measure depends only on the marginal distributions at dates s and t, not on individual trajectories. The decomposition above shows that this is both a strength and a limitation. Panel independence allows measurement from repeated cross-sections, but it also means that μ_α cannot separate systematic differential growth from the mechanical effect of idiosyncratic risk. In environments with substantial idiosyncratic volatility — as documented in the income dynamics literature — the dispersion component may be quantitatively important.

Proof. Define $P_\beta \equiv \mathrm{E}^G[x^\beta]$. By Itô's lemma, $d(x^\beta) = \beta\, x^\beta\, \mu(x)\, dt + \frac{1}{2}\beta(\beta-1)\, x^\beta\, \sigma^2(x)\, dt +$ martingale. Taking cross-sectional expectations:

$$\frac{d}{dt}\ln P_\beta = \beta\, E_\beta[\mu] + \frac{\beta(\beta-1)}{2}\, E_\beta[\sigma^2]. \tag{7}$$

Since $\ln a_\alpha = -\frac{1}{\alpha}\ln P_{-\alpha}$, substituting $\beta = -\alpha$:

$$\frac{d}{dt}\ln a_\alpha = -\frac{1}{\alpha}\left(-\alpha\, E_{-\alpha}[\mu] + \frac{(-\alpha)(-\alpha-1)}{2}\, E_{-\alpha}[\sigma^2]\right) = E_{-\alpha}[\mu] - \frac{\alpha+1}{2}\, E_{-\alpha}[\sigma^2],$$

establishing (5). For (6), note that $\bar{y} = P_1$, and (7) with $\beta = 1$ gives $d\ln\bar{y}/dt = E_1[\mu]$ (the second-order term vanishes since $\beta(\beta-1) = 0$). Subtracting yields (6). □

B Appendix for Section 3

B.1 Weighted Average

The multivariate framework nests the decomposition of a weighted average. Consider $\nu(G) = \mathrm{E}^G[x_1 x_2]$ (e.g., average productivity x_1 weighted by size x_2). Then:

$$IF_\nu(\boldsymbol{x}; G) = x_2 x_1 - \mathrm{E}^G[x_2 x_1]$$

$$\frac{d}{dt}\nu(G_t) = \underbrace{\mathrm{E}^{G_t}[x_1\mu_{1t} + x_2\mu_{2t}]}_{\text{Drift}} + \underbrace{\mathrm{E}^{G_t}[\rho_t\sigma_{1t}\sigma_{2t}]}_{\text{Dispersion}}$$

For the case of average productivity π weighted by size share $s = S/\mathrm{E}^G[S]$:

$$\frac{d}{dt}\nu(G_t) = \underbrace{\mathrm{E}^{G_t}[s\mu_{\pi t} + \pi\mu_{st}]}_{\text{Drift}} + \underbrace{\mathrm{E}^{G_t}[\sigma_{\pi t}\sigma_{st}]}_{\text{Dispersion}}$$

This recovers exactly the decomposition of Melitz and Polanec (2015). Note that for a weighted average (a linear functional), the dispersion term depends on the *covariance* between the two shocks, not on the variance of either shock individually.

B.2 Demography

For all population-invariant indices in Section 3, the discrete-time decomposition separates the change among survivors from the effect of entry and exit:

$$\nu(G_{t+\Delta t}) - \nu(G_t) = \underbrace{\nu(\widetilde{G}_{t+\Delta t}) - \nu(\widetilde{G}_t)}_{\text{Drift+Dispersion}} + \underbrace{\nu(G_{t+\Delta t}) - \nu(\widetilde{G}_{t+\Delta t}) + \nu(\widetilde{G}_t) - \nu(G_t)}_{\text{Demography}}$$

where $S = \Omega_t \cap \Omega_{t+\Delta t}$ is the set of survivors.

In continuous time, since the OWA and generalized entropy indices in Section 3 are population invariant (Section 2.1), the demography term reduces to:

$$\text{Demography} = (\eta_t + \delta_t) \int_X IF_\nu(x; G_t)\, d\Psi_{Bt}(x) - \delta_t \int_X IF_\nu(x; G_t)\, d\Psi_{Dt}(x)$$

For indices satisfying the transfer principle, IF_ν is convex with $\int_X IF_\nu \, dG_t = 0$: it is negative for low incomes and positive for high incomes. The demography term is therefore positive when entrants are richer than average or exiters are poorer than average.

OWA indices. Using integration by parts, the demography term can be written in terms of the CDFs Ψ_{Bt} and G_t:

$$\int_X IF_\nu(x; G_t)\, d\Psi_{Bt}(x) = \frac{1}{\mathrm{E}^{G_t}[x]} \int_X w(G_t(x))\, (\Psi_{Bt}(x) - G_t(x))\, dx + \nu(G_t) \left(1 - \frac{\mathrm{E}^{\Psi_{Bt}}[x]}{\mathrm{E}^{G_t}[x]}\right)$$

The first term captures the effect of rank differences between entrants and incumbents: it is positive when entrants are concentrated at ranks where w is large (e.g., the top for top shares). The second term captures the mean income difference, weighted by the index value. For the *top p share* ($w(r) = \mathbf{1}_{r \geq 1-p}/p$), this simplifies to:

$$\int_X IF_\nu \, d\Psi_{Bt} = \frac{1}{\mathrm{E}^{G_t}[x]} \left(\frac{\mathrm{E}^{\Psi_{Bt}}[\max(x - q_p, 0)]}{p} - \nu(G_t)\, \mathrm{E}^{\Psi_{Bt}}[x] \right)$$

where q_p denotes the $(1-p)$-th quantile of G_t (i.e., the threshold for the top p share). This is large when entrants have high incomes above q_p.

Comparison with Gomez (2022). For the top share, the discrete-time decomposition in this paper differs slightly from the one in Gomez (2022). In that paper, the "within" term tracks the wealth growth of individuals who were initially in the top group and survive, $\mathcal{P}_0 \setminus \mathcal{D}$. The "between" term captures reshuffling among survivors between the top and the rest. The "demography" term accounts for deaths—using a time-adjusted wealth for the deceased—and births into the top group, both measured relative to $q_{t+\Delta t}$, the quantile threshold at $t + \Delta t$.

In the present paper, the demography term instead operates at time t: removing the dead from the full population at t changes the survivor distribution $\widetilde{G}_t$, which promotes individuals from just below the threshold q_p into the top group of the survivor population, at their time-t wealth. The drift and dispersion terms then track *all* survivors—including these newly promoted individuals—from t to $t + \Delta t$.

The two decompositions have the same continuous-time limit and are quantitatively very close for short time intervals. To see why, note that the mass of promoted individuals is $O(\delta_t \Delta t)$. Their contribution to drift and dispersion over the interval Δt is therefore $O(\delta_t \Delta t \cdot \Delta t)$, which vanishes relative to Δt as $\Delta t \to 0$. The same reasoning applies to population growth. The advantage of the present formulation is that it holds as an exact accounting identity for any Δt and any death rate δ_t—it does not require the number of deaths to be infinitesimal.

Generalized entropy. Using the explicit influence function:

$$\int_X IF_\nu \, d\Psi_{Bt} = \frac{1}{\alpha(\alpha-1)} \left(\mathrm{E}^{\Psi_{Bt}}\left[\left(\frac{x}{\mathrm{E}^{G_t}[x]} \right)^{\alpha} \right] - 1 - \alpha \left(\frac{\mathrm{E}^{\Psi_{Bt}}[x]}{\mathrm{E}^{G_t}[x]} - 1 \right) \right)$$

This has the same functional form as ν_α itself, but evaluated at the entrant distribution Ψ_{Bt} relative to the incumbent mean $\mathrm{E}^{G_t}[x]$. It is zero when entrants have the same distribution as incumbents and positive (by convexity) whenever entrants differ.

Herfindahl-Hirschman Index. The HHI is not population invariant, so the full demography term from Proposition 1 applies. The influence function is:

$$IF_H(x;G) = \frac{x^2}{\left(\int_X x' \, dG\right)^2} - \frac{2\,H(G)\,x}{\int_X x' \, dG}$$

Note that $\int_X IF_H \, dG = H - 2H = -H < 0$, confirming the index is not population invariant. The full continuous-time decomposition is:

$$\frac{d}{dt} H(G_t) = \underbrace{2\,\mathrm{E}^{G_t}\left[\left(\frac{x}{\mathrm{E}^{G_t}[x]} \right)^2 (\mu_t(x) - \bar{\mu}_t) \right]}_{\text{Drift}} + \underbrace{\mathrm{E}^{G_t}\left[\left(\frac{x}{\mathrm{E}^{G_t}[x]} \right)^2 \sigma_t(x)^2 \right]}_{\text{Dispersion}}$$
$$+ \underbrace{(\eta_t + \delta_t) \int_X IF_H(x;G_t)\, d\Psi_{Bt}(x) - \delta_t \int_X IF_H(x;G_t)\, d\Psi_{Dt}(x)}_{\text{Demography}}$$

The demography term can be decomposed into a selection effect and a dilution effect:

$$\text{Demography} = \underbrace{(\eta_t + \delta_t) \left(\int_X IF_H \, d\Psi_{Bt} + H \right) - \delta_t \left(\int_X IF_H \, d\Psi_{Dt} + H \right)}_{\text{selection}} \underbrace{-\eta_t\, H}_{\text{dilution}}$$

The dilution term $-\eta_t H$ is always negative: population growth mechanically reduces concentration even when entrants are drawn from the same distribution as incumbents. This is the continuous-time analog of the $1/N$ effect.

The full discrete-time accounting framework is:

$$H(G_{t+\Delta t}) - H(G_t) = \underbrace{2N\,\mathrm{E}[s_{it}\,\Delta s_{it} \mid i \in S]}_{\text{Drift}} + \underbrace{N\,\mathrm{E}\left[(\Delta s_{it})^2 \mid i \in S\right]}_{\text{Dispersion}}$$
$$+ \underbrace{H(G_{t+\Delta t}) - H(\widetilde{G}_{t+\Delta t}) + H(\widetilde{G}_t) - H(G_t)}_{\text{Demography}}$$

where $s_{it} = x_{it} / \sum_j x_{jt}$ denotes market shares and N the number of survivors, consistent with Table 2.

B.3 Derivation of the GE Decomposition

Write $s = x/\mathrm{E}^G[x]$ for income shares, so that $\nu_\alpha = \frac{1}{\alpha(\alpha-1)}(\mathrm{E}^G[s^\alpha] - 1)$. Under the diffusion $dx = x\mu\,dt + x\sigma\,dW$, shares follow $ds/s = (\mu - \bar{\mu})\,dt + \sigma\,dW$, where $\bar{\mu} = \mathrm{E}^G[x\mu]/\mathrm{E}^G[x]$. By Itô's lemma,

$$d(s^\alpha) = s^\alpha\left[\alpha(\mu - \bar{\mu})\,dt + \alpha\sigma\,dW + \tfrac{1}{2}\alpha(\alpha-1)\sigma^2\,dt\right].$$

Taking expectations and dividing by $\alpha(\alpha - 1)$:

$$\frac{d}{dt}\nu_\alpha = \frac{1}{\alpha-1}\mathrm{E}^{G_t}[s^\alpha(\mu-\bar{\mu})] + \frac{1}{2}\mathrm{E}^{G_t}\left[s^\alpha\sigma^2\right] = \frac{1}{\alpha-1}\mathrm{E}^{G_t}\left[\left(\frac{x}{\mathrm{E}^{G_t}[x]}\right)^\alpha(\mu-\bar{\mu})\right] + \frac{1}{2}\mathrm{E}^{G_t}\left[\left(\frac{x}{\mathrm{E}^{G_t}[x]}\right)^\alpha\sigma^2\right].$$

B.4 Characterization of Generalized Entropy

Theorem 1 (Dynamic Characterization of Generalized Entropy)**.** *Let ν be a smooth, scale-invariant, population-invariant distributional index satisfying the transfer principle. Suppose the dispersion weight is multiplicatively separable in income level and distribution: there exist functions $h : \mathcal{F} \to \mathbb{R}_+$ and $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

$$x^2\,\partial_{xx} IF_\nu(x;G) = h(G)\,\phi(x) \quad \textit{for all } x \textit{ and } G,$$

where ϕ does not depend on G. Then ν belongs to the generalized entropy family, up to affine rescaling.

Proof. Scale invariance implies $IF_\nu(\lambda x; G_\lambda) = IF_\nu(x;G)$. Differentiating twice: $\lambda^2\,\partial_{xx} IF_\nu(\lambda x; G_\lambda) = \partial_{xx} IF_\nu(x;G)$. Substituting separability: $\phi(\lambda x)/\phi(x) = h(G)/[\lambda^2 h(G_\lambda)]$ is independent of x, so $\phi(x) = \phi(1)\,x^r$ for some $r > -2$. Setting $\alpha = r + 2$ and integrating twice: $IF_\nu(x;G) = C(G)\,x^\alpha + A(G)\,x + B(G)$. The scaling relation gives $C(G) = \kappa/\mathrm{E}^G[x]^\alpha$. Population invariance and scale invariance pin down $A(G)$ and $B(G)$. The resulting influence function is that of $\kappa\,\nu_\alpha(G)$. □

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and wealth distribution in macroeconomics: a continuous-time approach," *Review of Economic Studies*, 2022, *89* (1), 45–86.

Atkinson, Anthony B, "On the measurement of inequality," *Journal of Economic Theory*, 1970, *2* (3), 244–263.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Beare, Brendan K and Alexis Akira Toda, "Determination of Pareto exponents in economic models driven by Markov multiplicative processes," *Econometrica*, 2022, *90* (4), 1811–1833.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu, "The distribution of wealth and fiscal policy in economies with finitely lived agents," *Econometrica*, 2011, *79* (1), 123–157.

Bourguignon, Francois, "Decomposable income inequality measures," *Econometrica: Journal of the Econometric Society*, 1979, pp. 901–920.

Campbell, John Y, Tarun Ramadorai, and Benjamin Ranish, "Do the rich get richer in the stock market? Evidence from India," *American Economic Review: Insights*, 2019, *1* (2), 225–240.

Cowell, Frank A and Maria-Pia Victoria-Feser, "Robustness properties of inequality measures," *Econometrica*, 1996, *64* (1), 77–101.

Dalton, Hugh, "The measurement of the inequality of incomes," *The Economic Journal*, 1920, *30* (119), 348–361.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 1996, *64* (5), 1001–1044.

Firpo, Sergio, Nicole M Fortin, and Thomas Lemieux, "Unconditional quantile regressions," *Econometrica*, 2009, *77* (3), 953–973.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo, "Decomposition methods in economics," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1–102.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The dynamics of inequality," *Econometrica*, 2016, *84* (6), 2071–2111.

Genicot, Garance and Debraj Ray, "Measuring upward mobility," *American Economic Review*, 2023, *113* (11), 3044–3089.

Gomez, Matthieu, "Decomposing the Growth of Top Wealth Shares," 2022. Working paper.

— **and Emilien Gouin-Bonenfant**, "Wealth inequality in a low rate environment," *Econometrica*, 2024, *92* (4), 1225–1260.

Hampel, Frank R, "The influence curve and its role in robust estimation," *Journal of the american statistical association*, 1974, *69* (346), 383–393.

Hampel, Frank Rudolf, *Contributions to the theory of robust estimation*, University of California, Berkeley, 1968.

Kolm, Serge-Christophe, "Unequal inequalities. I," *Journal of Economic Theory*, 1976, *12* (3), 416–442.

Mehran, Farhad, "Linear measures of income inequality," *Econometrica: Journal of the Econometric Society*, 1976, pp. 805–809.

Melitz, Marc J. and Sašo Polanec, "Dynamic Olley-Pakes productivity decomposition with entry and exit," *The Rand journal of economics*, 2015, *46* (2), 362–375.

Piketty, Thomas and Emmanuel Saez, "Income inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, 2003, *118* (1), 1–41.

Saez, Emmanuel and Gabriel Zucman, "Wealth inequality in the United States since 1913: evidence from capitalized income tax data," *Quarterly Journal of Economics*, 2016, *131* (2), 519–578.

Shorrocks, Anthony F, "The class of additively decomposable inequality measures," *Econometrica*, 1980, *48* (3), 613–625.

— , "Inequality decomposition by population subgroups," *Econometrica*, 1984, *52* (6), 1369–1385.

— , "Revisiting the Sen poverty index," *Econometrica: Journal of the Econometric Society*, 1995, pp. 1225–1230.

Yaari, Menahem E, "A controversial proposal concerning inequality measurement," *Journal of Economic Theory*, 1988, *44* (2), 381–397.