Inequality and Immobility: Some Theoretical Models

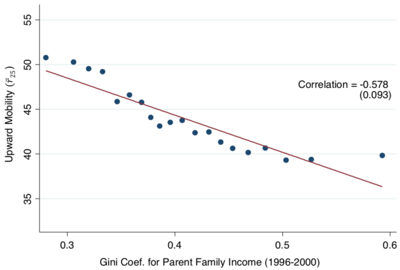
Lawrence Blume, Steven Durlauf, Aleksandra Lukina

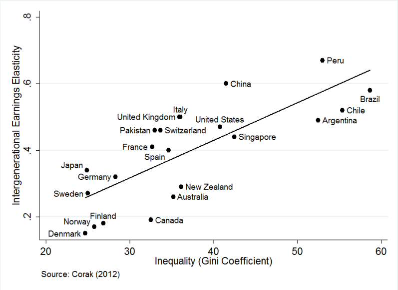
Cornell, University of Chicago, University of Chicago

The Great Gatsby Curve

$$\ln Y_i^{child} = \alpha + \beta \ln Y_i^{parent} + \epsilon_i$$

B. Upward Mobility vs. Gini Coefficient in CZ
The "Great Gatsby" Curve Within the U.S.





Modeling Strategies

- Becker-Tomes

  Positive intergenerational correlations in socioeconomic status can arise from the greater ability of richer parents to invest in their children's human capital, from genetic or cultural inheritance, or from all of the above.

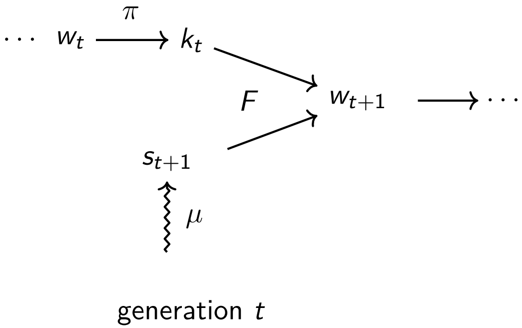
- Loury

  Persistence of inequality is modeled as a consequence of credit constraints that limit the human capital investment opportunities of poor families.

- Our Paper

  Microfoundations of Becker-Tomes models
  Becker Tomes dynamics — a critique of "poverty traps".
  A microeconomic foundation for the Great Gatsby idea — alternative observational strategies.
  A "social competition" model of inequality.

The evolution of wealth



The family optimization problem

Each family gets utility from their own consumption and their children's wealth. It has payoff function $U(c_t, w(t+1))$ and beliefs μ about s_{t+1}. Each family solves an optimization problem. The objective function is

$$V(c_t, k_t) \equiv E_\mu \left\{ U\big(c_t, F(k_t, s_{t+1})\big) \right\}.$$

The optimal policy is the correspondence $\pi : \mathbf{R}_+ \rightrightarrows \mathbf{R}_+$ given by

$$\begin{aligned}\pi\big(w_t\big) = \big\{ k_t : \text{there is a } c_t \geq 0 \text{ s.t.} \\ (c_t, k_t) \in \operatorname{argmax}_{c_t, k_t} \; V(c_t, k_t) \\ \text{s.t.} \quad c_t + C(k_t) \leq w(t) \\ c_t, k_t \geq 0. \big\}\end{aligned}$$

Transitions

- Neoclassical F: concave in k and $F(0) = 0$.
- Stepping-stone: Investment levels $k^0 = 0, k^1, \ldots, k^M$, and

$$F(k, s) = \tilde{w}^m(s) \text{ if } k^m \leq k < k^{m+1}.$$

- Becker-Tomes: $F(k, s) = s + (1 + r)k$.
- Lumpy neoclassical: neoclassical $F_1, \ldots, F_J$ given, for each j there is a k_j s.t. $F_j(k, s) > F_{j-1}(k, s)$ for $k > k_j$ and all s.

$$F(k, s) = \max_j F_j(k, s).$$

Assumptions

A.1. Utility $U(c, w)$ is continuous and strictly increasing in consumption c and child's wealth w.

A.2. Utility is concave.

A.3. $U\big(c, F(k, s)\big)$ is Lebesgue integrable for all $c, k \geq 0$.

A.4. $F(k, s)$ is non-negative, non-decreasing and continuous in k; and increasing in s.

A.5. $C(k)$ is continuous and strictly increasing.

A.6. Utility is supermodular ($U_{cw} \geq 0$ if it is differentiable).

Existence and Comparative Statics

Theorem 1. Assume A.1, A.3–A.5. For all $w \geq 0$, $\pi(w) \neq \emptyset$ and for any selection $k(w)$ from π, $V\big(w, k(w)\big)$ is upper semicontinuous.

Theorem 2. Assume A.1–6. π is increasing: If $w' > w''$, if k' is optimal for w' and k'' is optimal for w'', then $k' \geq k''$.

A Family of Markov Chains

A family of probability measures μ^ϵ with densities

$$\phi^\epsilon(s) = Z(\epsilon) \exp \frac{-h(s)}{\epsilon}$$

h has a min at 0 and is C^2 outside of 0. e.g. Normal distributions.

$\mathcal{W} = [0, w^*]$. F pushes wealth draws outside the interval to the nearest boundary.

$$P^\epsilon(w, A) = \mu^\epsilon(F(\pi^\epsilon(w), s) \in A).$$

For $A = [0, w']$ and $\epsilon > 0$,

$$P^\epsilon(w, A) = \begin{cases} \int_{-\infty}^{w' - G(\pi^\epsilon(w))} \phi^\epsilon(s) ds & \text{if } w' < w^*, \\ 1 & \text{if } w' = w^*. \end{cases}$$

The Equilibrium Distribution of Wealth

The equilibrium distribution of wealth — the invariant distribution of the process.

$$\mu^\epsilon(A) = \int P^\epsilon(w, A) d\mu^\epsilon(w).$$

Theorem: For all $\epsilon > 0$ the Markov processes $\{W_t^\epsilon\}$ have a single invariant distribution, and the time-t distribution of states converges to it from any initial distribution of states.

What do the invariant distributions look like?

- Masses at 0 and w^*.
- Absolutely continuous wrt Lebesgue measure in the interior.
- Shape of the density?

Deterministic Wealth Dynamics

When $\epsilon = 0$, wealth evolution is deterministic. It is governed by a difference equation whose right-hand side is the parental Engel curve:

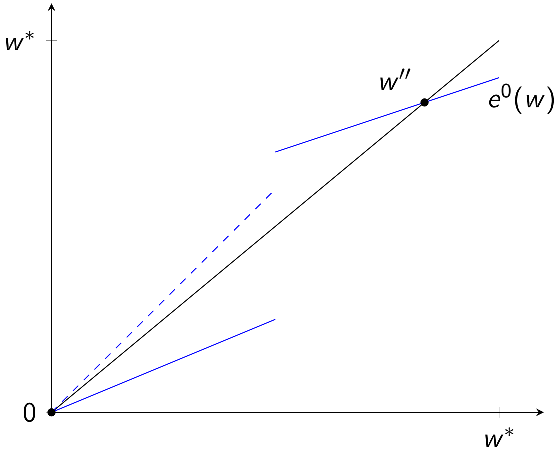
$$w_{t+1} = G(\pi^0(w_t)) \equiv e^0(w_t). \tag{1}$$

We know $e : \mathcal{W} \to \mathcal{W}$ is non-decreasing. Thus it will have fixed points:

Theorem. There is a $w' \in \mathcal{W}$ such that $e^0(w') = w'$.

Assume $e^0(w)$ has only a finite number of fixed points.

Deterministic Dynamics



Deterministic Dynamics

Definition. A fixed point w' is an attractor if there is an open interval $U \ni w'$ such that $\cap_{t \geq 0} e^{0t}(U) = \{w'\}$ and for all open $V \supset \{w'\}$ and t large enough, $e^t(U) \subset V$.

Theorem. e has at least one attractor.

The set of attractors is ordered. When $e(w)$ has more than one attractor, the smallest attractor is a poverty trap. Equally, it could be said that the highest attractor is an affluence trap.

Low-Noise Invariant Distributions

As $\epsilon \downarrow 0$, the process looks ever more deterministic. Does it settle on attractors? If so, which ones?

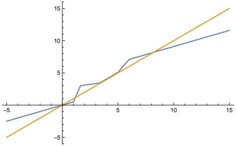
Theorem. There is a set $\mathcal{B} = \{w_1, \ldots, w_p\}$ of attractors such that for any closed set $D \subset \mathcal{W}$ disjoint from $\{w_i : i \in \mathcal{B}\}$ and for any sequence of invariant measures $\{\nu^\epsilon\}$ of the transition probabilities P^ϵ with $\epsilon \to 0$,

$$\lim_{\epsilon \to 0} \nu^\epsilon(D) = 0.$$

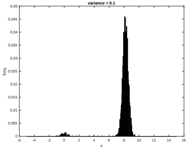
Idea

- Let $w_0, \ldots, w^T$ be a path from w_0 to w^T.
- The cost of traversing this path is $\sum_{t=0}^{T-1} h\big(w_{t+1} - e^0(w_t)\big)$.
- For attractors w_i and w_j, B_{ij} is the cost of the minimum-cost path from i to j.
- Consider a tree γ labelled with attractors and directed towards root w_i. The cost of γ is the sum of the costs of all edges in γ. The cost of attractor w_i is the minimum cost of all such graphs.
- Only min-cost attractors can have positive weight in the invariant limit.

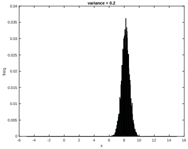
Simulations



(a) The deterministic system

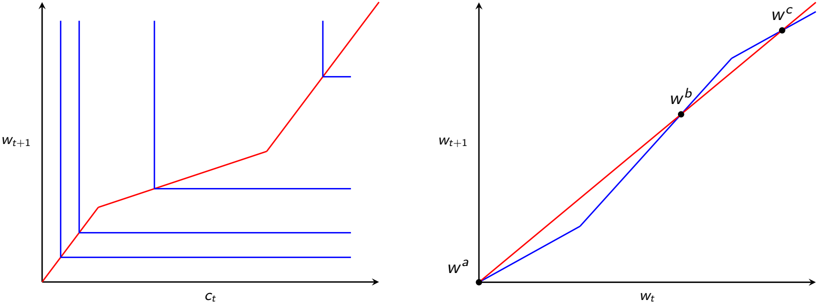


(b) $\sigma^2 = 0.1$



(c) $\sigma^2 = 0.2$

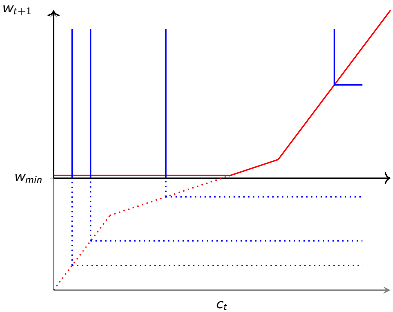
Examples



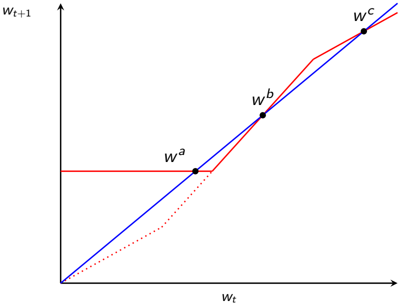
(a) An indifference map (b) The Engel curve

Figure: Example preferences and Engel curve.

A Wealth Floor



(a) An indifference map



(b) The Engel curve

Figure: Example preferences and Engel curve.

An Investment Subsidy

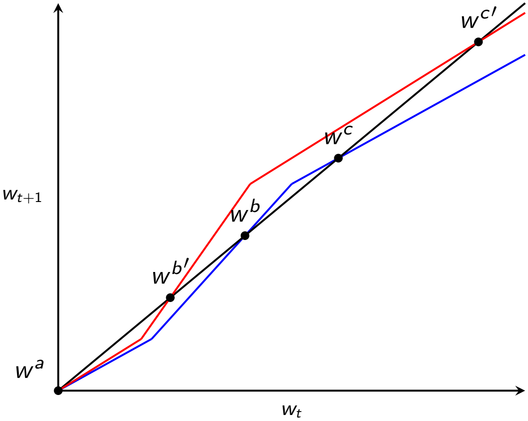


Figure: A wealth floor.

The Deterministic Stepping-Stone Model

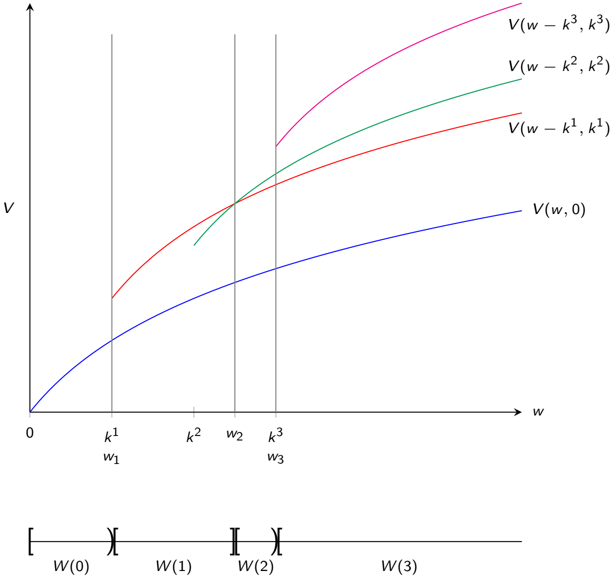
- A finite number of distinct investment levels.
- Some are fixed-points — stationary in the dynamics of the picture.
- Other investment levels are transitory, leading to one fixed-point or another.
- There are no cycles.
- This model may exhibit poverty traps — multiple inescapable steady-states.

Characterizing π

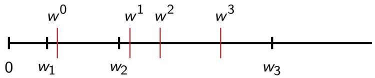
Characterization Theorem. The graph of π is characterized by no more than $N+1$ intervals with non-intersecting interiors (possibly degenerate) that cover $\mathbf{R}_+$, and an equal number of k-values. These characterize the graph of π as follows:

1. Associated with each interval $W(i)$ is a unique value k^i which is optimal on the interior of $W(i)$.
2. k^i is the unique optimum on the interior of $W(i)$.
3. k^i is optimal at the left endpoint, that is, each interval $W(i)$ contains its left endpoint.
4. π is increasing.

The deterministic stepping-stone model



Dynamics

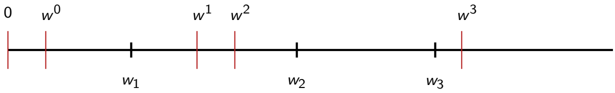


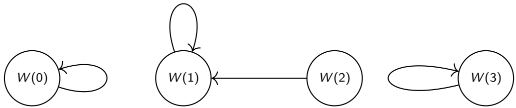
- A family with initial wealth less than w_1 invests 0. the next generation has wealth w^0.
- A family with initial wealth between w_1 and w_2 invests k^1. The next generation has wealth w^1.
- A family with wealth between w_2 and w_3 invests k^2, and all subsequent generations have wealth w^2 and invest k^2.
- A family with wealth $w \geq w_3$ invest k^2 and has wealth w^2.
- Dynastic wealth converges to w^2 in finite time.

Dynamics can be arbitrary, with multiple basins of attractions. The only constraint is that family fortunes cannot cross.

Dynamics

Another possibility





$w^1, w^2 \in W(1)$, $w^3 \in W(3)$.

The stochastic stepping-stone model

$$F(k) = \begin{cases} \tilde{w}^0(s) & \text{if } 0 \ \leq k < k^1, \\ \tilde{w}^1(s) & \text{if } k^1 \leq k < k^2 \\ \text{etc,} & \end{cases}$$

where the $\tilde{w}^n$ are non-negative random variables which strictly increase with i in the sense of first-order stochastic dominance.

Let $g^{k_i}(w)$ denote the density of $\tilde{w}^i$.

Theorem: The conclusions of the characterization theorem still hold.

The Markov process

All selections $\dot{\pi}$ from π differ from each other only at the wealths w_i, where multiple k^i are optimal. Each selection describes a Markov process. For measurable $A \subset W$,

$$P_{\dot{\pi}}(w_{t+1} \in A | w_0, \ldots, w_t) =$$
$$P_{\dot{\pi}}(w_{t+1} \in A | w_t) \equiv P_{\dot{\pi}}(w_t, A) = \int_A g^{\dot{\pi}(w_t)}(w)\, dw.$$

A parent in $W(i)$ chooses capital investment k^i. The child's wealth will be $\tilde{w}^i$, drawn from density g^i. Let p_{ij} denote the probability that $\tilde{w}^i \in W(j)$.

$$p_{ij} \equiv P_{\dot{\pi}}\big(w_{t+1} \in W(j) | w_t \in W(i)\big) = \int_{W(j)} g^{k_i}(w)\, dw.$$

The Markov Process

Choose a selection $\dot{\pi}$. Then

$$\nu_{t+1}(A) = \int P(w_t, A)\, d\nu_t = \sum_m \nu_t(W(m)) \int_A g^{k_m}(w)\, dw$$

and for any integrable function f,

$$\int f\, d\nu_{t+1} = \sum_m \nu_t(W(m)) \int f(w) g^{k_m}(w)\, dw.$$

ν_t matters only throughs $\bar{\nu}_t = \big(\nu_t(W(1), \ldots, \nu_t(W(N))\big)$.

$$\bar{\nu}_{t+1} = \bar{\nu}_t \cdot [p_{ij}].$$

The Markov Process

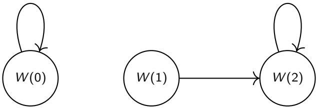
B.1. $[p_{ij}]$ is irreducible.

B.2. There is an i such that $\int_{W(i)} g^{k_i}(w)dw > 0$.

Theorem 3: Assume A.1–A.5, B.1 and B.2. The Markov process with transition probability $P_{\dot{\pi}}$ and any initial distribution is ergodic. The invariant distribution ν^* has density $\sum_i q_i^* g^i(w)$ where q^* is the invariant distribution of $[p_{ij}]$. Finally, $P_{\dot{\pi}}$ and ν^* are independent of the selection $\dot{\pi}$; that is, they depend only on π.

Stochastic stability of poverty traps

The idea



The transition matrix for this process is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The invariant probability distributions are the convex hull of $(1, 0, 0)$ and $(0, 0, 1)$.

Stochastic stability of poverty traps

The idea

Suppose the matrix is perturbed so that with small probability the state moves to a neighboring interval. The perturbed transition probability is

$$\frac{1}{1+\epsilon}\begin{pmatrix}1 & \epsilon & 0\\ \epsilon & 0 & 1\\ 0 & \epsilon & 1\end{pmatrix}.$$

The invariant measure for this matrix is $\nu_\epsilon = (1+2\epsilon)^{-1}(\epsilon, \epsilon, 1)$. And the limit as $\epsilon \to 0$ is $\nu = (0, 0, 1)$. State 3 is stochastically stable under this perturbation.

$$\nu(i) = \lim_{\epsilon\to 0} \nu_\epsilon(i) = \lim_{\epsilon\to 0}\lim_{t\to\infty} \Pr\{s(t) = i | s(0) = j\}.$$

This is not a graph property. If the 0s are replaced with ϵ, then $\nu(1) = 1/3$.

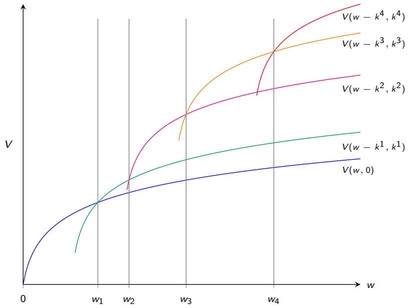
A Stepping-Stone Example

Take $U(c,w) = log(0.1 + c) + 1.5w$. This satisfies A.1–A.5.

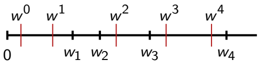
$$f(k) = \begin{cases} 0.3 & \text{if} \quad k < 1, \\ 1 & \text{if} \quad 1 \le k < 2, \\ 2.4 & \text{if} \quad 2 \le k < 3, \\ 3.5 & \text{if} \quad 3 \le k < 4.5, \\ 4.5 & \text{if} \; 4.5 \le k. \end{cases}$$

$$W(0) \approx [0; 1.44], \qquad W(3) \approx [3.14; 4.83],$$
$$W(1) \approx [1.44; 2.04], \qquad W(4) \approx [4.83; +\infty).$$
$$W(2) \approx [2.04; 3.14],$$

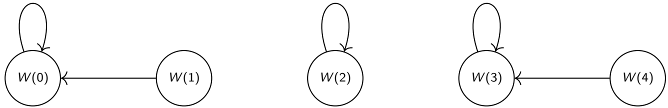
Example



$V(w-k^4,k^4)$
$V(w-k^3,k^3)$
$V(w-k^2,k^2)$
$V(w-k^1,k^1)$
$V(w,0)$
V
w
0
w_1
w_2
w_3
w_4



w^0
w^1
w^2
w^3
w^4
0
w_1
w_2
w_3
w_4

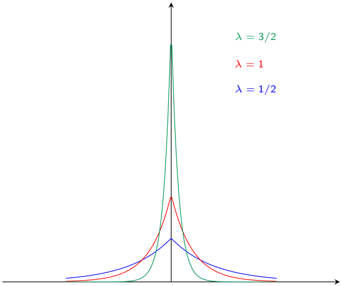


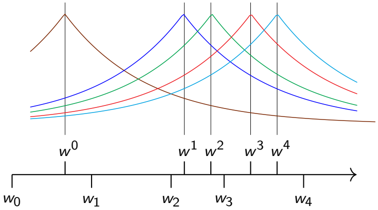
$W(0)$
$W(1)$
$W(2)$
$W(3)$
$W(4)$

Example

Stochastic stepping-stone model

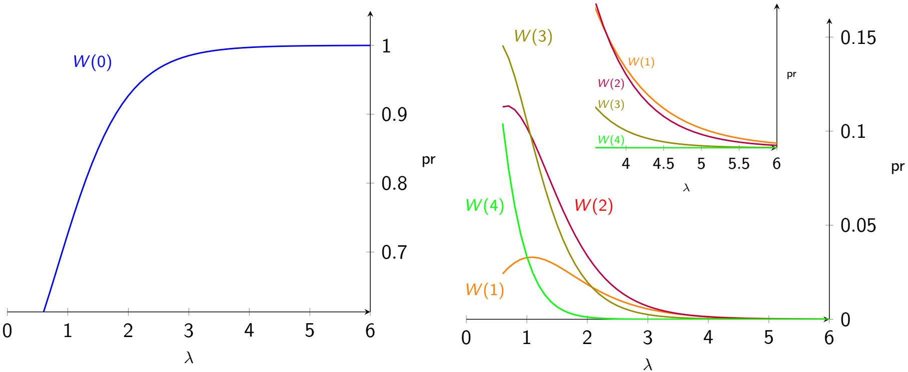
$\tilde{w}^i = \max\{w^i + s, 0\}$. The random variable s has a Laplace distribution, i.e. the density of s is $h_\lambda(s) = \frac{\lambda}{2} exp(-\lambda|x|)$.

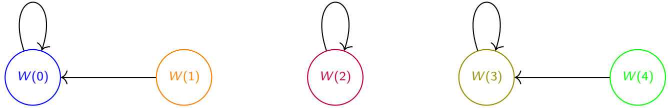




Example

Convergence to the limit invariant measure





Example

Convergence to the limit invariant measure

The calculations show two facts:

- $W(0)$ is the only stochastically stable state.
- For all $i > 2$, $\nu_\lambda\big(W(1)\big)/\nu_\lambda\big(W(i)\big) \to \infty$, although $W(1)$ is not a steady state of the deterministic dynamics.

Stochastic stability of poverty traps

As $\lambda \to \infty$, the distribution of s converges weakly to point mass at 0. The boundaries w_i are functions of this distribution. Write $w_i(\lambda)$.

B.3. h is C^2 at 0.

B.4. The boundaries $w_i(\lambda)$ converge to the deterministic boundaries as $\lambda \to \infty$.

Theorem. Assume A.1–5 and B.1–4. If $W(m)$ is stochastically stable, then it is an attractor in the deterministic dynamics. For "most" h there is a unique stochastically stable $W(i)$.

The Gatsby Curve

Stopping times:

τ_y is the time of the first visit to y.

τ_x^+ is the time of the first return to x.

Theorem:

$$\nu^*(W(i)) \cdot \Pr^{W(i)}\{\tau_{W(j)} < \tau^+_{W(i)}\} = \nu^*(W(j)) \cdot \Pr^{W(j)}\{\tau_{W(i)} < \tau^+_{W(j)}\}$$

Gatsby Curve

$$\frac{\nu^*(W(i))}{\nu^*(W(j))} = \frac{\Pr^{W(j)}\{\tau_{W(i)} < \tau^+_{W(j)}\}}{\Pr^{W(i)}\{\tau_{W(j)} < \tau^+_{W(i)}\}}$$

Attractors are easy to enter, and hard to leave.

Suppose the income distribution is skewed so that lower-income states have higher invariant probabilities. Then it is more likely that a high-wealth dynasty will fall to a lower-wealth dynasty before returning to high wealth is greater than the probability that a low-wealth dynasty will experience high wealth before returning to low wealth. Low wealth is stickier than high wealth.

Gatsby Curve

From an initial condition, how long does it take to get near the invariant distribution?

$$\tau_M(\epsilon) = \inf_t \sup_{\nu_0} \left|\left|\nu_0 \cdot [p_{ij}]^t - \nu^*\right|\right|_{\mathrm{var}} < \epsilon.$$

- Mixing times are long when poverty traps are deep, shorter when they are not.
- There are a number of relationships between mixing times and entrance times that we plan to exploit.

Poverty traps are a source of the Great Gatsby Curve.