Models of Intergenerational Inequality

Larry Blume

Cornell University & IHS & The Santa Fe Institute & HCEO

FCS 2018

广州市

Plan

- Introduction
  - The family optimization problem
  - Comparative statics
- The deterministic stepping-stone model
  - The optimal investment correspondence
  - Dynamics
- The stochastic stepping-stone model
  - Description and the optimal policy
  - The Markov process
  - Stochastic stability of poverty traps

INTRODUCTION

The Intergenerational Transmission of Inequality

Preferences

- Parents invest in children — *inter vivos, testamentary*
- Preferences can take several forms
  - Consumption-based: $U_t(c_t, c_{t+1})$.
    Arrow (1973), Bernheim and Ray (1987)
  - Income-based: $U(c_t, y_{t+1})$.
    Becker and Tomes (1979)
  - Utility of giving: $U(c_t, k_t)$.
    Banerjee and Newman (1993).
  - Child's Utility-based: $U(c_t, U_{t+1})$.
    Loury (1981).

How distinct are these?

The Intergenerational Transmission of Inequality

Transition function

- Parents make some investment, bequest, whatever, in children ..., k_t. This becomes wealth for children in the next generation, w_{t+1}.
- Production functions can take several forms
  - Neoclassical deterministic or stochastic f, concave and $f(0) = 0$.
  - Deterministic stepping-stone:

$$f(k) = \begin{cases} w^0 & \text{if } k < k^1, \\ w^1 & \text{if } k^1 \leq k < k^2, \\ \vdots & \vdots \end{cases}$$

  - Stochastic stepping-stone: $\tilde{w}^i$ are ordered by f.o.s.d.
  - Becker-Tomes: $f(k, s) = s + (1 + r)k$,
  - Lumpy 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. $f(k, s) = \max_j f_j(k, s)$.

How the Analysis Works

A dynasty is a sequence of families. Parents at time t were children at time $t-1$, and divide their wealth between their own consumption and investment in their children. This is in the style of Becker and Tomes (1979). Perhaps also Loury (1981)?

Children's wealth at time $t+1$ is determined (perhaps with noise) by their parent's investment. As parents, they in turn . . . .

- We analyze the family optimization problem.
- We chain the optimization problems together in a dynasty to analyze the distribution of wealth through time; within period inequality and the persistence of inequality over time.
- We allow dynasties to compete for social position and wealth, and examine the effects of this competition on the inequality questions (maybe).

How the Analysis Works

A dynasty is a sequence of families. Parents at time t were children at time $t-1$, and divide their wealth between their own consumption and investment in their children. This is in the style of Becker and Tomes (1979). Perhaps also Loury (1981)?

Children's wealth at time $t+1$ is determined (perhaps with noise) by their parent's investment. As parents, they in turn . . . .

- We analyze the family optimization problem.
- We chain the optimization problems together in a dynasty to analyze the distribution of wealth through time; within period inequality and the persistence of inequality over time.
- We allow dynasties to compete for social position and wealth, and examine the effects of this competition on the inequality questions (maybe).

Loury is different from the others because his family optimization problem becomes recursive dynamic programming problem. Policies that perturb the investment function perturb utility directly by changing the value function.

Today we won't discuss Loury.

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

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

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

$$V(c(t), k(t)) = E_\mu \left\{ U\big(c(t), f(k(t), s(t+1))\big) \right\}.$$

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

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

What do we want to ask about π?

- Does it exist?
- What does it look like?

What will de do with π? How can we use it to build an intergenerational model?

Assumptions are next.

Assumptions

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

A.2. Utility is supermodular ($U_{12} \geq 0$ if it is differentiable).

A.3. Utility is strictly concave in consumption c.

A.4. Utility $U(c, w)$ and wealth $f(k, s)$ are both upper-semicontinuous.

A.5. Wealth $f(k, s)$ is non-decreasing in investment k for each s.

Assumptions

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

A.2. Utility is supermodular ($U_{12} \geq 0$ if it is differentiable).

A.3. Utility is strictly concave in consumption c.

A.4. Utility $U(c, w)$ and wealth $f(k, s)$ are both upper-semicontinuous.

A.5. Wealth $f(k, s)$ is non-decreasing in investment k for each s.

A.4. covers both the stepping-stone and the other production functions listed above.

Discuss super-modularity.

- We could have assumed u was jointly concave. This leads to comparative statics through the implicit function theorem. Weird stuff.
- Super-modularity is a different approach to comparative statics.
- Super-modularity does not imply that c and k are complements. $U(c, k) = c + k$ is super-modular.
- CES utility. For which values of elasticity of substitution σ is u super-modular?

Back to Questions!

Existence of solutions

Theorem 1. For all $w \geq 0$, $\pi(w) \neq \emptyset$. Furthermore, π is upper-hemicontinuous at every continuity point of f.

Proof Since U is increasing in w, $U\big(c, f(k, s)\big)$ is upper semi-continuous in (c, k) for every s. So, therefore, is $V(c, k)$, and maxima of usc functions on compact sets exist. The rest is the Berge Maximum Theorem.

Theorem 1. For all $w \geq 0$, $\pi(w) \neq \emptyset$. Furthermore, π is upper-hemicontinuous at every continuity point of f.

Proof Since U is increasing in w, $U(c, f(k, s))$ is upper semi-continuous in (c, k) for every s. So, therefore, is $V(c, k)$, and maxima of usc functions on compact sets exist. The rest is the Berge Maximum Theorem.

Proof of semi-continuity of the composition: Suppress s since the theorem's proof takes it as fixed. Let $(c_n, k_n) \to (c, k)$. We have that $\limsup_n f(k_n) \leq f(k)$. Since U is increasing, $U(c_n, f(k_n)) \leq U(c_n, f(k) + \epsilon)$ for any $\epsilon > 0$ and all sufficiently large n. Consequently, and due to the upper semi-continuity of U,

$$\limsup_n U(c_n, f(k_n)) \leq \limsup_n U(c_n, f(k) + \epsilon) \leq U(c, f(k) + \epsilon).$$

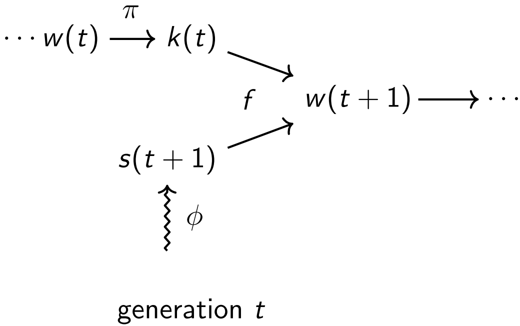
Since this holds for all $\epsilon > 0$, $\limsup_n U(c_n, f(k_n)) \leq U(c, f(k))$, which was to be proved. □

Comparative statics

What does π look like?

- π is a non-decreasing function except at isolated points.
- At those points it jumps up.
- At a jump point w, the higher point is always in the graph of π. The lower point may or may not be.
- If the lower point is in the graph, π is uhc at w.

The evolution of wealth



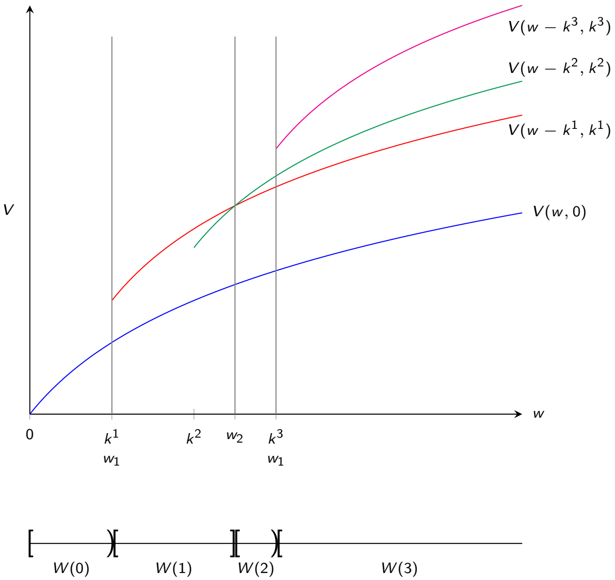
The deterministic stepping-stone model

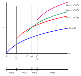
In this model there are a finite number of distinct investment levels. Some of these are self-sustaining, fixed points in that if a family makes an investment at that level, the return is such that the child will want to and be able to invest the same amount for the next generation.

Other investment levels are transitory. Low and high investment levels may be self-sustaining. Dynasties with enough more wealth than the low level make mid-level investments, thereby "stepping up" through the generations to the high level. Families with more wealth than the low level, but not enough more, make mid-level investments, thereby "stepping down" through the generations to the low level.

This model may exhibit poverty traps, multiple inescapable steady-states.

The deterministic stepping-stone model

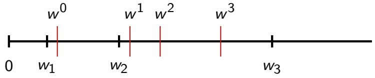




Failure of uhc at w_1, w_3 because the constraint correspondence is not lhc.

Dynamics

Another example



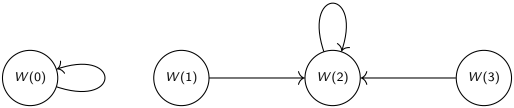
- A family with initial wealth less than w_1 invests 0. the next generation has wealth w^0.
- A family with initial wealth between w_1 and w_2 invests k^1. The next generation has wealth w^1.
- A family with wealth between w_2 and w_4 invests k^2, and all subsequent generations have wealth w^2 and invest k^2.
- A family with wealth $w \geq w_4$ invest k^2 and ha 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 wealth paths are monotone.

Dynamics

For $n = 0, \ldots, N$, define w_n and w^n as above. Define the intervals $W(0) = [0, w_1\}$ and $W(i) = [w_i, w_{i+1})]$ for $i \leq N$, with $w_{N+1} = +\infty$, where the last)] is open and the remaining right delimiters are either open or closed. Assume that no w^i is on an interval boundary. Dynamics can be described by a graph.

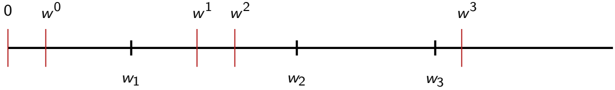
Example:

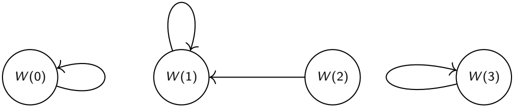


$w^1, w^2, w^3 \in W(2)$. $W(0)$ is a poverty trap. For initial wealth exceeding w_1, wealth converges to w^2. This graph is the deterministic graph of the dynamics.

Dynamics

Another possibility





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

- What do we learn from this model?
- What's missing?

- Poverty traps
- Interaction across families.
- N.B. This is what both Loury and Becker-Tomes do.

The stochastic stepping-stone model

The stochastic stepping-stone model

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

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

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

The stochastic stepping-stone model

The optimal policy

Key Fact: π looks just as it did before.

- nondecreasing
- jumps only up
- uhc from the right.

The stochastic stepping stone model

The Markov process

All selections $\dot{\pi}$ from π differ from each other only at the wealths w_i, where both k^{i-1} and k^i are optimal. Each selection describes a Markov process. For measurable $A \subset W$,

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

- 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 $\tilde{w}^i \in W(j)$.

$$p_{ij} = \int_{W(j)} g^i(w)\, dw.$$

We can define social classes by the $W(i)$.

All selections $\hat{\pi}$ from π differ from each other only at the wealths w_{i*} where both k^{i-1} and k^i are optimal. Each selection describes a Markov process. For measurable $A \subset W$,

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

- 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 $\tilde{w}^i \in W(j)$.

$$p_{ij} = \int_{W(j)} g^i(w)\, dw.$$

We can define social classes by the $W(i)$.

What do we want to know about this Markov process?

- Formal questions:
  - Invariant distribution — Existence and uniqueness?
  - Convergence rates.
- Substantive questions:
  - Inequality — shape of invariant distributions.
  - Mobility — convergence rates, transition probabilities.

The stochastic stepping stone model

The Markov process

The Markov processes derived with selections from π are not "textbook" because of the jumps at the w_i. Nonetheless all such processes have a unique invariant distribution, which does not depend on the selection. Furthermore, the generation t marginal distributions of wealth converge weakly to this invariant distribution from any initial condition.

Key idea: The process which records generation t's $W(i)$ is Markov.

Choose a selection $\dot{\pi}$ and define the $W(i)$ by correctly assigning the endpoints to the proper class. Then for $w_t \in W(i)$,

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

The stochastic stepping stone model

Therefore,

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

and

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

ν_t matters only through the probs $\big((\nu_t(W(1), \ldots, \nu_t(W(N))\big)$. Define $p_{ij} = \int_{W(j)} g^i(w)\, dw$, the probability of moving from $W(i)$ to $W(j)$. Then $[p_{ij}]$ is a Markov matrix, and

$$\big(\nu_{t+1}(W(1), \ldots, \nu_{t+1}(W(N))\big) = \big(\nu_t(W(1), \ldots, \nu_t(W(N))\big) \cdot [p_{ij}].$$

The stochastic stepping stone model

If $[p_{ij}]$ is irreducible, then it has an invariant probability vector q^* and the stepping stone process has an invariant distribution ν^* on $\mathbf{R}_+$ given by

$$\nu^*(A) = \sum_k q_k^* \int_A g^k(w)\, dw.$$

If $[p_{ij}]$ is primitive, then in addition, the sequence $(\nu_t(W(1)), \ldots, \nu_t(W(N)))$ converges to q^* and so the sequence of distributions ν_t converges to ν^*.

Finally, note that the p_{ij} do not depend upon the choice of $\dot{\pi}$ because the probability of drawing a multi-valued w is 0.

If $[p_{ij}]$ is irreducible, then it has an invariant probability vector q^* and the stepping stone process has an invariant distribution ν^* on $\mathbf{R}_+$ given by

$$\nu^*(A) = \sum_k q_k^* \int_A g^k(w)\, dw.$$

If $[p_{ij}]$ is primitive, then in addition, the sequence $(\nu_t(W(1)), \ldots, \nu_t(W(N)))$ converges to q^* and so the sequence of distributions ν_t converges to ν^*.

Finally, note that the p_{ij} do not depend upon the choice of $\hat{\tau}$ because the probability of drawing a multi-valued w is 0.

In fact, this shows that ν_t converges to its limit ν^* in the variation norm.

The stochastic stepping-stone model

Assumption: $[p_{ij}]$ is irreducible.

Assumption: There is a g^i which is strictly positive on some open interval in $W(i)$.

The purpose of the first assumption is to make the process irreducible. The second makes it aperiodic.

Assumption: $[p_{ij}]$ is irreducible.

Assumption: There is a g^j which is strictly positive on some open interval in $W(i)$.

The purpose of the first assumption is to make the process irreducible. The second makes it aperiodic.

- Because there are only a finite number of steps, the matrix $[p]$ is finite, so invariant distributions exist.
- irreducibility makes it unique.
- primitivity guarantees convergence.

Measures of mobility

- A possible measure of mobility is the expected length of time a dynasty remains in a given class.

$$E\{\text{duration of } W(i)\} = \frac{1}{1 - p_{ii}} - 1 = \frac{p_{ii}}{1 - p_{ii}}.$$

- The second largest eigenvalue magnitude measures convergence rates. Bottleneck inequalities.

Measures of mobility

- A possible measure of mobility is the expected length of time a dynasty remains in a given class.

$$E\{\text{duration of } W(i)\} = \frac{1}{1-p_{ii}} - 1 = \frac{p_{ii}}{1-p_{ii}}.$$

- The second largest eigenvalue magnitude measures convergence rates. Bottleneck inequalities.

There is no necessary connection between inequality and mobility. Choose a transition probability P with invariant distribution q^*. Then q^* is invariant for $\alpha I + (1-\alpha)P$, but duration can be made arbitrarily long by choosing α near 1.

Notice that the basins of attraction are not "special" with respect to persistence. The g^i can be such that the probability of leaving a deterministic attractor is more than the probability of leaving some other region in its basin of attraction. But for densities with sufficiently small variance this will not be true.

Poverty Traps

What does a poverty trap look like in a stochastic model?

- The states $W(i)$ of $[p]$ can be partitioned into $K+1$ groups: Group 0 states are transient. Groups 1 through K are recurrence classes. Once a group is entered, it is never left. Iff $[p]$ is irreducible, $K=1$.
- Some states or groups of states may be meta-stable: They may be very durable, or they may be entered very frequently.

Poverty Traps

What does a poverty trap look like in a stochastic model?

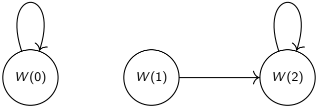
- The states $W(i)$ of $[\rho]$ can be partitioned into $K+1$ groups: Group 0 states are transient. Groups 1 through K are recurrence classes. Once a group is entered, it is never left. Iff $[\rho]$ is irreducible, $K=1$.
- Some states or groups of states may be meta-stable: They may be very durable, or they may be entered very frequently.

- Reducibility is not interesting. An insurmountable barrier??
- Two ways to talk about meta-stability: How easy is it to get stuck, and where does the process spend most of its time.

Poverty Traps

Natural candidates for poverty traps are the low-wealth stationary states of the deterministic model. What happens if we add a little noise?

Stochastic Stability of stationary states.



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

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. No matter how perturbations are introduced, if they are all of the same order, only state 3 survives in the limit. $\nu(i) = \lim_{\epsilon\to 0} \lim_{T\to\infty} \Pr\{s(t) = i | s(0) = j\}$.

Stochastically stable states are identified by introducing a family of random perturbations into a deterministic system in a reasonable way, and finding the limit of invariant distributions as the perturbations shrink to 0.

Stochastic stability

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. No matter how perturbations are introduced, if they are all of the same order, only state 3 survives in the limit. $\nu(i) = \lim_{\epsilon \to 0} \lim_{T \to \infty} \Pr\{s(t) = i | s(0) = j\}$.

Stochastically stable states are identified by introducing a family of random perturbations into a deterministic system in a reasonable way, and finding the limit of invariant distributions as the perturbations shrink to 0.

Reversing the order of the double limits gives a very different answer. Notice too that the answer is independent of j.

Stochastic stability

Define a rich class of examples. Suppose $h : \mathbf{R} \to \mathbf{R}_+$ is C^2 at its minimum, 0, and $h''(0) > 0$.

- $\tilde{w}_n = \max\{w^n + s, 0\}$
- s has a density $h_\lambda(s)$ on $\mathbf{R}$, where

$$h_\lambda(s) = \exp\{-\lambda h(s)\}/Z(\lambda)$$

  where $Z(\lambda)$ is a normalizing constant.

  For the rest of this section, $[p]$ and everything associated with it will be parametrized by λ.
- Then

$$p_{ij}(\lambda) = \int_{s \in \{z - w^i, z \in W(j)\}} h_\lambda(s)\, ds \equiv \int_{w_j(\lambda)}^{w_{j+1}(\lambda)} \exp -\lambda h^i(s)\, ds$$

  for any $w \in W(i)$. (Take $W(0) = (-\infty, w_1(\lambda)]$.)

Stochastic stability

Define a rich class of examples. Suppose $h : \mathbf{R} \to \mathbf{R}_+$ is C^2 at its minimum, 0, and $h''(0) > 0$.

- $\tilde{w}_s = \max\{w^* + s, 0\}$
- s has a density $h_\lambda(s)$ on $\mathbf{R}$, where

  $$h_\lambda(s) = \exp\{-\lambda h(s)\}/Z(\lambda)$$

  where $Z(\lambda)$ is a normalizing constant.

  For the rest of this section, $[p]$ and everything associated with it will be parametrized by λ.
- Then

  $$p_{ij}(\lambda) = \int_{s \in \{z - w^i, z \in W(j)\}} h_\lambda(s)\, ds = \int_{w_j(\lambda)}^{w_{j+1}(\lambda)} \exp -\lambda h^i(s)\, ds$$

  for any $w \in W(i)$. (Take $W(0) = (-\infty, w_1(\lambda)]$.)

$$Z(\lambda) \approx \sqrt{2\pi/\lambda h''(0)}$$

Recall $h^i(s) = h(s - w^i)$.

Note that $W(i)$ now depends upon λ.

Stochastic stability

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 $w_i(\lambda)$.

A.6. The boundaries $w_i(\lambda)$ converge to the deterministic boundaries as $\lambda \to \infty$.

A sufficient condition for this is that U is continuous.

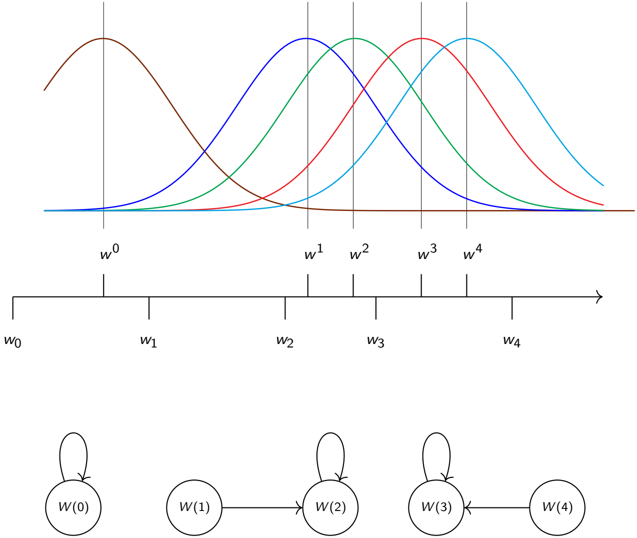
Write $w_i(\lambda)$ and note that for each i, $\lim_{\lambda \to \infty} w_i(\lambda) \to w_i$.

Stochastic Stability

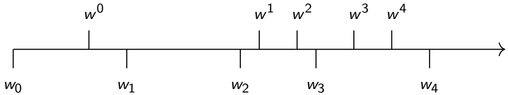
Denote by ν_λ the invariant distribution of $[p(\lambda)]$.

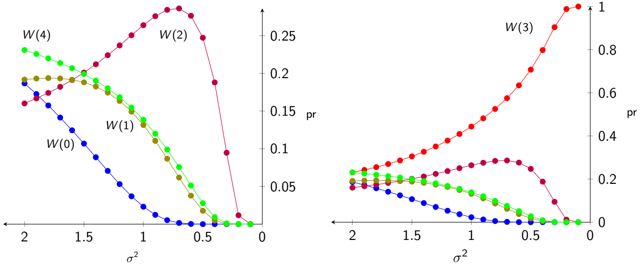
Theorem. For every state $W(i)$, $\lim_{\lambda \to \infty} \nu_\lambda(W(i)) = 0$ or 1. If the limit is 1, then $W(i)$ is a stationary point of the deterministic dynamic. Typically there will be only one state $W(i)$ with limit 1. The stochastically stable attractor is unique.

The stochastically stable attractor is not determined by the graph. That is, there are graphs for which the identity of the unique stochastically stable state depends on the shape of h.

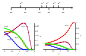

w^0
w^1
w^2
w^3
w^4
w_0
w_1
w_2
w_3
w_4
$W(0)$
$W(1)$
$W(2)$
$W(3)$
$W(4)$

Stochastic stability





Invariant distribution probabilities, $h(x) = x^2$ and $\lambda = 1/2\sigma^2$.



Invariant distribution probabilities, $h(x) = x^2$ and $\lambda = 1/2\sigma^2$.

It's so nice when computations are consistent with a theorem.

Stochastic Stability

What do we learn?

- The deterministic model does not give a good description of long-run behavior of the model with even a modest amount of noise.
- $$\lim_{T\to\infty} \frac{1}{T}\#\big\{t : w_t \in W(i)\big\} = \nu_\lambda\big(W(i)\big).$$

  This does not capture poverty-trap behavior.
- Computation shows that the limit picture sheds light on the finite-λ picture.

Poverty Traps

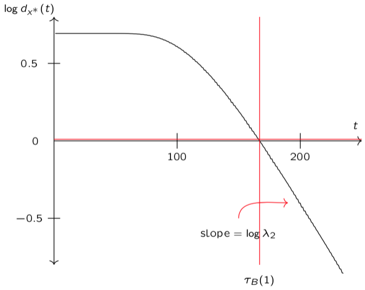
- In light of this analysis, what should we think of poverty traps in deterministic models?
- Typically, the odds ratio $\nu_\lambda\big(W(i)\big)/\nu_\lambda\big(W(j)\big)$ converges to either 0 or ∞. $W(i)$ is more stable than $W(j)$ iff $\nu_\lambda\big(W(i)\big)/\nu_\lambda\big(W(j)\big) = \infty$. States can be ordered by the "more stable than" relation. This can give a coarse description of traps.

Measuring Mobility

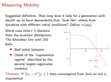
Suggested definition. How long does it take for a generation with wealth w_0 to have descendents that "look like" others from dynasties with different initial conditions? Define $\pi_t(w_0)$.

Worst-case time-t l_1-distance from the invariant distribution. The Ehrenfest Urn with 150 balls.

- Bad initial behavior
- Onset of the "exponential regime" described by the second largest eigenvalue modulus.



Theorem: If $||\pi_t - \pi^*||_1 \leq 1$ then convergence from here on out is exponential.



- Two points:
  - Expected pic on left vs weird possibility on the right.
  - Why l_{1x} is the right norm.
- Eigenvalues are $1, \lambda_2, \dots, \lambda_N$ ordered by decreasing modulus. Slope of the bottom piece is $1 - |\lambda_2|$.

References I

Arrow, Kenneth. 1973. "Rawls Principle of Just Saving." *Swedish Journal of Economics* 75: 323–335.

Banerjee, Abhijit V. and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101 (2): 274–298.

Becker, Gary and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy* 87 (6): 1153–1189.

Bernheim, B. Douglas and Debraj Ray. 1987. "Economic growth with intergenerational altruism." *Review of Economic Studies* 54: 227–43.

Loury, Glenn. 1981. "Intergenerational transfers and the distribution of earnings." Econometrica 49 (4): 843–67.