1 The Steady-State Distribution of Permanent Income

Our economy begins at date 1 with a population of size L_0 living at locations indexed by ℓ.[1] Immediately prior to date 1, all agents have a value of 'permanent income' $\wp_{0,\ell} = 1$.

Here we derive the dynamics of the distribution of $\wp_{t,\ell}$ across locations, and the limiting distribution as $t \to \infty$, under the assumption that each period the agent at each location receives an idiosyncratic shock to $\wp = \log \wp$ until 'death,' at which point the agent at that location is replaced by a 'newborn.'[2]

1. The first thing that happens at date τ is that agents living at a randomly selected constant proportion $\mho$ of locations die and are replaced by newborns; $\{\mho_\tau\}$ is the set of such locations. Locations where death does not occur are in the set $\{\Lambda_\tau\}$ (death has been canceled - for now), and of course $\Lambda = 1 - \mho$ and $0 \leq \mho < 1$.

2. The initial permanent income of newborns is $\wp = 1$ (formally, $\wp_{\tau,\ell} = 1 \ \forall \ \ell \ \in \ \{\mho_\tau\}$)

3. Next, every location is subjected to a mean-one lognormally distributed shock

$$\psi \equiv \log \Psi \sim \mathcal{N}(\mu_\psi, \sigma_\psi^2) \tag{1}$$

where $\mu_\psi = -\sigma_\psi^2/2$. That is, for all $\ell \in \Lambda_\tau$,

$$\wp_\tau = \wp_{\tau-1} + \psi_\tau \tag{2}$$

while for $\ell \in \mho$

$$\wp_\tau = \psi_\tau \tag{3}$$

---

[1]These will be represented computationally by a simulated population of integer size L_0 in which each member of the population at any date τ occupies a slot at integer location ℓ).

[2]This is a discrete version of the 'perpetual youth' model of Blanchard (1985).

4. Under these assumptions, $\mathbb{E}_{\{\Lambda_\tau\}}[\wp_{\tau+1,\ell}] = \mathbb{E}_{\{\Lambda_\tau\}}[\wp_{\tau,\ell}\Psi_{\tau+1,\ell}] = \wp_{\tau,\ell}$ (where $\mathbb{E}_{\{[\mathrm{set}]_\tau\}}$ denotes the expectation of the value of its argument at the *locations* occupied in period τ for the agents in the set, while $\mathbb{E}[\wp_{\tau,\ell}]$ denotes the expectation across all persons and locations)

5. Because dying agents are selected at random, the expectation of $\wp_{\tau,\ell}$ for period-τ survivors must match the expectation of $\wp_\tau$ for the entire population:

$$\mathbb{E}_{\{\Lambda_\tau\}}[\wp_{\tau,\ell}] = \mathbb{E}[\wp_{\tau,\ell}] \qquad (4)$$

Thus for an economy in which the average value of $\wp$ is 1 in period τ, the population-mean level of $\wp$ in period $\tau + 1$ will remain at 1:

$$\mathbb{E}[\wp_{\tau+1}] = \mho \overbrace{\mathbb{E}_{\{\Lambda_\tau\}}[\wp_{\tau,\ell}\Psi_{\tau+1,\ell}]}^{\text{survivors have } \wp = 1} + (1 - \mho) \overbrace{\mathbb{E}_{\{\mho_\tau\}}[1 \times \Psi_{\tau+1,\ell}]}^{\text{newborns have expected } \wp = 1}$$
$$= \mho + (1 - \mho)$$

At t we denote the locations occupied by persons who have been alive for exactly m periods as the set $\{\ell_t[m]\}$. Agents for whom $m = 1$ will have experienced only a single shock, the one at the beginning of the period; for $m > 1$ the agents at those locations (that is, for $\ell \in \{\ell_t\}[m]$),

$$\wp_{t,\ell} = \sum_{i=1}^{m} \psi_{i,\ell}. \qquad (5)$$

For legibility defining $\{\wp_t[m]\} = \{\wp_{t,\{\ell_t[m]\}}\}$ as the values of $\wp$ across the set of agents who in period t have been alive for exactly m periods, the independence of the ψ's means that for every $\ell \in \{\ell_t[m]\}$,

$$\wp_{t,\ell}[m] \sim \mathcal{N}(m\mu_\psi, m\sigma_\psi^2) \qquad (6)$$

because such agents have experienced m IID permanent

shocks (recall that the variance of the sum of independent variables is equal to the sum of their variances).

So the PDF of an agent who has been alive for m periods is the PDF of $\mathcal{N}(m\mu_\psi, m\sigma_\psi^2)$,

$$\text{PDF}(\wp_t[m]) = \left(\sigma_\psi\sqrt{m2\pi}\right)^{-1}\exp\left(-\frac{(\wp - \mu_\psi m)^2}{2m\sigma_\psi^2}\right) \quad (7)$$

The set of all values of $\wp$ in period t will be the union of the $\wp$'s of persons in the sets of people who have been alive for each possible number of periods (1, for newborns):

$$\{\wp_t\} \;=\; \bigcup_{m=1}^{t} \{\wp_t[m]\} \tag{8}$$

but since each of the individual $\wp_{t,\ell}[m]$ objects is identical (Gaussians), the set as a whole will be distributed according to that same distribution with a weight equal to the size of its population. The cohort's population mass starts at size $(1 - \Lambda)$ and shrinks by a factor Λ each period:

$$k_m = (1 - \Lambda)\Lambda^{m-1} \tag{9}$$

and the total population mass after t periods is the sum of the masses of the living cohorts:

$$K_t = \sum_{m=1}^{t} k_m$$

$$= (1 - \Lambda)\sum_{m=1}^{t} \Lambda^{i-1}$$

$$= (1 - \Lambda)\left(\frac{1 - \Lambda^t}{1 - \Lambda}\right)$$

$$= 1 - \Lambda^t$$

and we use ϱ for the proportion of the population of the population at date t constituted by agents who have lived m periods:

$$\varrho_{t,m} = k_m / K_t \tag{10}$$

$$= \left(\frac{(1 - \Lambda)\Lambda^{m-1}}{1 - \Lambda^t}\right)$$

For example, in the first period of operation of the economy, $t = m = 1$ and $\varrho_{t,m} = 1$.

Thus $\wp_t$ is distributed as a weighted sum of gaussians:

$$\wp_t \sim \sum_{m=1}^{t} \varrho_{t,m} \mathcal{N}(m\mu_\psi, m\sigma_\psi^2) \qquad (11)$$

We are finally in a position to derive the results we need.

The set of values of permanent income at all locations at date $t, \{\wp_t\}$, is a weighted collection of normally distributed random variables, with weights corresponding to the population masses of the cohorts of persons who have lived m periods ($\varrho_{t,m}$). If the first generation had been introduced with a population size of $1 - \Lambda$ the PDF of the entire distribution would be the weighted sum of the PDFs of the constituent populations:

$$\text{PDF}(\wp_t) = \frac{\sum_{m=1}^{t} \Lambda^{m-1} \exp\left(-\frac{(\wp_m - m\mu_\psi)^2}{2m\sigma_\psi^2}\right)}{\sigma_\psi \sqrt{2\pi} \underbrace{\sum_{m=1}^{t} \Lambda^m}_{(1-\Lambda^{t-1})/(1-\Lambda)}} + \qquad (12)$$

$$= \frac{(1-\Lambda) \sum_{m=1}^{t} \Lambda^{m-1} \exp\left(-\frac{(\wp_t - m\mu_\psi)^2}{2m\sigma_\psi^2}\right)}{(1-\Lambda^{t-1}) \, \sigma_\psi \sqrt{2\pi}} \qquad (13)$$

$$= \sum_{m=1}^{t} (1-\Lambda)\Lambda^{m-1} \text{PDF}[\mathcal{N}(m\mu_\psi, m\sigma_\psi), \wp] \qquad (14)$$

so the probability mass associated with the tth (oldest) cohort would be

$$\mathcal{L}(\wp_t[t]) = (1-\Lambda)\Lambda^{t-1} \qquad (15)$$

but we actually introduced the first cohort with a population size of 1 rather than $(1-\Lambda)$, so its remaining population mass

will instead be Λ^{t-1}. So the correct expression would be

$$\mathrm{PDF}(\wp_t) = \Lambda^{t-1}\mathrm{PDF}[\mathcal{N}(t\mu_\psi, t\sigma_\wp), \wp] + \sum_{m=1}^{t-1}(1-\Lambda)\Lambda^{m-1}\mathrm{PDF}[\mathcal{N}(m\mu_\psi, m\sigma_\psi), \wp]$$

(16)

Our objective is to determine how far to continue the summation in (16), with the natural criterion for judgment being that we want to proceed up to some number of periods $FutPer$ such that the proportion of the population constituted by the "founder" generation is negligible.

If we are conducting a companion Monte Carlo exercise with a population size of N, a reasonable choice would to calculate the t such that

$$\Lambda^t N = 1$$
$$t \log \Lambda + \log N = 0$$
$$t = -\frac{\log N}{\log \Lambda}$$

That is, we are capturing the correct distribution of the probabilty mass at the t such that only one remaining member of the original population is expected to remain alive.

2 Simulation Algorithm

This describes the simulation algorithm.

2.1 A Useful Trick

We will repeatedly be using a trick that greatly reduces the number of agents necessary to get well-behaved results. If there is a normally distributed random variable ψ that will be drawn for a population of size $K \in \mathbb{N}^+$, we first construct the $K-$point equiprobable approximation to the normal dis-

tribution, where the probability mass of each element is $1/K$ and the value at location ℓ is $\psi[\ell]$.

The trick is that, rather than actually drawing a value of ψ independently for each ℓ, instead we just randomly reshuffle the order of the elements in $\boldsymbol{\psi}$ (the vector of values of ψ. In the limit as $k \uparrow \infty$ this approaches the asymptote defined by the true distribution. But in a finite population, we guarantee that the population-mean level and variance of the collection of draws of $\boldsymbol{\psi}$ matches the theoretical values of these objects:

$$\mu = (1/K) \sum_{\ell=1}^{K} \psi_\ell \tag{17}$$

$$\sigma_\psi^2 = (1/K) \sum_{\ell=1}^{K} (\psi_\ell - \mu)^2 \tag{18}$$

Call a version of $\boldsymbol{\psi}$ that has been reshuffled at date τ $\tilde{\boldsymbol{\psi}}$, and define

$$\tilde{\boldsymbol{\psi}}_\Lambda = \text{reshuffled } \psi \text{ list of size of suvivors} \tag{19}$$

$$\tilde{\boldsymbol{\psi}}_\mho = \text{reshuffled } \psi \text{ list of size of newborns} \tag{20}$$

2.2 Monte Carlo Simulation Method

The first step is to determine how many agents to simulate.

We choose a value of $\mho$ so that $1/\mho$ is an integer. Our population size needs to be at least $\Xi = 3/\mho$ because this guarantees that there will be at least three newborns in the population, which allows us to use the trick above to construct the minimal sensible approximation to a normal distribution - with three points each of which is realized with probability one third.

To simulate a population that has existed for t periods, loop over the following steps t times:

1. Select a random set $\{\mho_\tau\}$, which corresponds to an integer number Ξ locations; these are the locations at which the agents will be replaced

2. For the locations in $\{\mho_\tau\}$, set the values of $\wp$ to $\tilde{\boldsymbol{\psi}}_\mho$

3. For the locations in $\{\Lambda_\tau\}$ (the complement of $\mho_\tau$), add to the existing value of $\wp$ the list of values $\tilde{\boldsymbol{\psi}}_\mho$

   - Note that in a finite sample, the covariance between $\wp$ and $\tilde{\boldsymbol{\psi}}_\Lambda$ will not be zero, which means that the sample mean and variance will not be exactly equal to their expectations.

4. Adjust the distribution (see the algorithm for details) so that its mean and variance exactly match the theoretical expectation

3 Discretizing the Transition Matrix

Soudjani and Abate (2014) and Soudjani and Abate (2015) show how to bound the errors that will be associated with discrete approximations of Markov processes on continuous state spaces in discrete time. Notation and setup here follow them.

The setup is identical to Soudjani and Abate (2014)

3.1 The Continuous-State-Space Problem We Want to Discretize

Consider a discrete time Markov process $\mathcal{M}$ defined over a continuous state space $\mathcal{S}$ which is characterized by a pair $(\mathcal{S}, \mathfrak{T}_s)$. We denote by $(\mathcal{S}, \mathcal{B}(\mathcal{S}), \mathcal{P})$ the probability structure on $\mathcal{S}$, with $\mathcal{B}(\mathcal{S})$ being the associated $\sigma-$algebra, and $\mathcal{P}$ is

$\mathfrak{T}$ for 'transition'

a probability measure to be defined shortly. $\mathfrak{T}_s$ is a conditional stochastic kernel that assigns to each point $s \in \mathcal{S}$ a probability measure $\mathfrak{T}_s(\bullet|s)$ so that for any measurable set $A \in \mathcal{B}(\mathcal{S}), \mathcal{P}(s(1) \in A|s(0) = s) \equiv \int_A \mathfrak{T}_s(d\bar{s}|s)$ ($s(\tau)$ is a measurement at date τ). We assume that the stochastic kernel $\mathfrak{T}$ admits a density function $\mathfrak{t}_s$, namely $\mathfrak{T}_s(d\bar{s}|s) = \mathfrak{t}_s(\bar{s}|s)d\bar{s}$.[3]

For generality, Soudjani and Abate (2014, 2015) assume that the initial state of the Markov process $\mathcal{M}$ is random and distributed according to the density function $\pi_0 : \mathcal{S} \to \mathbb{R}^{\geq 0}$. We will instead start with a well-defined initial density function (also π_0), because this makes debugging the code easier. The state distribution of $\mathcal{M}$ at time $\tau \in \mathbb{N} \doteq \{1, 2, 3, ...\}$ is characterized by a density function $\pi_\tau : \mathcal{S} \to \mathbb{R}^{\geq 0}$ such that for all $A \in \mathcal{B}(\mathcal{S})$,

$$\mathbb{P}(s(\tau) \in A) \equiv \int_A \pi_\tau(s)ds \qquad (21)$$

where $\mathbb{P}$ indicates the probability associated to events over the product space $\mathcal{S}^{\tau+1}$ with elements $\boldsymbol{s} = [s(0), s(1), \ldots, s(\tau)]$, where $\boldsymbol{s}$ indicates a vector of individual values of s (the same notation applies hereafter).

Because the transition process is assumed to be time invariant, the conditional state probability functions $\pi_\tau(\bullet)$ can be computed recursively:

$$\pi_{\tau+1}(\bar{s}) = \int_{\mathcal{S}} \mathfrak{t}_s(\bar{s}|s)\pi_\tau(s)ds \qquad \forall \, \bar{s} \in \mathcal{S} \qquad (22)$$

We make further assumptions that correspond to Assumptions 1 and 2 in Soudjani and Abate (2014).

---

[3]This rules out probability mass points; it seems likely that they could be incorporated by an extension of the argument that associates a Dirac δ function with each such point, so long as all such points are always incorporated in A.

Assumption 1. *For given sets $\Gamma \subset \mathcal{S}^2$ and $\Lambda_0 \subset \mathcal{S}$, there exist positive constants ϵ and ε_0, such that $\mathsf{t}_s(\bar{s}|s)$ and $\pi_0(s)$ satisfy the following conditions:*

$$\mathsf{t}_s(\bar{s}|s) \leq \epsilon \quad \forall (s, \bar{s}) \in \mathcal{S}\backslash\Gamma, \text{ and } \pi_0(s) \leq \varepsilon_0. \qquad (23)$$

(Note that this rules out point masses, for which the $\mathsf{t}_s(\bar{s}|s)$ and ε_0 would be unbounded. However, it is hard to see anywhere that the logic below would fail if there were point masses as long as they were always in the interior of Γ.)

Assumption 2. *The density functions $\pi_0(s)$ and $\mathsf{t}(\bar{s}|s)$ are (globally) Lipschitz continuous, namely there exist finite constants λ_0 and λ_f such that the following Lipschitz continuity assumptions hold:*

$$|\pi_0(s) - \pi_0(s')| \leq \lambda_0, \forall s, s' \in \Gamma_0, \qquad (24)$$
$$|\mathsf{t}(\bar{s}|s) - \mathsf{t}(\bar{s}'|s)| \leq \lambda_f, \forall s, \bar{s}, \bar{s}' \in \mathcal{S}.$$

That is,

1. The difference in probability densities across points in the initial distribution is bounded (by λ_0); no density is infinitely more likely than another (no point masses).

2. From any point s, the difference in the probability of transitioning to point $\bar{s}$ cannot be arbitrarily larger than the probability of transitioning to point $\bar{s}'$.

Under these assumptions, their Theorem 3 holds:

Theorem 3. *The functions $\pi_\tau(\bullet)$ can be approximated by piecewise constant functions $\hat{\pi}(\bullet)$, defined as*

$$\hat{\pi}_\tau(s) = \sum_{k=1}^n \frac{q_\tau(k)}{\mathcal{L}(A_k)} \mathbb{I}_{A_k}(s) \quad \forall \tau \in \mathbb{N}^+ \qquad (25)$$

where $\mathbb{I}_A(\bullet)$ is the indicator function of a set $A \subseteq \mathcal{S}$ (that is, $\mathbb{I}_A(s) = 1$ if $s \in A$; else $\mathbb{I}_A(s) = 0$).

We consider a finite Markov chain $\mathcal{M}_p$ with $n+1$ discrete abstract states over this space characterized by the transition probability matrix $\hat{P} = [\hat{P}_{ij}]$ where the probability of jumping from state i to state j ($\hat{P}_{i,j}$) is computed as:[4]

$$\hat{P}_{i,j} = \left(\frac{1}{\mathcal{L}(A_i)}\right) \int_{A_i} \left(\int_{A_j} \mathsf{t}_s(\bar{s}|s)d\bar{s}\right) ds \quad \forall j \in \mathbb{N}_n \tag{26}$$

$$\hat{P}_{(n+1),j} = \delta_{(n+1),j}$$

for all $j \in \mathbb{N}_{(n+1)}$, and where $\delta_{(n+1),j}$ is the Kronecker delta function (the abstract state $(n+1)$ of $\mathcal{M}_p$ is absorbing)[5] and $\mathcal{L}(\bullet)$ denotes the Lebesgue measure of a set.[6] That is, as the process evolves, a measurable (but arbitrarily small) proportion of the probability mass is in the portion not characterized by the transition matrix; but for longer (fixed) time intervals the matrix changes to accommodate the ever-expanding universe of allowed values of $A \in \mathcal{S}$).

The initial distribution of $\mathcal{M}_p$ is given by the probability mass function (pmf) $\boldsymbol{q}_0 = [q_0(1), q_0(2), \ldots, q_0(n+1)]$ obtained from π_0 as the average value of q_0 in each of the i intervals: $q_0(i) = \int_{A_i} s\pi_0(s)ds$, $\forall i \in \mathbb{N}_n$.[7] Under Soudjani and Abate (2014)'s further assumptions, the pmf associated

$\hat{P}_{(n+1),j}$ captures the probability mass of all values of $s \in \mathcal{S}$ but not in $A = \cup A_i$ - values above or below the cutoff values that defined A as a bounded subset of the original unbounded space

---

[4]The order of integration is reversed compared to Soudjani and Abate (2014), to map better the computational procedure. We integrate over j first, then i.

[5]absorbing: when you leave the bounded space you never return; but, why doesn't all the mass eventually 'leak out' through this hole? answer: we are considering this calculation as occurring over a fixed time interval, and the key result is that *over the chosen time interval* the amount of mass in the excluded area is bounded by the Lipschitz conditions. As we consider longer time intervals, the boundaries of the space over which the computations are done must be increased (thanks to the recursion Γ evolves also to incorporate enough of the missing mass for the limiting results to hold. (This is the most subtle part of the argument)).

[6]If there are any point masses in the set, they are captured as a case where $\underline{A} = \bar{A}$; that is, all point masses are allocated an "interval" of measure zero at the location of the point mass.

[7]If there is a point mass at A_i, $q_0(i) = \int_{A_i} \delta(i)\pi_{0,i} = \mathbb{P}_0(s) > 0$ where δ is the Dirac δ).

to the state distribution of $\mathcal{M}_p$ at time $\tau \in \mathbb{N} \doteq \{1, 2, 3, ...\}$ can be computed as $\boldsymbol{q}_\tau = \boldsymbol{q}_0 \hat{P}^\tau$. $\boldsymbol{q}_\tau$ is used to construct an approximation function, denoted by $\hat{\pi}_\tau$, of the density function π_τ, where $\hat{\pi}_\tau$ is a piecewise constant approximation (with values that are the entries of the pmf $\boldsymbol{q}_\tau$ normalized by the Lebesgue measure of the associated partition set) of the original density function π_τ.

The corresponding approximation points for each Markov bin i are given by

$$s_i = \int_{\underline{A}_i}^{\bar{A}_i} s\pi(s)ds \tag{27}$$

and the vector of such points is $\boldsymbol{s}$.

3.2 Interpretation and Implementation

It is clear from the foregoing results that, so long as an invariant limiting distribution exists, that distribution should be approximable to any desired degree of accuracy by using the methods described above (so long as the A_i regions can be arbitrarily small). (See section 8 of Soudjani and Abate (2015) for a proof and details).

For our model's specification of permanent income, we have an analytical expression for the limiting invariant distribution, (12). But the expression for computing it is unboundedly long.

We need a truncation rule to tell us where to stop in the construction of the function.

A natural choice is given by (17): We use the approximation to the distribution that is accurate up to the point where the population would be expected to contain less than one living person from the original population.

A convenient assumption to complete the model is that

whenever you are the last member of your cohort, you die of loneliness. (But you do not anticipate that you will die of loneliness – your behavior continues to reflect the solution to the "perpetual youth" model).

Suppose that in the limiting distribution of π_∞ the measure of the unbounded portion of $\mathcal{S}$ lying above $\bar{A}$ is

$$q_\infty(n+1) = \int_{\mathcal{S}\backslash\Gamma} \pi_\infty(s) \qquad (28)$$

with conditional mean

$$s_{n+1} = \left(\frac{1}{\mathcal{L}(\mathcal{S}\backslash\Gamma)}\right) \int_{\mathcal{S}\backslash\Gamma} s\pi_\infty(s) \qquad (29)$$

while the conditional mean and measure of the points $s \in A_n$ are s_n and $q_\infty(n)$. The mean of the two sets together (that is, those included, and those excluded from the finite approximation to the transition matrix) is $q_\infty(n)s_n + q_\infty(n+1)s_{n+1}$.

4 Permanent Income Shocks

We now adapt the foregoing to the case of calculating an approximation to the steady state distribution of permanent income. We specialize the space $\mathcal{S}$ to contain points on the real number line so that for all $s \in \mathcal{S}$, $s \in \mathbb{R}$.

Suppose we have partitioned the space into a set of $(n+1)$ successive intervals indexed by i such that $i \in \mathbb{N}_{(n+1)} \doteq \{1, 2, \ldots, n, (n+1)\}$, whose boundaries are given by $\{\underline{A}_i, \bar{A}_i\}$ where $\underline{A}_1$ is the lower bound of the space A and $\bar{A}_n$ is the upper bound, and the upper bound of A_i is the lower bound of A_{i+1} (that is, $\bar{A}_i = \underline{A}_{i+1}$). The set A_i therefore contains all $s \in [\underline{A}_i, \underline{A}_{i+1})$, and $A_{(n+1)}$ is the mass of points (if any) excluded by the bounding process.

Leaves any point mass at $\bar{A}_i$ in $i+1$

We now have the notation to describe our discrete approx-

imation to the distribution of the level of permanent income $\wp$.

Per (26), we can calculate the transition probabilities from:

$$\hat{P}_{i,j} = \left(\underbrace{\frac{1}{\mathcal{L}(A_i)}}_{A_i \text{ Pop mass}} \right) \overbrace{\int_{A_i}}^{s\in[\underline{A}_i,\bar{A}_i)} \left(\underbrace{\overbrace{\int_{A_j}}^{\bar{s}\in[\underline{A}_j,\bar{A}_j)} \mathfrak{t}_s(\bar{s}|s)d\bar{s}}_{\equiv\bar{b}(j|s)} \right) ds \quad \forall i \in \mathbb{N}_n$$

(30)

where $\bar{b}(j|s)$ is the probability mass for being in interval j next period, given that you are at point s this period.

Defining $\mathcal{F}^\bullet$ as the cumulative distribution function of $\bullet$, now recall from (2) that for $s \in \Lambda_{\tau+1}$ (locations where the occupant survives),

$$\bar{s} = s + \psi \tag{31}$$

so

$$\mathfrak{t}_s(\bar{s}|s, \bar{s} \in \Lambda_{\tau+1}) = d\mathcal{F}^\psi(\bar{s} - s). \tag{32}$$

In words: To get from s to $\bar{s}$ requires a shock of size $\psi = \bar{s}-s$ whose probability density is measured by $d\mathcal{F}^\psi(\bar{s} - s)$. For a given s the probability mass for the set of draws of ψ that put you in A_j is therefore:

$$\bar{\sharp}(j|s, \bar{s} \in \Lambda_{\tau+1}) = \int_{\underline{A}_j - s}^{\bar{A}_j - s} d\mathcal{F}^\psi \tag{33}$$
$$= \mathcal{F}^\psi(\bar{A}_j - s) - \mathcal{F}^\psi(\underline{A}_j - s)$$

For locations whose occupant is replaced, the calculation is even simpler. At those locations, permanent income is reset to its starting value of 0 and then a shock is drawn, so

$$\bar{s} = \psi \tag{34}$$

so

$$\bar{\bar{\sharp}}(j|s, \bar{s} \notin \Lambda_{\tau+1}) = \int_{\underline{A}_j}^{\bar{A}_j} d\mathcal{F}^\psi \tag{35}$$

$$= \mathcal{F}^\psi(\bar{A}_j) - \mathcal{F}^\psi(\underline{A}_j)$$

so the unconditional transition process is a weighted mixture of the two distributions:

$$\bar{\sharp}(j|s) = \Lambda\bar{\sharp}(j|s, \bar{s} \in \Lambda_{\tau+1}) + (1 - \Lambda)\bar{\bar{\sharp}}(j|s, \bar{s} \notin \Lambda_{\tau+1}). \tag{36}$$

The overall transition probability structure is therefore

$$\hat{P}_{i,j} = \left(\frac{1}{\mathcal{L}(A_i)}\right) \int_{\underline{A}_i}^{\bar{A}_i} \bar{b}(j|s) d\mathcal{F}^{\pi_\tau}(s) \tag{37}$$

$$= \left(\frac{1}{\mathcal{L}(A_i)}\right) \int_{\underline{A}_i}^{\bar{A}_i} \bar{\bar{\sharp}}(j|s) d\mathcal{F}^{\hat{\pi}_\tau}(s). \tag{38}$$

The idea here is that we are replacing the respective set of shocks that would be sufficient, in the continuous state space, to move from each point in any partition i to some point in partition j (weighted by each shock's probability) by, in the discretized state space, a single shock with that same probability that moves the (now assumed to be identical) occupants of discrete bin i to discrete bin j.

4.1 Intervals

There are many possible ways of choosing the set of intervals A_i. A particularly transparent choice is the equiprobable case: The intervals are chosen so that in the steady state there should be an equal population mass in each interval. If we have the "true" CDF of the limiting distribution $\mathcal{F}^\mathcal{P}$ then we can define:

Define a set of points from $\sharp_0$ to $\sharp_n$ on the $[0, 1]$ interval as

the elements of the set $\sharp = \{0, 1/n, 2/n, \ldots, 1\}$.[8] Call the inverse of the $\mathcal{P}$ distribution $(\mathcal{F}^{\mathcal{P}})^{-1}$, and define the points $s_i = (\mathcal{F}^{\mathcal{P}})^{-1}(\sharp_i)$. Then the conditional mean of s in each of the intervals numbered 1 to n is:

$$\wp_i \equiv \mathbb{E}[\wp | \sharp_{i-1}^{-1} \leq \wp < \sharp_i^{-1}] = \int_{\underline{A}_i}^{\bar{A}_i} \wp \, d\mathcal{F}^{\wp}(\wp) \quad (39)$$

4.2 What Exactly Are We Approximating?

In the original, truly-continuous-state-space economy, the homotheticity of the problem allowed us to eliminate a state variable by dividing the problem by the level of permanent income. This required that the anticipated distribution of the shocks to (log) permanent income be identical no matter what the consumer's level of permanent income.

In our approximation, however, the size of the jump in s cannot be independent of the level of s. This is most obvious at the outermost bins of the distribution: It is not possible to jump to a level of permanent income that is lower than the lowest bin, or higher than the highest bin, while those two possibilities are equally likely for the point at the midpoint of the $\mathcal{F}$.

There are two possible interpretations of what we are doing:

- We are modeling a 'true' economy in which, in truth, the magnitude of the permanent shocks is independent of their level

- We are imagining that we have moved the population of the true economy to a new and different economy that follows the structure we have derived as an approximation to the true economy

---

[8]These points define intervals that constitute a partition of the domain of $\mathcal{F}^{\mathcal{P}}$.

In the former case, we think of our model consumers as actually living in the true economy. Their expectations, and the "true" evolution of the population, reflect the homotheticity/scaling assumptions we made earlier. In the latter case, consumers would need to keep track of their permanent income as a state variable, because the true characteristics of the shocks they would experience would depend on the level of their permanent income.

5 Markov Transition for a

Designating a next-period variable $\bullet$ by $\grave{\bullet}$, for a person who ends the current period with assets a, market resources at the beginning of the next period will be

$$\grave{m} = \grave{\mathcal{R}}a + \grave{\xi}. \tag{40}$$

We assume that we have found converged consumption and asset functions $\mathrm{a}(m) = m - \mathrm{c}(m)$.

Now we want to calculate the discrete Markov approximation to the true continuous evolution of the economy.

$$\grave{a} = \mathrm{a}(\grave{\mathcal{R}}a + \grave{\xi}) \tag{41}$$

The period-τ distribution of $\mathcal{M}_p$ is defined by a probability mass function (pmf) $\mathbf{b}_\tau = [\mathbf{b}_\tau(1), \mathbf{b}_\tau(2), \ldots, \mathbf{b}_\tau(n+1)]$.[9] The transition matrix is

$$\hat{P}_{i,j} = \left(\frac{1}{\mathcal{L}(\mu_i)}\right) \overbrace{\int_{\underline{\mu}_i}^{\bar{\mu}_i}}^{a\in[\underline{\mu}_i,\bar{\mu}_i)} \underbrace{\overbrace{\int_{\underline{\mu}_j}^{}}^{\grave{a}\in[\underline{\mu}_j,\bar{\mu}_j)} \mathbf{t}_a(\grave{a}|a)d\grave{a}}_{\equiv \grave{b}(j|a)}\, da \quad \forall i \in \mathbb{N}_n \tag{42}$$

where $\grave{b}(j|a)$ is the probability of being in interval j next period, given that you are at point a this period. There are two consequential structural differences with the case above: The transition function $\mathbf{t}_a$ now depends on two shocks, $\grave{\psi}$ and $\grave{\xi}$, instead of just one ($\grave{\psi}$); and there is no closed-form solution for the density function.[10]

We now want to inspect $\grave{b}(j|a)$ more closely:

$$\grave{\sharp}(j|a) = \int_{\underline{\mu}_j}^{\bar{\mu}_j} \mathbf{t}_a(\grave{a}|a)d\grave{a} \tag{43}$$

---

[9]If the consumption rule has a target $\breve{a}$, a natural starting point is to have all of the population mass at $\breve{a}$.

[10]There is no such solution for a sum of lognormals, much less for the sum of a lognormal and the inverse of a lognormal.

In words, we integrate over the combinations of realizations of $\grave{\psi}$ and $\grave{\xi}$ such that, if we start at a, $\grave{a} \in \mu_j$. For the moment, let's take $\grave{\mathcal{R}}$ as fixed and consider which values of ξ would lead to some $\grave{a} \in \mu_j$. Because a($\bullet$) is not an analytical function, for any $\grave{\mathcal{R}}(\psi)$ (equivalently, for any ψ) we must search numerically for the $\{\grave{\underline{\xi}}(a,\psi,j), \grave{\bar{\xi}}(a,\psi,j)\}$ such that $\grave{a}$ arrives respectively at $\{\underline{\mu}_j, \bar{\mu}_j\}$.[11]

Various characteristics of the a function (continuity; monotonicity; concavity; limits) mean that $\grave{a}$ is monotonically increasing in ξ, so the pmf $\grave{b}(j|a)$ for arriving in state j as a function of ψ is simply:

$$\grave{b}(j|a,\grave{\psi}) = \mathcal{F}^\xi(\bar{\xi}(j|a,\grave{\psi})) - \mathcal{F}^\xi(\underline{\xi}(j|a,\grave{\psi})). \qquad (44)$$

so the probability mass for arrival in j is

$$\grave{b}(j|a) = \int_{\underline{\psi}}^{\bar{\psi}} \grave{b}(j|a,\grave{\psi})d\mathcal{F}^{\grave{\psi}} \qquad (45)$$

We now have the pmf of μ_j for any fixed a,[12] so in order to calculate the probability of arriving in μ_j from any point in μ_i all we need to do is integrate the conditional probability over all possible values of $a \in \mu_i$:

$$\hat{P}_{i,j} = \left(\frac{1}{\mathcal{L}(\mu_i)}\right) \int_{\underline{\mu}_i}^{\bar{\mu}_i} \grave{b}(j|a)d\mathcal{F}^a. \qquad (46)$$

Our approximation scheme relies upon the substitution, for each $a \in \mu_i$ a point that is the weighted mean, $\hat{a}_i = \int_{\mu_i} a d\mathcal{F}^a$. But without having the CDF $\mathcal{F}$ of the continuous distribution, we have no way to compute these values.

Fortunately, Soudjani and Abate (2014) describe a recursive procedure that is guaranteed to result in bounded errors.

Theorem 1. The functions π_τ satisfy the bound

π is small for a in $\mathcal{S}$ that are excluded from Λ

---

[11] See the appendix for a sketch of a procedure for doing this.

[12] We are assuming that $\log \xi \sim \mathcal{N}(\mu, \sigma^2)$, so we can use the fact that: $\mathcal{F}^\xi(\xi) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{\xi - \mu}{\sigma\sqrt{2}}\right)\right)$

19

$$0 \le \pi_\tau(a) \le \varepsilon_\tau \quad \forall a \in \mathcal{S} \backslash \Lambda_\tau \qquad (47)$$

where the quantities $\{\varepsilon_\tau\}_{\tau=0}^N$ are defined recursively by

$$\varepsilon_{\tau+1} = \epsilon + \mathrm{M}_p \varepsilon_\tau, \qquad (48)$$

whereas the support sets $\{\Lambda\}_{\tau=0}^N$ are computed as

$$\Lambda_{\tau+1} = \Pi_{\grave{a}} \left(\Gamma \cap (\Lambda \times \mathcal{S}) \right) \qquad (49)$$

where $\Pi_{\grave{a}}$ denotes the projection map along the second set of coordinates. This means that we can start with an arbitrary distribution, and if we iterate long enough we can construct support sets that are large enough to guarantee the errors are bounded arbitrarily tightly.

Concretely, what we do is to begin with a support set Λ_0 defined by the bounds of our discretization of the shock process. Concretely, we construct Λ_0 as $a \in [\underline{\xi} - \mu_\xi, \bar{\xi} - \mu_\xi]$ with pmf identical to the pmf of the shock ξ. Then we proceed by continuing to expand. That is, each of the nodes in Λ_0 yields an array of other points as that location experiences the full distribution of outcomes from the discretized process. So, for instance, the bottommost point in Λ_1 will match the result of having twice drawn the worst possible realization of the shock. We iterate for enough periods that we judge to give a wide enough dispersion of points for the final approximation, then stop.

6 Backwards

We now consider an economy in which there is a last period T.

The main thing that makes this problem different from the former one is that the consumption function now depends on time, a_τ. In the last period $\mathsf{a}_\mathsf{T}(m) = 0$, and for prior

periods it will be obtained by solving the dynamic stochastic optimization problem.

Assuming that the distribution of the shocks does not change with the time period,

$$\hat{P}_{i,j}[\tau] = \int_{\underline{\mu}_i}^{\bar{\mu}_i} \grave{\flat}[\tau](j|a)d\mathcal{F}_\tau^a \qquad (50)$$

Rewriting (**??**)

$$a_{\tau+1} = a_{\tau+1}(\mathcal{R}_{\tau+1}a_\tau + \xi_{\tau+1}) \qquad (51)$$

we again must solve numerically for $\bar{\xi}[\tau](j|a_\tau, \psi_{\tau+1})$ and $\underline{\xi}[\tau](j|a_\tau, \psi_{\tau+1})$ functions, which will now yield a

$$\grave{\flat}[\tau](j|a_\tau, \psi_{\tau+1}) = \mathcal{F}^\xi(\bar{\xi}[\tau](j|a_\tau, \psi_{\tau+1})) - \mathcal{F}^\xi(\underline{\xi}[\tau](j|a_\tau, \psi_{\tau+1})), \qquad (52)$$

and the formula for $\grave{\flat}$ is again given by (**??**)

7 Intervals

As before, suppose we have a prespecified set of boundaries on the real number line: $\underline{\mu}_i, \bar{\mu}_i$ for $i \in \mathbb{N}_{n+1}$ points that define the boundaries of the intervals we want to use in constructing our Markov chain. Then the transition probabilities are again

$$a_i \equiv \int_{\underline{\mu}_i}^{\bar{\mu}_{i+1}} a d\mathcal{F}^a \qquad (53)$$

$$\grave{\flat}[\tau](j|a_\tau, \psi_{\tau+1}) = \qquad (54)$$

We need to figure out how to divide the continuous state space up into appropriately sized intervals for which each of the

8 Anticipation

Per the discussion above, our consumers do not believe themselves to be living in a discretization of their economy; they believe they are living in the original continuous economy.

For any period-τ ending value of a_τ, expected marginal utility will be:

$$\mathfrak{v}'_\tau(a_\tau) \equiv \mathbb{E}_\tau[\mathrm{u}'\left(\mathrm{c}(\mathcal{R}_{\tau+1}a_\tau + \xi_{\tau+1})\right)] \tag{55}$$

which as usual yields consumption and beginning-of-period market resources from

$$c_\tau = (\mathrm{u}')^{-1}(\mathfrak{v}'_\tau) \tag{56}$$

$$m_\tau = a_\tau + c_\tau \tag{57}$$

References

BLANCHARD, OLIVIER J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93(2), 223–247.

SOUDJANI, SADEGH ESMAEIL ZADEH, AND ALESSANDRO ABATE (2014): "Precise approximations of the probability distribution of a markov process in time: An application to probabilistic invariance," in *Tools and algorithms for the construction and analysis of systems*, ed. by Erika Abraham, and Klaus Havelund, pp. 547–561, Berlin, Heidelberg. Springer Berlin Heidelberg.

——— (2015): "Quantitative approximation of the probability distribution of a Markov process by formal abstractions," *Logical Methods in Computer Science*, 11, 1–29.

A Numerical Search for ξ

But we do have analytical lower and upper bounds for c, which in turn can be used to give us lower and upper bounds for the search region.

Since $\grave{a}$ is continuous and strictly increasing in the realization of ξ and $\mu_j \in [\underline{\mu}_j, \bar{\mu}_j]$, for any given $\grave{\mathcal{R}}$ we can find the lower bound such that $\grave{a} = \underline{\mu}_j$ and the upper bound such that $\grave{a} = \bar{\mu}_j$

$$\underline{\mu}_j(\grave{\mathcal{R}}) = \grave{\mathcal{R}}a + \underline{\grave{\xi}} - c(\grave{\mathcal{R}}a + \underline{\grave{\xi}}) \tag{58}$$

$$\bar{\mu}_j(\grave{\mathcal{R}}) = \grave{\mathcal{R}}a + \bar{\grave{\xi}} - c(\grave{\mathcal{R}}a + \bar{\grave{\xi}}) \tag{59}$$

Because c is not an analytical expression, this is an implicit function, but for 'small' ξ can be approximated by a first

order Taylor expansion:

$$\mathrm{c}(\mathcal{R}a + \xi) \approx \mathrm{c}(\mathcal{R}a + 1) + (\xi - 1)\mathrm{c}'(\mathcal{R}a + 1) \qquad (60)$$

and shortening $\mathrm{c}'(\mathcal{R}a + 1)$ to c',

$$\check{a} \approx \left(\grave{\mathcal{R}}a - \mathrm{c}(\grave{\mathcal{R}}a + 1)\right) + \grave{\xi} + (\grave{\xi} - 1)\mathrm{c}'(\mathcal{R}a + 1) \qquad (61)$$

$$\check{a} - (\mathrm{c}' + 1)\grave{\xi} \approx \left(\grave{\mathcal{R}}a - (\mathrm{c}(\grave{\mathcal{R}}a + 1))\right) - \mathrm{c}' \qquad (62)$$

$$(\mathrm{c}' + 1)\grave{\xi} \approx \check{a} - \left(\grave{\mathcal{R}}a - (\mathrm{c}(\grave{\mathcal{R}}a + 1))\right) + \mathrm{c}' \qquad (63)$$

$$\grave{\xi} \approx (1 + \mathrm{c}')^{-1} \left\{ \check{a} - \left(\grave{\mathcal{R}}a - (\mathrm{c}(\grave{\mathcal{R}}a + 1))\right) + \mathrm{c}' \right\} \qquad (64)$$

which can serve as the starting point for a numerical search for the 'true' value of $\grave{\xi}$ that satisfies (**??**).