ONLINE APPENDIX

"Risk, Uncertainty, and the Dynamics of Inequality"

by Kenneth Kasa and Xiaowen Lei

Appendix A. Proof of Proposition 3.1

After substituting the h policy function in eq. (3.6) into the HJB equation in eq. (3.5) we get

$$(\rho+\delta)V(w) = \max_{c,\alpha}\left\{\frac{c^{1-\gamma}}{1-\gamma} + [(r+\alpha(\mu-r))w - c]V'(w) + \frac{1}{2}\alpha^2\sigma^2 w^2 V''(w) - \frac{1}{2}\varepsilon\alpha^2\sigma^2 w^2 (V'(w))^2\right\} \quad (A.1)$$

Except for the last term multiplying ε, this is a standard Merton consumption/portfolio problem, with a well known solution. Hence, this suggests a perturbation approximation around ε. To obtain this, we posit

$$V(w) \approx V^0(w) + \varepsilon V^1(w)$$

Our goal is to solve for $V^0(w)$ and $V^1(w)$. With this approximation, a first-order approximation of the $c(w)$ policy function is

$$\begin{aligned} c(w) &= V_w^{-\frac{1}{\gamma}} \\ &\approx [V_w^0 + \varepsilon V_w^1]^{-\frac{1}{\gamma}} \\ &\approx \left(V_w^0\right)^{-\frac{1}{\gamma}} - \varepsilon\frac{1}{\gamma}\left(V_w^0\right)^{-\frac{1}{\gamma}-1} V_w^1 \\ &\equiv c_0 + \varepsilon c_1 \end{aligned} \quad (A.2)$$

and a first-order approximation of the $\alpha(w)$ policy is

$$\begin{aligned} \alpha(w) &= -\frac{(\mu-r)}{\sigma^2}\left[\frac{V_w}{w(V_{ww} - \varepsilon V_w^2)}\right] \\ &\approx -\frac{(\mu-r)}{\sigma^2}\left[\frac{wV_w^0}{w^2V_{ww}^0} + \varepsilon\frac{wV_w^1V_{ww}^0 - wV_w^0(V_{ww}^1 - (V_w^0)^2)}{w^2(V_{ww}^0)^2}\right] \\ &\equiv \alpha_0 + \varepsilon\alpha_1 \end{aligned} \quad (A.3)$$

Substituting these approximations into HJB equation in (A.1) gives us

$$\begin{aligned} (\rho+\delta)[V^0 + \varepsilon V^1] = {} & \frac{c_0^{1-\gamma}}{1-\gamma} + \varepsilon c_0^{-\gamma}c_1 + (V_w^0 + \varepsilon V_w^1)[(r + (\alpha_0 + \varepsilon\alpha_1)(\mu-r))w - (c_0 + \varepsilon c_1)] + \frac{1}{2}(\alpha_0 + \varepsilon\alpha_1)^2\sigma^2 w^2 [V_{ww}^0 + \varepsilon V_{ww}^1] \\ & - \frac{1}{2}\varepsilon(\alpha_0 + \varepsilon\alpha_1)^2\sigma^2 w^2 (V_w^0 + \varepsilon V_w^1)^2 \end{aligned}$$

Matching terms of equal order and dropping terms of order $O(\varepsilon^2)$ gives us

$$\begin{aligned} \varepsilon^0: \quad (\rho+\delta)V^0 &= \frac{c_0^{1-\gamma}}{1-\gamma} + V_w^0[(r + \alpha_0(\mu-r))w - c_0] + \frac{1}{2}\alpha_0^2\sigma^2w^2V_{ww}^0 \\ \varepsilon^1: \quad (\rho+\delta)V^1 &= c_0^{-\gamma}c_1 + V_w^1[(r+\alpha_0(\mu-r))w - c_0] + V_w^0[\alpha_1(\mu-r)w - c_1] + \sigma^2w^2[\alpha_0\alpha_1V_{ww}^0 + \alpha_0^2V_{ww}^1] - \frac{1}{2}\alpha_0^2\sigma^2w^2(V_w^0)^2 \end{aligned}$$

Note this system is recursive. We can first solve the ε^0 equation for V^0, and then substitute this into the ε^1 equation. Solving for V^0 just gives the Merton solution. In particular, we conjecture $V^0(w) = \frac{A}{1-\gamma}w^{1-\gamma}$. Note, this implies $c_0 = A^{-1/\gamma}w$ and $\alpha_0 = \frac{(\mu-r)}{\gamma\sigma^2}$. After canceling the common $w^{1-\gamma}$ term we can solve for A. This produces the expression for A_0 stated in Proposition 3.1.

To solve the ε^1 equation we conjecture $V^1(w) = \frac{B}{\epsilon}w^{\epsilon}$, and try to solve for B and ϵ. From (A.26) and (A.27), this guess implies $c_1 = -\gamma^{-1}A^{-1/\gamma-1}Bw^{2-\gamma}$ and $\alpha_1 = -\left(\frac{\mu-r}{\sigma^2}\right)\left(\frac{A^2+(\gamma-1)B}{A\gamma^2}\right)w^{1-\gamma}$. Note, these now depend on w. Substituting these into the ε^1 equation, we find that if $\epsilon = 2(1-\gamma)$, we can cancel out the terms in w and solve for B. Doing so produces the expression for A_1 stated in Proposition 3.1. □

Appendix B. Proof of Corollary 3.2

This follows immediately from the proof of Proposition 3.1. Here we fill in some of the details omitted in the above proof. Substituting the expressions for V^0 and V^1 into the expression for $\alpha(w)$ in eq. (A.3) gives

$$\begin{aligned}\alpha(w) &\approx \alpha_0 - \varepsilon\left(\frac{\mu - r}{\sigma^2}\right)\left[\frac{-\gamma A_0 A_1 w^{-3\gamma} - A_0 A_1(1-2\gamma)w^{-3\gamma} + A_0^3 w^{-3\gamma}}{\gamma^2 A_0^2 w^{-2\gamma-1}}\right] \\ &= \alpha_0 - \varepsilon\alpha_0\left[\frac{-\gamma A_1 - (1-2\gamma)A_1 + A_0^2}{\gamma A_0}\right] w^{1-\gamma}\end{aligned}$$

One can readily verify this is the same expression as stated in Corollary 3.2. Next, substituting the expressions for V^0 and V^1 into the expressions for $c(w)$ in eq. (A.2) gives

$$\frac{c(w)}{w} = A_0^{-1/\gamma} - \varepsilon\frac{1}{\gamma}\left(A_0^{-1/\gamma-1}A_1\right)w^{1-\gamma}$$

This then implies the expression for $s(w) = 1 - [c(w)/w]$ stated in Corollary 3.2. □

Appendix C. Recursive Preferences I

The main text considers a traditional robust control problem with observable states where an agent conditions on a given benchmark model, and then formulates policies that are robust to local unstructured perturbations around this model. As noted by Hansen, Sargent, Turmuhambetova, and Williams (2006), continuous-time versions of this problem are observationally equivalent to Duffie and Epstein's (1992) Stochastic Differential Utility (SDU) model of recursive preferences. Hence, risk aversion is not separately identified from ambiguity/uncertainty aversion.

This observational equivalence has sparked a more recent literature that attempts to distinguish risk aversion from both ambiguity aversion and intertemporal substitution. Hansen and Sargent (2008, chpts. 18 and 19) note that the key to separating risk aversion from ambiguity aversion is to introduce hidden state variables, which the agent attempts to learn about. Early robust control methods were criticized because they abstracted from learning. Ambiguity is then defined by distortions of the agent's estimates of the hidden states.[1]

Hidden states can be used to represent a wide range of unobservables. For example, time invariant hidden states can index alternative models. Here we assume the hidden state is an unobserved, potentially time-varying, mean investment return. In principle, we could allow the agent to be uncertain about both the dynamics conditional on a particular mean growth rate, as well as the mean itself. However, for our purposes it is sufficient to assume the agent is only uncertain about the mean.[2]

Distorted beliefs about the hidden state can be interpreted from the perspective of the Klibanoff, Marinacci, and Mukerji (2005) (KMM) model of smooth ambiguity aversion. In the KMM model an agent prefers act f to act g if and only if

$$\mathbb{E}_\mu\phi(\mathbb{E}_\pi u \circ f) \geq \mathbb{E}_\mu\phi(\mathbb{E}_\pi u \circ g)$$

where $\mathbb{E}$ is the expectation operator, π is a probability measure over outcomes conditional on a model, and μ is a probability measure over models. Ambiguity aversion is characterized by the properties of the ϕ function, while risk aversion is characterized by the properties of the u function. If ϕ is concave, the agent is ambiguity averse. KMM refer to $\mathbb{E}_\pi$ as 'first-order beliefs', while $\mathbb{E}_\mu$ is referred to as 'second-order beliefs'. Note that when ϕ is nonlinear, the implicit compound lottery defined by selecting a model with unknown parameters cannot be reduced to a single lottery over a 'hypermodel', as in Bayesian decision theory, so the distinction between models and parameters becomes important. Also note that from the perspective of smooth ambiguity aversion, evil agents and entropy penalized drift distortions are just a device used to produce a particular distortion in second-order beliefs about continuation values, i.e., where $\phi(V) \approx -\exp(-\varepsilon V)$.

The original KMM model was static. Klibanoff, Marinacci, and Mukerji (2009) extend it to a recursive, dynamic setting. However, their implicit aggregator is additive, so risk and intertemporal substitution cannot be distinguished. In response, Hayashi and Miao (2011) propose a model of generalized smooth ambiguity aversion by combining KMM with an Epstein-Zin aggregator. Unfortunately, as noted by Skiadas (2013), this model does not extend to continuous-time with Brownian information structures. Intuitively, first-order uncertainty (risk) is $O(dt)$, whereas second-order uncertainty (ambiguity) is $O(dt^2)$, and so it evaporates in the continuous-time limit.

---

[1] This definition of ambiguity aversion is based on the axiomatization of Ghirardato and Marinacci (2002), which defines ambiguity aversion as deviations from subjective expected utility. Epstein (1999) proposes an alternative definition based on deviations from probabilistic sophistication.

[2] In the language of Hansen and Sargent (2008, chpt. 18), we activate the T^2-operator by setting $\theta_2 < \infty$, while deactivating the T^1-operator by setting $\theta_1 = \infty$. This is a subtle distinction, since at the end-of-the-day they both produce drift distortions. However, they do this in different ways.

In response, Hansen and Sargent (2011) propose a trick to retain ambiguity aversion, even as the sampling interval shrinks to zero. In particular, they show that if the robustness/ambiguity-aversion parameter is scaled by the sampling interval, then ambiguity aversion will persist in the limit. Intuitively, even though second-order uncertainty becomes smaller and smaller as the sampling interval shrinks, because the agent effectively becomes more ambiguity averse at the same time, ambiguity continues to matter.

In what follows we outline a heuristic combination of the discrete-time generalized KMM preferences of Hayashi and Miao (2011) and the continuous-time scaling trick of Hansen and Sargent (2011). As far as we know, there are no formal decision-theoretic foundations for such a combination, at least not yet.

With recusive preferences, the agent's problem becomes

$$V_t = \max_{c,\alpha} \min_{h} E_t \int_t^\infty f(c_s, V_s) ds$$

where $f(c_s, V_s)$ is the (normalized) Duffie-Epstein aggregator,

$$f(c, V) = \varphi(1-\gamma)V \left[\log(c) - \frac{1}{1-\gamma}\log((1-\gamma)V)\right]$$

and where for simplicity we've assumed the the elasticity of intertemporal substitution is unity. The effective rate of time preference is $\varphi = \rho + \delta$, and the coefficient of relative risk aversion is $\gamma \neq 1$.[3] The budget constraint is the same as before

$$dw = \{[r + \alpha(\mu - r)]w - c + \alpha\sigma w h\}\, dt + \alpha\sigma w dB$$

The HJB equation is

$$0 = \max_{c,\alpha} \min_{h} \left\{ f(c, V) + \frac{1}{2\varepsilon}h^2 + ([r + \alpha(\mu - r)]w - c + \alpha\sigma w h)\, V'(w) + + \frac{1}{2}\alpha^2\sigma^2 w^2 V''(w) \right\}$$

Note that discounting is embodied in the properties of the aggregator. Also note that in contrast to Bayesian learning models, where the drift is regarded as an unknown parameter and its current estimate becomes a hedgeable state variable, here the drift is viewed as a *control* variable, which is selected by the agent to produce a robust portfolio.

The first-order conditions for (α, h) are the same as before, while the first-order condition for c becomes:

$$c = \frac{\varphi(1-\gamma)V}{V'(w)}$$

If these are substituted into the HJB equation we get:

$$0 = f[c(V, V'), V] + (rw - c)V' - \frac{1}{2}\frac{(\mu - r)^2 (V')^2}{[V'' - \varepsilon(V')^2]\sigma^2}$$

Our goal is to compute the following first-order approximation

$$V(w) \approx V^0(w) + \varepsilon V^1(w)$$

By inspection, it is clear that when $\varepsilon = 0$ it is natural to guess

$$V^0(w) = \frac{\hat{A}_0}{1-\gamma} w^{1-\gamma}$$

Note that this implies $c_0 = \varphi w$ and $\alpha_0 = (\mu - r)/\gamma\sigma^2$. Substituting into the HJB equation and cancelling out the common $w^{1-\gamma}$ term gives the following equation for $\hat{A}_0$

$$0 = \varphi \left[\log(\varphi) - \frac{1}{1-\gamma}\log(\hat{A}_0)\right] + (r - \varphi) - \frac{1}{2}\frac{(\mu - r)^2}{\gamma\sigma^2} \qquad \Rightarrow \qquad \hat{A}_0 = \exp\left\{(1-\gamma)\left[\log(\varphi) + \frac{r - \varphi}{\varphi} - \frac{1}{2}\frac{(\mu - r)^2}{\varphi\gamma\sigma^2}\right]\right\}$$

Next, matching the $O(\varepsilon)$ terms in the HJB equation yields the following ODE for $V^1(w)$

$$0 = \varphi(1-\gamma)\left\{V^1\left[\log(c_0) - \frac{1}{1-\gamma}\log((1-\gamma)V^0)\right] + V^0\left(\frac{c_1}{c_0} - \frac{1}{1-\gamma}\frac{V^1}{V^0}\right)\right\} + (r - \varphi)wV^1_w - \frac{1}{2}\alpha_0^2\sigma^2(wV^0_w)^2 + \frac{1}{2}\sigma^2 w^2[2\alpha_0\alpha_1 V^0_{ww} + \alpha_0^2 V^1_{ww}]$$

where

[3] We could easily allow $\gamma = 1$ by slightly modifying the aggregator. However, this would be uninteresting, since preferences would then collapse to additive form given that we've already assumed the elasticity of intertemporal substitution is unity.

$$
\begin{aligned}
c_1 &= \varphi(1-\gamma)\frac{V^1 V_w^0 - V^0 V_w^1}{(V_w^0)^2} \\
\alpha_1 &= -\gamma\alpha_0\frac{wV_w^1 V_{ww}^0 - wV_w^0(V_{ww}^1 - (V_w^0)^2)}{w^2(V_{ww}^0)^2}
\end{aligned}
$$

Note that the expression for α_1 is the same as before. Also as before, note that the system is recursive, with the above solutions for (V^0, c_0, α_0) becoming inputs into the V^1 ODE. If you stare at this ODE long enough, you will see that a function of the following form will solve this equation

$$V^1(w) = \frac{\hat{A}_1}{\epsilon} w^\epsilon$$

where as before $\epsilon = 2(1-\gamma)$. Substituting in this guess, cancelling the common w^ϵ terms, and then solving for $\hat{A}_1$ gives

$$\hat{A}_1 = \frac{-\frac{1}{2}\alpha_0^2\sigma^2\hat{A}_0}{-\gamma\varphi\log(\varphi) + (1-\gamma)(r-\varphi) + \frac{\varphi\gamma}{2(\gamma-1)} - \varphi + \alpha_0^2\sigma^2[\frac{1}{2}(\gamma^2-1) + (\gamma-1)^2]}$$

From here, the analysis proceeds exactly as in the main text. We just need to replace (A_0, A_1) with $(\hat{A}_0, \hat{A}_1)$.The approximate saving rate now becomes

$$s(w) = 1 - \varphi + \varepsilon\frac{\varphi\hat{A}_1}{2\hat{A}_0}w^{1-\gamma}$$

If $\gamma > 1$, then this is increasing in w as long as $\hat{A}_1 < 0$, since $\hat{A}_0 > 0$.

To examine the quantitative properties of the model with recursive preferences, we use the same parameter values as those in Table 1, with three minor exceptions. First, since the model with recursive preferences seems to be somewhat less sensitive to the robustness parameter, we increased ε from 0.045 to 0.45. Second, we increased γ slightly from 1.31 to 1.5. Finally, we increased μ slightly, from 5.86% to 5.95%. These parameter values remain consistent with available empirical estimates. Figure 1 displays the resulting portfolio shares and savings rates.

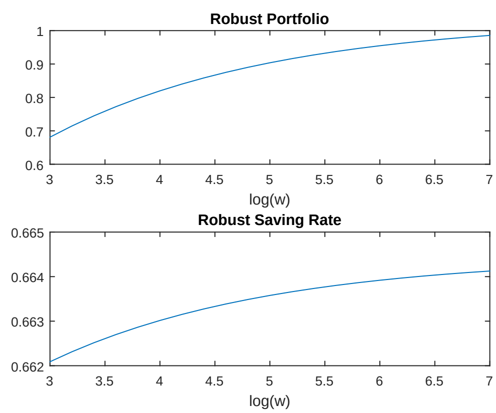

Figure 1. Recursive Preferences

The portfolio shares are similar to those in Figure 2, though wealth dependence is somewhat weaker. The key difference here is the saving rate. Now it is increasing in wealth. However, as in Figure 3, wealth dependence is very weak.

Appendix D. Proof of Proposition 4.1

Substituting the policy functions into the budget constraint gives

$$dw = [(r + \alpha(w)(\mu - r))w - c(w)]dt + \alpha(w)\sigma w dB$$

where $\alpha(w)$ and $c(w)$ are the approximate policy functions given in Corollary 3.2. Collecting terms, individual wealth dynamics can be described

$$\frac{dw}{w} = a(w;\varepsilon)dt + b(w;\varepsilon)dB$$

where the drift coefficient is given by

$$\begin{aligned} a(w;\varepsilon) &= r + \frac{(\mu-r)^2}{\gamma\sigma^2} - A_0^{-1/\gamma} + \varepsilon\left\{\frac{-\gamma^2\sigma^2\alpha_0^2[A_0^2+(\gamma-1)A_1)] + \gamma A_1 A_0^{-1/\gamma}}{\gamma^2 A_0}\right\} w^{1-\gamma} \qquad \text{(D.4)} \\ &\equiv \bar{a}_0 + \varepsilon\bar{a}_1 w^{1-\gamma} \end{aligned}$$

and the diffusion coefficient is given by

$$\begin{aligned} b(w;\varepsilon) &= \sigma\alpha_0 - \varepsilon\left\{\alpha_0\sigma\gamma\frac{A_0^2+(\gamma-1)A_1}{\gamma^2 A_0}\right\} w^{1-\gamma} \qquad \text{(D.5)} \\ &\equiv b_0 + \varepsilon b_1 w^{1-\gamma} \end{aligned}$$

In a standard random growth model, these coefficients would be constant. Hence, the wealth dependence here reflects the 'scale dependence' of our model. Next, let $x = \log(w) \equiv g(w)$ denote log wealth. Ito's lemma implies

$$\begin{aligned} dx &= g'(w)dw + \frac{1}{2}w^2 b(w)^2 g''(w)dt \\ &= \frac{dw}{w} - \frac{1}{2}\left(b_0 + \varepsilon b_1 w^{1-\gamma}\right)^2 dt \end{aligned}$$

Substituting $e^{(1-\gamma)x} = w^{1-\gamma}$ into the right-hand side and dropping $O(\varepsilon^2)$ terms gives

$$dx = [\bar{a}_0 - \frac{1}{2}b_0^2 + \varepsilon(\bar{a}_1 - b_0 b_1)e^{(1-\gamma)x}]dt + (b_0 + \varepsilon b_1 e^{(1-\gamma)x})dB$$

Finally, defining $a_0 \equiv \bar{a}_0 - \frac{1}{2}b_0^2$ and $a_1 = \bar{a}_1 - b_0 b_1$ gives the result stated in Proposition 4.1. □

Appendix E. Proof of Proposition 4.4

For convenience, we start by reproducing the KFP equation in (4.14)

$$\frac{\partial f}{\partial t} = -\frac{\partial[(a_0 + \varepsilon a_1 e^{(1-\gamma)x})f]}{\partial x} + \frac{1}{2}\frac{\partial^2[(b_0 + \varepsilon b_1 e^{(1-\gamma)x})^2 f]}{\partial x^2} - \delta f + \delta\zeta_0$$

Evaluating the derivatives gives

$$\frac{\partial f}{\partial t} = (-a_x + b_x^2 + bb_{xx})f + (2bb_x - a)\frac{\partial f}{\partial x} + \frac{1}{2}b^2\frac{\partial^2 f}{\partial x^2} - \delta f + \delta\zeta_0$$

where $a(x)$ and $b(x)$ are the drift and diffusion coefficients defined in equations (D.4) and (D.5) after the change of variables $w = e^x$. In general, these sorts of partial differential equations are not fun to solve. However, this PDE is *linear*, which opens the door to transform methods. The first step is to evaluate the derivatives of the $a(x)$ and $b(x)$ functions, and then drop the $O(\varepsilon^2)$ terms. Then we take the Laplace transform of both sides, with x as the transform variable. When doing this we use the following facts

$$\mathcal{L}\{\frac{\partial f}{\partial t}\} = \frac{\partial F(s)}{\partial t} \qquad \mathcal{L}\{\frac{\partial f}{\partial x}\} = sF(s) \qquad \mathcal{L}\{\frac{\partial^2 f}{\partial x^2}\} = s^2 F(s) \qquad \mathcal{L}\{e^{\beta x}f\} = F(s-\beta) \qquad \mathcal{L}\{\zeta_0\} = 1$$

where $\mathcal{L}\{f(x)\} \equiv F(t,s) \equiv \int_{-\infty}^{\infty} f(t,x)e^{-sx}dx$ defines the (two-sided) Laplace transform. The first result follows from interchanging differentiation and integration, while the second and third results follow from integration by parts (using the boundary conditions $f(-\infty) = f(\infty) = 0$). The fourth result is called the 'shift theorem', and follows from the change of variable $s \to s - \beta$. The last result is more subtle. The fact that the Laplace transform of a delta function is just equal to 1 uses results from the theory of generalized functions.[4]

Following these steps produces equation (4.15) in the text, which we repeat here for convenience

$$\frac{\partial F}{\partial t} = \Lambda(s)F(t,s) + \varepsilon\Phi(s-\beta)F(t,s-\beta) + \delta \qquad \text{(E.6)}$$

where $\beta \equiv 1 - \gamma$ and

$$\begin{aligned} \Lambda(s) &= \frac{1}{2}b_0^2 s^2 - a_0 s - \delta \\ \Phi(s) &= b_0 b_1 s^2 + (2b_0 b_1\beta - a_1)s + \beta(b_0 b_1\beta - a_1) \end{aligned}$$

[4]Kaplan (1962) provides a good discussion of Laplace transform methods.

To solve equation (E.6) we use the approximation $F(t, s-\beta) \approx F(t,s) - \beta \frac{\partial F}{\partial s}$. This gives us

$$F_t = \Lambda(s)F(t,s) + \varepsilon\Phi(s-\beta)\left[F(t,s) - \beta F_s\right] + \delta \tag{E.7}$$

where subscripts denote partial derivatives. Note that we are now back to solving a PDE. We solve this using a standard guess-and-verify/separation-of-variables strategy. In particular, we posit a solution of the following form

$$F(t,s) = H(t)G(s) + F_\infty(s)$$

Loosely speaking, $H(t)$ captures transition dynamics, $G(s)$ captures initial and boundary conditions, and $F_\infty(s)$ captures the new stationary distribution. Here we focus on this last component. Plugging this guess into (E.7) gives

$$H'G = \Lambda[HG] + \varepsilon\Phi[HG - \beta HG'] + \left\{[\Lambda + \varepsilon\Phi]F_\infty - \varepsilon\Phi\beta F'_\infty + \delta\right\} \tag{E.8}$$

where for convenience we have suppressed function arguments. The key point to notice here is that the last term in parentheses is independent of time, so we can solve it separately. Doing so gives us the robust stationary distribution. Another important observation here is that ε multiplies the derivative F'_∞. This 'singular perturbation' term makes conventional first-order perturbation approximation unreliable. To deal with this term we employ the change of variables $\hat{s} = s/\varepsilon$. With this change of variable we can write the ODE in parentheses as follows

$$F'_\infty = \frac{1}{\beta}\left(\varepsilon + \frac{\Lambda}{\Phi}\right)F_\infty + \frac{\delta}{\beta\Phi}$$

We can eliminate the nonhomogeneous term's dependence on $\hat{s}$ by defining $Q(\hat{s}) = \Phi(\hat{s}-\beta)F_\infty(\hat{s})$, which implies $F'_\infty = Q' - (\Phi'/\Phi)Q$. This delivers the following ODE

$$Q' = \left(\frac{\varepsilon + \Lambda + \beta\Phi'}{\beta\Phi}\right)Q + \frac{\delta}{\beta}$$

The general solution of this linear ODE is the sum of a particular solution to the nonhomogeneous equation and the solution of the homogeneous equation. However, we can ignore the homogeneous solution, since we know that a stationary distribution does not exist when $\delta = 0$.[5] Stated in terms of F_∞, the particular solution is

$$F_\infty(\hat{s}) = \frac{-\delta}{\varepsilon + \Lambda(\hat{s}) + \beta\Phi'(\hat{s}-\beta)}$$

After expanding the denominator polynomial into partial fractions we obtain the result stated in Proposition 4.4. To prove the correspondence principle stated in Corollary **??** we can just reverse the change of variables back to s. Let (R_1, R_2) denote the two roots of $F(\hat{s})$, (ϕ_1, ϕ_2) denote the roots of $\Lambda(s)$, and (r_1, r_2) denote the roots of $\Phi(s)$. After substituting s/ε for $\hat{s}$ and then multiplying numerator and denominator by ε^2 we get

$$\frac{\varepsilon^2}{(s-\varepsilon R_1)(s-\varepsilon R_2)} = \frac{-\varepsilon^2\delta}{\varepsilon^3 + (s-\varepsilon\phi_1)(s-\varepsilon\phi_2) + \varepsilon\beta[s - \varepsilon r_1 + s - \varepsilon r_2]}$$

Taking limits we obtain $\lim_{\varepsilon\to 0}(R_1, R_2) = (\phi_1, \phi_2)$, which is the stated correspondence principle. □

Appendix F. Proof of Proposition 5.1

This follows directly from the proof of Proposition **??**. Having solved for the stationary distribution, F_∞, the PDE in eq. (E.8) becomes

$$\tilde{F}_t = [\Lambda(s) + \varepsilon\Phi(s-\beta)]\tilde{F} - \varepsilon\beta\Phi(s-\beta)(G'/G)\tilde{F}$$

where $\tilde{F} \equiv HG$. Assuming $O(\beta) = O(\varepsilon)$, the last term can be dropped since it is second-order. □

Appendix G. Recursive Preferences II

To examine inequality dynamics with recursive preferences we just need to replace the expressions for (A_0, A_1) in the main text with the expressions for $(\hat{A}_0, \hat{A}_1)$ derived in Appendix C, and then replace the expressions for (a_0, a_1) stated in Proposition 4.1 with the following expressions for $(\hat{a}_0, \hat{a}_1)$:

$$\begin{aligned}\hat{a}_0 &= r - \varphi + \gamma\sigma^2\alpha_0^2 - \frac{1}{2}b_0^2 \\ \hat{a}_1 &= \frac{\frac{1}{2}\gamma^2\hat{A}_1 - (\sigma\gamma\alpha_0)^2(\hat{A}_0^2 + (\gamma-1)\hat{A}_1)}{\hat{A}_0\gamma^2} - b_0 b_1\end{aligned}$$

We can then follow the exact same approximation strategy as outlined in the main text. Figure 2 displays the resulting stationary distributions and convergence rates. We use the same parameter values as those in Table 1,

[5] Also note that since $\hat{s} = s/\varepsilon$, as $\varepsilon \to 0$ we know $\hat{s} \to \infty$. From the 'initial value theorem' we know $\lim_{s\to\infty} F(s) = 0$ since $f(x)$ is bounded at the switch point.

with the exceptions noted in Appendix C. In particular, we increase ε from 0.045 to 0.45, we increase γ slightly from 1.31 to 1.5, and increase μ slightly, from 5.86% to 5.95%.

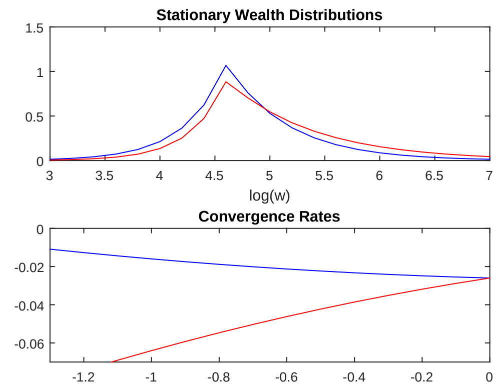

Figure 2. Recursive Preferences

Given the higher value of ε and the reinforcing effect of the saving rate, it is perhaps not too surprising that we now see a greater increase in inequality. The top 1% wealth share increases from 13.1% to 39.1%. Although 39.1% is very close to current estimates, 13.1% is somewhat smaller than its 1980 value. Finally, and most importantly, we find that robust convergence rates are even higher than those reported in the main text. At the mean level of wealth, the nonrobust convergence rate is 1.6%, close to its original value of 1.14%. However, now the robust convergence rate becomes 6.4%, more than 200 basis points higher than before, and four times greater than its nonrobust value.

One concern with using a higher value of ε is that it induces an overly pessimistic drift distortion and implausibly small detection error probability. However, we find that $\varepsilon = 0.45$ still produces maximal drift distortions around 1%, and detection error probabilities above 40%.

References

Epstein, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608.

Ghirardato, P., and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–89.

Hansen, L. P., and T. J. Sargent (2011): "Robustness and Ambiguity in Continuous Time," *Journal of Economic Theory*, 146, 1195–223.

Hansen, L. P., T. J. Sargent, G. A. Turmuhambetova, and N. Williams (2006): "Robust Control and Model Misspecification," *Journal of Economic Theory*, 128, 45–90.

Hayashi, T., and J. Miao (2011): "Intertemporal Substitution and Recursive Smooth Ambiguity Preferences," *Theoretical Economics*, 6, 423–72.

Kaplan, W. (1962): *Operational Methods for Linear Systems*. Addison-Wesley.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making Under Ambiguity," *Econometrica*, 73, 1849–92.

——— (2009): "Recursive Smooth Ambiguity Preferences," *Journal of Economic Theory*, 144, 930–76.

Skiadas, C. (2013): "Smooth Ambiguity Aversion toward Small Risks and Continuous-Time Recursive Utility," *Journal of Political Economy*, 121, 775–92.