

Robust Permanent Income and Pricing

LARS PETER HANSEN

University of Chicago, NBER and NORC

THOMAS J. SARGENT

Hoover Institution and Stanford University

and

THOMAS D. TALLARINI, JR.

Carnegie Mellon University

First version received July 1993; final version accepted January 1999 (Eds.)

“... I suppose there exists an extremely powerful, and, if I may so speak, malignant being, whose whole endeavours are directed toward deceiving me.” René Descartes, *Meditations*, II.¹

1. INTRODUCTION

This paper studies consumption and savings profiles and security market prices in a permanent income model when consumers are robust decision makers. Robust decision makers and expected utility maximizers share a common probabilistic specification of the income shocks. But robust decision makers suspect specification errors and want decisions to be insensitive to them. We show how a preference for robustness lies concealed within the quantity implications of the permanent income model and how it can be revealed by market-based measures of “risk-aversion”. We aim to show that large market-based measures of risk aversion can emerge from concern about small specification errors.

We reinterpret the decision rules for saving and consumption from a rational expectations version of Hall’s (1978) permanent income model with habit persistence. We show how a robust decision maker with a lower discount factor would use those same decision rules for saving and consumption.² Increasing the preference for robustness stimulates a precautionary motive for savings,³ an effect that an appropriate decrease of the discount factor cancels.⁴

1. Descartes (1901, p. 227).

2. Our setting relates to the max-min utility theory of Gilboa and Schmeidler (1989) and Epstein and Wang (1994). A robust decision maker uses rules that work well for a specific stochastic environment, but that are also insensitive to small perturbations of the probabilistic specification (see Zames (1981), Francis (1987), and Zhou, Glover and Doyle (1996)). Similarly, by ascribing a family of possible probability laws to a decision maker, the literature draws a sharp distinction between Knightian uncertainty and risk. Knightian uncertainty corresponds to the perturbations in the probabilistic specification envisioned by the robust control theorists.

3. Under a rational expectations interpretation, Hall’s model excludes precautionary savings, as emphasized by Zeldes (1989).

4. In effect, we are solving a particular “robust control” version of an “inverse optimal decision” problem. Versions of such problems have played an important role in the development of rational expectations theory. See Muth (1960). See Hansen and Sargent (1983) and Christiano (1987) for developments building on Muth’s work.

Our empirical strategy comes from the preceding observational equivalence result. To determine all but two parameters of the model, we estimate the rational expectations version of a habit-persistent version of Hall's model from aggregate U.S. time series on consumption and investment. By construction, our model with a preference for robustness must fit these quantity data as well as Hall's. But it has different implications about prices of risky assets. In particular, at the consumption/savings plan associated with Hall's model, the shadow prices of a robust decision maker put the market price of risk much closer to empirical estimates. After estimating Hall's model from the quantity data, we use some asset prices to calibrate the discount factor and a robustness parameter, while preserving the implications for saving and consumption.

In contrast to models in the spirit of Bewley (1977), market incompleteness plays no role in our decentralization of the permanent income model. Instead, following Hansen (1987), we interpret the permanent income decision rule in terms of a planning problem whose consumption and investment processes are equilibrium allocations for a competitive equilibrium. We then deduce asset prices as did Lucas (1978) and Epstein (1988) by finding shadow prices that clear security markets. These asset prices encode information about the slopes of intertemporal indifference curves passing through the equilibrium consumption process, and therefore measure the risk aversion of the consumer. To accommodate robustness, our decentralization copies Epstein and Wang's (1994).⁵

To model robust decision making requires formulating a class of misspecifications that worry the decision maker. We obtain a workable class of misspecifications by using the literature on risk-sensitive control started by Jacobson (1973, 1977) and extended by Whittle (1982, 1983, 1989, 1990) and ourselves (1995). Originally this literature did not seek to model robustness but rather sought to magnify responses to risk under rational expectations. The idea was to induce bigger effects of risk on decision rules (*i.e.* greater departures from certainty equivalence) by altering a single risk-sensitivity parameter that influences the intertemporal objective function. But risk-sensitive preferences can be reinterpreted as embedding a wish for robustness against a class of perturbations of the transition dynamics. For undiscounted linear-quadratic control problems, Glover and Doyle (1988) showed how a particular kind of concern for robustness connects to the risk-sensitive formulation of preferences. They showed how the risk sensitivity parameter measures the size of the class of misspecifications against which robustness is sought. We use a discounted version of James' (1995) notion of robustness. In this paper, we prefer to interpret our results in terms of a decision maker's concern for robustness. However, because we use a formulation of robust decision theory induced by the risk-sensitivity parameterization, an interpretation in terms of risk-sensitive preferences is also available.⁶

The remainder of this paper is organized as follows. Section 2 summarizes the necessary decision theory. We link risk-sensitive and robust decision theories by displaying two closely connected value functions associated with superficially different problems. The problems lead to identical decision rules. The second problem embodies a preference for robustness, provides links to Gilboa–Schmeidler's version of Knightian uncertainty, and explains the quote from Descartes. In Sections 3 and 4, we describe and estimate our

5. See Melino and Epstein (1995) for an alternative attack on this same question. They use a recursive formulation of an ε -contamination specification adapted from the theory of robust statistics.

6. To avail ourselves of this interpretation requires that we model risk sensitivity with discounting in a recursive manner, as in Epstein (1988), Weil (1989), Epstein and Zin (1989) and Hansen and Sargent (1995). Epstein and Zin (1989) developed a version of recursive utility theory that raises the market price of risk without altering the intertemporal substitution elasticity. Van Der Ploeg (1993) introduced risk sensitivity into a permanent income model, but not in a recursive manner.

permanent income model. The observational equivalence proposition of Section 4 motivates a two part strategy for using the quantity and asset price data. Section 5 exploits the links between robustness and risk-sensitivity in developing asset pricing formulas in terms of probability measures induced by “pessimistic” views of laws of motion that emerge as by-products of robust decision making. These formulas prepare the way for our interpretations of the market price of risk in terms of robustness. Section 6 quantifies the amount of preference for robustness required to push up the market price of risk. Section 7 measures intertemporal mean-risk trade-offs associated with different amounts of concern with robustness. Section 8 concludes.

2. RECURSIVE RISK SENSITIVE CONTROL

The theory rests on two closely related recursive linear quadratic optimization problems. We describe a distortion of beliefs away from rational expectations that induces the same behaviour as a particular modification of preferences toward risk. The equivalence of these two problems lets us interpret a “risk sensitivity” parameter as measuring a preference for robustness.

The recursive risk sensitive control problem

The state transition equation is

$$x_{t+1} = Ax_t + Bi_t + Cw_{t+1}, \quad (1)$$

where i_t is a control vector, x_t is the state vector, and w_{t+1} is an i.i.d. Gaussian random vector with $Ew_{t+1} = 0$, and $Ew_{t+1}w'_{t+1} = I$. Let J_t be the sigma algebra induced by $\{x_0, w_s, 0 \leq s \leq t\}$. The one-period return function is

$$u(i, x) = -i'Qi - x'Rx,$$

where Q is positive definite and R is positive semidefinite. Following Epstein and Zin (1989), Weil (1993), and Hansen and Sargent (1995), we use the following recursion to induce intertemporal preferences

$$U_t = u(i_t, x_t) + \beta \mathcal{R}_t(U_{t+1}), \quad (2)$$

where

$$\mathcal{R}_t(U_{t+1}) \equiv \frac{2}{\sigma} \log E \left[\exp \left(\frac{\sigma U_{t+1}}{2} \right) \middle| J_t \right]. \quad (3)$$

When $\sigma = 0$ we take $\mathcal{R}_t \equiv E(U_{t+1} | J_t)$, and we have the usual von Neumann–Morgenstern form of state additivity. When $\sigma \neq 0$, the operator \mathcal{R}_t makes an additional risk adjustment over and above that induced by the shape of $u(\cdot, \cdot)$. Values of σ less than zero correspond to more aversion to risk *vis a vis* the von Neumann–Morgenstern specification.⁷ As emphasized by Hansen and Sargent (1995), the (log, exp) specification links the general recursive utility specification of Epstein and Zin (1989) to risk-sensitive control theory. Weil’s (1993) permanent income model used the same (log, exp) specification but did not exploit connections to the risk-sensitive control literature.

7. As in Kreps and Porteus (1978), this recursive utility formulation overturns the indifference to the timing of the resolution of uncertainty inherent in state-separable preferences. The additional risk adjustment for $\sigma < 0$ implies a preference for early resolution of uncertainty.

The *risk sensitive control problem* is to maximize the time zero utility index U_0 by choosing a control process i_t adapted to J_t . Let $W(x)$ denote the optimum value function for this problem, so that $U_0^e = W(x_0)$ where the e superscript is used to distinguish the efficient or optimal utility index. Hansen and Sargent (1995) extended the Jacobson–Whittle risk-sensitive control theory to provide formulas for Ω and ρ in the following representation of the value function

$$U_t^e = W(x_t) = x_t' \Omega x_t + \rho. \quad (4)$$

Let $i = -Fx$ denote the optimal decision rule. Let $A^* = A - BF$ be the closed loop transition matrix (*i.e.* with $i_t = -Fx_t$ substituted into the original transition law). We display explicit formulas for the distorted expectation operator below.

We shall have cause to evaluate $\mathcal{R}_t(U_{t+1})$ for the quadratic value function (4) where Ω is a negative semidefinite matrix of real numbers and ρ is a nonpositive real number. It follows from Jacobson (1973) that

$$\mathcal{R}_t(U_{t+1}^e) = x_t' \hat{\Omega} x_t + \hat{\rho}, \quad (5)$$

where

$$\hat{\Omega} = A^{*'} [\Omega + \sigma \Omega C (I - \sigma C' \Omega C)^{-1} C' \Omega] A^*, \quad (6a)$$

and

$$\hat{\rho} = \rho - (1/\sigma) \log [\det (I - \sigma C' \Omega C)], \quad (6b)$$

so long as the matrix $(I - \sigma C' \Omega C)$ is positive definite, which we assume.

Robustness reinterpretation

We can reinterpret risk-sensitive preferences in terms of a decision maker with ordinary preferences who fears specification errors. The robustness interpretation is based on a recursive formulation of a zero-sum two-player Lagrange multiplier game whose value function $\tilde{W}(x)$ relates to $W(x)$. Parameterizing the game in terms of a fixed Lagrange multiplier makes a sequential version of the game, under the Markov perfect equilibrium concept, have the same outcome as a version where players can precommit at time zero.⁸

In this game, one player chooses decision rules for the control vector $\{i_t\}$, with two differences *vis a vis* the single agent risk-sensitive control problem. First, a maximizing player makes no risk adjustment in the utility function. Second, another minimizing player injects a distortion each time period into the conditional mean of the shock process. Thus, the first player maximizes a utility index $\tilde{U}_0 = E_0 \sum_{t=0}^{\infty} \beta^t u(i_t, x_t)$ by choice of state-feedback rules for $\{i_t\}$ and subject to the distorted law of motion

$$x_{t+1} = Ax_t + Bi_t + C(w_{t+1} + v_t), \quad (7)$$

8. Anderson, Hansen and Sargent (AHS) (1998) describe a different class of specification errors that leads to the same risk adjustments (3). AHS permit specification errors in the form of perturbations to a controlled Markov process. AHS use a constraint on the size relative entropy to parameterize the admissible class of misspecifications. Their formulation applies to nonquadratic objective functions and nonlinear laws of motion. They also formulate the connection between risk-sensitivity and robustness in continuous time.

where v_t distorts the mean of the innovation. The second player chooses a feedback rule for v_t to minimize \tilde{U}_0 subject to

$$\hat{E}_t \sum_{j=0}^{\infty} \beta^j v_{t+j} \cdot v_{t+j} \leq \eta_t, \quad (8a)$$

$$\eta_{t+1} = \beta^{-1}(\eta_t - v_t \cdot v_t), \quad (8b)$$

where η_0 is given and η_t serves as a continuation pessimism bound at date t . In (8a), $\hat{E}_t(\cdot)$ denotes the conditional expectation taken with respect to the law of motion (7), which relative to (1) is distorted by the presence of v_t .

The second player is introduced as a device to determine the conditional mean distortions $\{v_t\}$ in a way that delivers a particular form of robustness. Letting v_t feed back on x_t , including its endogenous components, allows for a wide class of misspecifications. We want the feedback rule for i_t to be insensitive to mistakes v_t in the conditional mean of w_{t+1} . To promote insensitivity, we make the second player malevolent and instruct him to minimize \tilde{U}_0 over state feedback rules for v_t .

We impose restriction (8b) by formulating a multiplier game. In particular, we let $-1/\sigma \geq 0$ be a Lagrange multiplier on the time t constraint (8a) and require that the continuation pessimism level η_t be such that the multiplier is constant over time.⁹ Condition (8b) accomplishes this. This leads to a recursive formulation of the game. The Markov perfect equilibrium has a value function that satisfies

$$\begin{aligned} \tilde{W}(x) &= \inf_v \sup_i \left\{ -i'Qi - x'Rx + \beta \left[-\frac{1}{\sigma} v'v + E\tilde{W}(Ax + Bi + C(w + v)) \right] \right\} \\ &= x'\Omega x + \tilde{\rho}, \end{aligned} \quad (9)$$

where the E operator integrates w with respect to a normal distribution with mean zero and covariance matrix I . Hansen and Sargent (1998) show that the value functions W and \tilde{W} share the same matrix Ω in their quadratic forms, but have different constants ρ and $\tilde{\rho}$. Let $i = -Fx$, $v = Gx$ denote the policy rules that solve (9); the rules are linear, and the rule for i also solves the risk-sensitive control problem.¹⁰

The relationship between the two value functions and the decision rules for i establishes how the risk-sensitive preference specification induces the same behaviour that would occur *without* the risk-sensitivity adjustment to preferences, but *with* the pessimistic view of the conditional mean of innovations (the v_t 's) reflected in (9). The risk-sensitivity parameter σ sets the constant Lagrange multiplier $-\sigma^{-1}$ on restriction (8). Notice how η_0 indexes the degree of pessimism, *i.e.* the size of the domain of sequences from which the malevolent opponent selects adverse v_t 's. Hansen and Sargent (1998) describe in detail why it is convenient computationally to parameterize pessimism in this way.

“Uncertainty aversion” or robustness

The Markov perfect equilibrium summarized by (9) is the value function for a single decision maker whose decisions are governed by a “worst case” analysis. By using a feedback rule for i_t that solves (9), the robust controller does better for some appropriately constrained mistake sequences $\{v_t\}$ while sacrificing utility when these mistakes are absent. Our treatment of this robustness and its connection to risk sensitivity follows James'

9. See Hansen and Sargent (1998) for more details.

10. Hansen and Sargent (1998) discuss how the particular parameterization of “uncertainty aversion” embedded in (9)—in which the “Lagrange multiplier” $-\sigma^{-1}$ is time invariant—requires choosing the continuation pessimism bounds η_t in a way to make the opponent’s decision problem recursive.

(1995) recent survey of robust control, except that we have incorporated discounting into the risk sensitive formulation of the problem and into the corresponding constraints on the model misspecification.

There is a closely related literature in economics originating with the work of Gilboa and Schmeidler (1989) and Epstein and Wang (1994). The decision theory axiomatized by Gilboa and Schmeidler generalizes expected utility theory by studying a setting where decisions are based on a “maxmin” criterion because beliefs are described by a family of probability measures rather than a single probability measure. In our setup, there is a “nominal model” corresponding to setting $v_t = 0$ for all t . Alternative specification error sequences $\{v_t\}$ constrained by (8) deliver the resulting family of stochastic processes used in the state evolution equation. Hence our decision maker can be viewed as having preferences represented by the maxmin utility theory of Gilboa and Schmeidler. Following Epstein and Wang (1994), we can interpret the nonuniqueness of the stochastic constraints as depicting a form of Knightian uncertainty: an ambiguity of beliefs not fully specified in probabilistic terms but described by the set of specification errors $\{v_t\}$ defined by restriction (8).

In intertemporal contexts, Epstein and Wang (1994) use a Markov formulation of the two-player game to avoid inducing a form of time inconsistency. We follow the literature on robust control by holding fixed the Lagrange multiplier $-\sigma^{-1}$ on the specification error constraint over time. Below, we shall compute the v_t 's and use them to measure the amount of uncertainty aversion associated with alternative values of σ . We avail ourselves of a formula for the matrix G in $v = Gx$.

Solution for v

The solution for v within the Markov perfect equilibrium satisfies

$$\hat{v}_t = \sigma(I - \sigma C' \Omega C)^{-1} C' \Omega A^* x_t, \quad (10)$$

where $x_{t+1} = A^* x_t + C w_{t+1}$ under the optimal control law for the risk-sensitive problem ($A^* = A - BF$). (Here we are assuming that the parameter σ is sufficiently small that the matrix $(I - \sigma C' \Omega C)$ is positive definite.)¹¹

Below we shall compute \hat{v}_t and study how it alters measures of risk aversion extracted from asset prices.

Modified certainty equivalence

Whittle (1981) pointed out how the solution for v supports a modified version of certainty equivalence. This version asserts the equivalence of two ways of evaluating time-invariant decision rules $i_t = -Fx_t$, one under rational expectations and risk-sensitive preferences; the other under distorted expectations and ordinary ($\sigma = 0$) quadratic preferences. Recall that $A^* = A - BF$, and let $R^* = R + F'QF$. The two valuation algorithms are:

- (1) $U_t^e = -x_t' R^* x_t + \beta \mathcal{R}_t U_{t+1}^e$, where \mathcal{R}_t is defined in (3), and where the conditional expectation operator in (3) is computed with respect to the (true) law of motion

11. Although the matrix Ω depends implicitly on σ , it can be shown that the requisite positive definiteness will be satisfied for small values of σ . The risk-sensitive control theory literature draws attention to the *breakdown point* under which this positive definiteness property ceases to hold (e.g. see Glover and Doyle (1988)). At such points, the risk-adjusted recursive utility is $-\infty$ regardless of the controller's action. The general equilibrium aspects of our analysis lead us to look at much smaller risk corrections than are tolerated by the breakdown analysis.

$x_{t+1} = A^*x_t + Cw_{t+1}$. The criterion can be represented as the translated quadratic form $U_t^e = x_t' \Omega x_t + \rho$, where the matrix Ω and the scalar ρ are fixed points of operators defined by Hansen and Sargent (1995).

- (2) $\tilde{W}(x_t) = \tilde{U}_t = -x_t' R^* x_t + \beta \tilde{E}_t \tilde{U}_{t+1} - (\beta/\sigma) \hat{v}_t' \hat{v}_t$, where \tilde{E}_t is an expectation operator computed with respect to the distorted law of motion

$$x_{t+1} = \hat{A}x_t + Cw_{t+1}, \quad (11)$$

where

$$\hat{A} = [I + \sigma C(I - \sigma C' \Omega C)^{-1} C' \Omega] A^*, \quad (12)$$

and \hat{v}_t is given by (10). The formula for \hat{A} is derived by adding $C\hat{v}_t$ to A^* , where \hat{v}_t satisfies (10). The criterion \tilde{U}_t has the representation $\tilde{U}_t = x_t' \Omega x_t + \tilde{\rho}$, where Ω is the same matrix occurring in the first representation.

Evidently, these two evaluations yield the same ordering over time-invariant decision rules $i_t = -Fx_t$. This is the modified certainty equivalence principle. Notice the appearance of Ω , computed from the first formulation, in the construction of the distorted law of motion (12). We shall use \hat{A} from (12) again in computing asset prices.

3. ROBUST PERMANENT INCOME THEORY

Hall (1978), Campbell (1987), Heaton (1993), and Hansen, Roberds, and Sargent (1991) studied how closely a permanent income model approximates aggregate data on consumption and investment. We formulate a risk-sensitive version of the permanent income model with habit persistence, estimate it from data on consumption and investment, then use it to compare the implications of risk-sensitivity for consumption, investment, and asset prices. We demonstrate an observational equivalence proposition asserting that the consumption and investment data alone are insufficient simultaneously to identify the risk-sensitivity parameter σ and the subjective discount factor β . This observational equivalence substantiates our claim to be reinterpreting decision rules from a habit-persistence version of Hall's model in terms of robust decision making. Adding knowledge of the risk-free rate, which is constant in this model, does not achieve identification. But later we will show that the risk-sensitivity parameter has strong effects on other asset prices, including the market price of risk.

The lack of identification from consumption and investment data emerges as follows. For a given specification of shocks, introducing risk sensitivity provides an additional precautionary motive for saving. In terms of implications for savings, this motive can be offset by diminishing the subjective discount factor to make saving less attractive. In terms of effects on the valuation of risky assets, these changes are *not* offsetting.

The model

We formulate the model in terms of a planner with preferences over consumption streams $\{c_t\}_{t=0}^\infty$, intermediated through the service stream $\{s_t\}$. Preferences are ordered by the utility index U_0 , defined through the recursion

$$U_t = -(s_t - b_t)^2 + \beta \mathcal{R}_t(U_{t+1}), \quad (13)$$

where $\mathcal{R}_t(U_{t+1})$ is defined by (3).

In (13), s_t is a scalar household service produced by the scalar consumption c_t via the household technology

$$s_t = (1 + \lambda)c_t - \lambda h_{t-1}, \quad (14a)$$

$$h_t = \delta_h h_{t-1} + (1 - \delta_h)c_t, \quad (14b)$$

where $\lambda > 0$ and $\delta_h \in (0, 1)$. In (13), $\{b_t\}$ is an exogenous preference shock process. System (14) accommodates habit persistence or rational addiction as in Ryder and Heal (1973), Becker and Murphy (1988), Sundaresan (1989), Constantinides (1990) and Heaton (1993). By construction, h_t is a geometric weighted average of current and past consumption. Setting $\lambda > 0$ induces intertemporal complementarities. Consumption services depend positively on current consumption, but negatively on a weighted average of past consumptions, an embodiment of “habit persistence”.

There is a linear production technology

$$c_t + i_t = \gamma k_{t-1} + d_t,$$

where the capital stock k_t at the end of period t evolves according to

$$k_t = \delta_k k_{t-1} + i_t,$$

i_t is time t gross investment, and $\{d_t\}$ is an exogenously specified endowment process. The parameter γ is the (constant) marginal product of capital, and δ_k is the depreciation factor for capital. Solving the capital evolution equation for investment and substituting into the linear production technology gives

$$c_t + k_t = (\delta_k + \gamma)k_{t-1} + d_t. \quad (15)$$

We define

$$R \equiv \delta_k + \gamma,$$

which is the physical (gross) return on capital taking account of the fact that capital depreciates over time. When the economy is decentralized, R will also coincide with the gross return on a risk free asset. We impose that the components of the solution for $\{c_t, h_t, k_t\}$ belong to L_0^2 , the space of stochastic processes $\{y_t\}$ defined as

$$L_0^2 = \{y: y_t \text{ is in } J_t \text{ for } t = 0, 1, \dots \text{ and } E \sum_{t=0}^{\infty} R^{-t} (y_t)^2 | J_0 < +\infty\}.$$

We suppose that the endowment and preference shocks (d_t, b_t) are governed by $b_t = U_b z_t$, $d_t = U_d z_t$ where

$$z_{t+1} = A_{22} z_t + C_2 w_{t+1}.$$

Here w_{t+1} is independent of $J_t = \{w_t, w_{t-1}, \dots, w_1, z_0\}$, the eigenvalues of A_{22} are bounded in modulus by unity, and w_{t+1} is normally distributed with mean zero and covariance matrix I .

Given k_0 , the planner chooses a process $\{c_t, k_t\}$ with components in L_0^2 to maximize U_0 subject to (14), (15).¹²

12. We can convert this problem into a special case of the control problem posed in Section 2 as follows. Form a composite state vector x_t by stacking h_{t-1} , k_{t-1} and z_t , and let the control i_t be given by $s_t - b_t$. Solve (14a) for c_t as a function of $s_t - b_t$, b_t and h_{t-1} and substitute into equations (14b) and (15). Stack the resulting two equations along with the state evolution equation for z_t to form the evolution equation for x_{t+1} .

Solution of model and identification of σ

To establish observational equivalence for the quantity observations, we proceed constructively. First, we compute a solution for $\sigma = 0$ and $\beta R = 1$, *i.e.* a permanent income economy, without risk sensitivity. Then we use the allocation for this $\sigma = 0$ economy to construct an equivalence class of alternative (σ, β) 's that generate the same allocation, for fixed values of all the other parameters. This demonstrates that the pair (σ, β) is not identified from quantity observations alone.

The $\sigma = 0, \beta R = 1$ benchmark case

To produce a permanent income model in the $\sigma = 0$ special case, we follow Hall (1978) and impose that $\beta R = 1$. When $\sigma = 0$, (13) and (3) reduce to

$$U_0 = E_0 \sum_{t=0}^{\infty} \beta^t \{-(s_t - b_t)^2\}. \quad (16)$$

Formulate the planning problem as a Lagrangian by putting random Lagrange multiplier processes of $2\beta' \mu_{st}$ on (14a), $2\beta' \mu_{ht}$ on (14b), and $2\beta' \mu_{ct}$ on (15). First-order necessary conditions are

$$\mu_{st} = b_t - s_t, \quad (17a)$$

$$\mu_{ct} = (1 + \lambda)\mu_{st} + (1 - \delta_h)\mu_{ht}, \quad (17b)$$

$$\mu_{ht} = \beta E_t [\delta_h \mu_{ht+1} - \lambda \mu_{st+1}], \quad (17c)$$

$$\mu_{ct} = \beta R E_t \mu_{ct+1}, \quad (17d)$$

and also (14), (15). When $\beta R = 1$, equation (17d) implies that μ_{ct} is a martingale; then (17b) and (17c) solved forward imply that μ_{st} , μ_{ht} are also martingales. This implies that μ_{st} has the representation

$$\mu_{st} = \mu_{st-1} + v' w_t, \quad (18)$$

for some vector v .

Use (17a) to write $s_t = b_t - \mu_{st}$, substitute this into the household technology (14), and rearrange to get the system

$$c_t = \frac{1}{1 + \lambda} (b_t - \mu_{st}) + \frac{\lambda}{1 + \lambda} h_{t-1}, \quad (19a)$$

$$h_t = \tilde{\delta}_h h_{t-1} + (1 - \tilde{\delta}_h)(b_t - \mu_{st}), \quad (19b)$$

where $\tilde{\delta}_h = (\delta_h + \lambda)/(1 + \lambda)$. Equation (19b) can be used to compute

$$E_t \sum_{j=0}^{\infty} \beta^j h_{t+j-1} = (1 - \beta \tilde{\delta}_h)^{-1} h_{t-1} + \frac{\beta(1 - \tilde{\delta}_h)}{(1 - \beta \tilde{\delta}_h)} E_t \sum_{j=0}^{\infty} \beta^j (b_{t+j} - \mu_{st+j}). \quad (20)$$

For the purpose of solving the first-order conditions (17), (14) and (15) subject to the side condition that $\{c_t, k_t\} \in L_0^2$, treat the technology (15) as a difference equation in $\{k_t\}$, solve forward, and take conditional expectations on both sides to get

$$k_{t-1} = \sum_{j=0}^{\infty} R^{-(j+1)} E_t (c_{t+j} - d_{t+j}). \quad (21)$$

Use (19a) to eliminate $\{c_{t+j}\}$ from (21), then use (18) and (20). Solve the resulting system for μ_{st} , to get

$$\mu_{st} = (1 - R^{-1}) \sum_{j=0}^{\infty} R^{-j} E_t b_{t+j} + \psi_0 \sum_{j=0}^{\infty} R^{-j} E_t d_{t+j} + \psi_1 h_{t-1} + \psi_2 k_{t-1}, \quad (22)$$

where ψ_0 , ψ_1 , ψ_2 are constants. Equations (22), (19), and (15) represent the solution of the planning problem.

Notice that (22) makes μ_{st} depend on a geometric average of current and future values of b_t . Therefore, both the optimal consumption service process and optimal consumption depend on the difference between b_t and a geometric average of current and expected future values of b . So there is no “level effect” of the preference shock on the optimal decision rules for consumption and investment. However, the level of b_t will affect equilibrium asset prices.

Observational equivalence (for quantities) of $\sigma = 0$ and $\sigma \neq 0$

At this point, we state the following

Observational Equivalence Proposition. *Fix all parameters except β and σ . Suppose $\beta R = 1$. There exists a $\underline{\sigma} < 0$ such that the optimal consumption-investment plan with $\sigma = 0$ is also the optimal consumption-investment plan for any σ satisfying $\underline{\sigma} < \sigma < 0$ and a smaller discount factor $\hat{\beta}(\sigma)$ that varies directly with σ .*

This proposition means that, so far as the quantities $\{c_t, k_t\}$ are concerned, the risk-sensitive ($\sigma < 0$) version of the permanent income model is observationally equivalent to the benchmark ($\sigma = 0$) version. This insight will guide our estimation strategy, because it sharply partitions the impact of risk-sensitivity into *real* and *pricing* parts.

The proof of the proposition is by construction.

Proof. This is the plan of the proof. Begin with a solution $\{\bar{s}_t, \bar{c}_t, \bar{k}_t, \bar{h}_t\}$ for a benchmark $\sigma = 0$ economy. Form a comparison economy with a $\sigma \in [\underline{\sigma}, 0]$, where $\underline{\sigma}$ is the boundary of an admissible set of σ 's to be described below. Fix all parameters except (σ, β) the same as in the benchmark economy. Conjecture that $\{\bar{s}_t, \bar{c}_t, \bar{k}_t, \bar{h}_t\}$ is also the optimal allocation for the $\sigma < 0$ economy. Finally, construct a $\beta = \hat{\beta}$ that verifies this conjecture.

Here are the details of the construction. The optimality of the allocation implies that $E_t \mu_{ct+1} = \mu_{ct}$, and that (18) and (22) are satisfied for the (−) benchmark allocation, where E_t is the expectation operator under the *correct* probability measure. The key idea is to form the distorted expectation operator \hat{E}_t , then choose $\beta = \hat{\beta}$ to make the distorted version of the Euler equation for μ_{ct} hold at the benchmark ($\sigma = 0$) allocation.

To compute the distorted expectation operator, we follow the recipe given in formulas (9) and (12). First, we have to evaluate the utility index U_0 by using (9). We want to evaluate (13) with $s_t - b_t \equiv -\mu_{st}$ and μ_{st} given by the law of motion (18), which we take as exogenous because the allocation is frozen. We take μ_{st} as the state. Since there is no control, (9) collapses to

$$\Omega x^2 = -x^2 + \beta \min_v \left(-\frac{1}{\sigma} v^2 + \Omega(x + \theta v)^2 \right), \quad (23)$$

and we write $\mu_{st} = \mu_{st-1} + \theta(v + w)$, where $\theta^2 = v'v$ and v is the specification error chosen by the “opponent” in the fictitious game. The scalar Ω that solves (23) is

$$\Omega(\beta) = \frac{\beta - 1 + \sigma\theta^2 + \sqrt{(\beta - 1 + \sigma\theta^2)^2 + 4\sigma\theta^2}}{-2\sigma\theta^2}. \quad (24)$$

It follows from (12) that the distorted law of motion for μ_{st} is

$$\hat{E}_t \mu_{st+1} = \hat{\zeta} \mu_{st}, \quad (25)$$

where

$$\hat{\zeta} = \hat{\zeta}(\beta) = 1 + \frac{\theta^2 \sigma \Omega(\beta)}{1 - \sigma \theta^2 \Omega(\beta)}. \quad (26)$$

Since μ_{ct} is proportional to μ_{st} , it follows that

$$\hat{E}_t \mu_{ct+1} = \hat{\zeta} \mu_{ct}, \quad (27)$$

with the same $\hat{\zeta}$ given by (26). In terms of the distorted expectation operator, the Euler equation for capital is

$$\hat{\beta} R \hat{E}_t \mu_{ct+1} = \mu_{ct},$$

or

$$\hat{\beta} R \hat{\zeta}(\hat{\beta}) = 1. \quad (28)$$

Let $\underline{\sigma}$ be the lowest value for which the solution of (24) is real. Then given $\sigma \in (\underline{\sigma}, 0]$, there exists a $\hat{\beta}$ satisfying (28) such that for $(\sigma, \hat{\beta})$ the benchmark allocation solves the risk-adjusted problem. Therefore equations (24), (26), and (28) define a locus of $(\sigma, \hat{\beta})$'s, each point of which is observationally equivalent to $(0, \beta)$ for (c_t, k_t) observations, because each supports the benchmark allocation.

Furthermore, according to the asset pricing theory to be developed shortly and (28), the price of a sure claim on consumption one period ahead is R^{-1} for all t and all $(\sigma, \hat{\beta})$ in the locus. Therefore, these different parameter pairs are also observationally equivalent with respect to the risk-free rate.¹³ ||

In Figure 1, we report the (σ, β) pairs that are observationally equivalent for our maximum likelihood estimates for the remaining parameters, which we are about to describe.

The observational equivalence depicted in Figure 1 shows that by lowering the discount factor, we can make investment less attractive and thereby offset the precautionary savings motive. As an indication of the important precautionary role for savings in this model, suppose that future endowments and preference shifters could be forecast perfectly. Then consumers would choose to draw down their capital stock. Investment would be sufficiently unattractive that the optimal linear rule would eventually have both consumption and capital cross zero.^{14,15} Thus our robust control interpretation of the permanent income decision rule delivers a form of precautionary savings absent under the usual interpretation.

13. In this model, the technology (15) ties down the risk-free rate. For a version of the model with quadratic costs of adjusting capital, the risk-free rate comes to depend on σ , even though the observations on quantities are nearly independent of σ . See Hansen and Sargent (1996).

14. Introducing nonnegativity constraints in capital and/or consumption would induce nonlinearities into the consumption and savings rules, especially near zero capital. But investment would remain unattractive in the presence of those constraints for experiments like the ones we are describing here. See Deaton (1991) for a survey and quantitative assessment of consumption models with binding borrowing constraints.

15. As emphasized by Carroll (1992), even when the discount factor is small relative to the interest rate, precautionary savings can emerge when there is a severe utility cost for zero consumption. Such a utility cost is absent in our formulation.

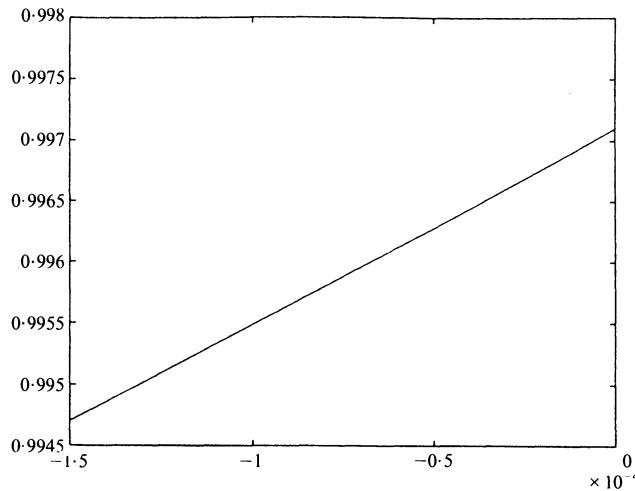


FIGURE 1

Observationally equivalent (σ, β) pairs for maximum likelihood values of identified parameters; σ is the ordinate, β the coordinate

For any given pair (σ, β) depicted in Figure 1, the permanent income decision rule reflects either risk sensitivity or a concern for robustness. The familiar version of the precautionary savings motive focuses on the role of variation in the shocks. This version is delivered in our setup by the risk sensitive decision theoretic formulation. In contrast, the precautionary notion delivered by robust control theory emerges because consumers guard against mistakes in conditional means of shocks. Thus concern for robustness shifts emphasis from second to first moment properties of shocks.

4. ESTIMATION

Different observationally equivalent (σ, β) pairs identified by our proposition bear different implications about (i) the pricing of risky assets; (ii) the amounts required to compensate the planner for confronting different amounts of risk; (iii) the amount of model misspecification used to justify the planner's decisions if risk sensitivity is reinterpreted as aversion to Knightian uncertainty. To evaluate these implications, we first choose parameters, including noise variances, by estimating a $\sigma = 0$ version of our permanent income model, conditioning the likelihood function only on U.S. post-war quarterly consumption and investment data. We estimated the permanent-income model with habit persistence using U.S. quarterly data on consumption and investment for the period 1970I–1996III.¹⁶

Consumption is measured by nondurables plus services, while investment is measured by the sum of durable consumption and gross private investment.¹⁷ We applied the model

16. Our choice of starting the sample in 1970 corresponds to the second subsample analyzed by Burnside, Eichenbaum and Rebelo (1990). Thus we have omitted the earlier period of "higher productivity". We initially estimated a version of the model with a stochastic preference shock over the entire post-war time period, but we found that the "productivity slowdown" was captured in our likelihood estimation by an initial slow decline in the preference shock process followed by a slow increase. Our illustrative permanent income model is apparently not well suited to capture productivity slowdowns. Given the empirical results reported in Burnside, Eichenbaum and Rebelo (1990), the same could be said of the commonly used stochastic specification of Solow's growth model.

17. We used "old data", not chain-weighted indexes.

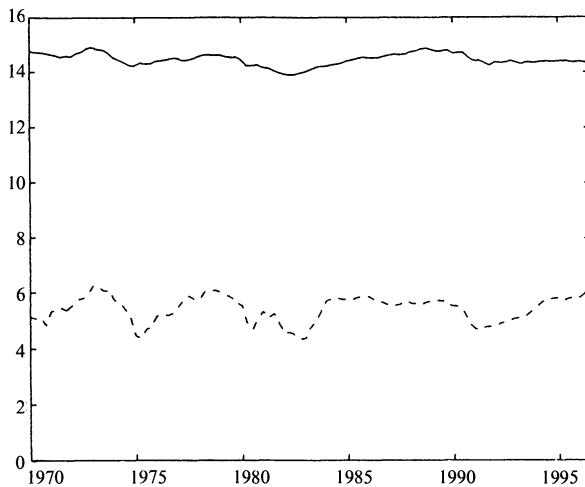


FIGURE 2
Detrended consumption and investment (dotted line) data

to data that have been scaled through multiplication by 1.0033^{-t} . The scaled time series are plotted in Figure 2. We estimated the model from data on (c_t, i_t) , setting $\sigma = 0$, then deduced pairs (σ, β) that are observationally equivalent. We estimated parameters by climbing a Gaussian likelihood function. We formed the likelihood function recursively, and estimated the unobserved part of the initial state vector using procedures described by Hansen and Sargent (1996).

Under our robustness interpretation, this approach to estimation may be justified in one of two ways. First, economic agents may allow for model misspecification when making their decisions, even though in fact the model is specified correctly during the sample period. Alternatively, economic agents use the (misspecified) maximum likelihood criterion for selecting a baseline model around which they entertain small specification errors. Under this second interpretation, the formal statistical inference formulas for maximum likelihood estimation require modification (see White (1982)).

We specified a constant preference shifter $b_t = \mu_b$ and a bivariate stochastic endowment process: $d_t = \mu_d + d_t^* + \hat{d}_t$.¹⁸ Because we are modelling two observed time series as functions of two shock processes, the model would lose its content were we to permit arbitrary cross correlation between the two endowment processes. Therefore, we assumed that these processes are orthogonal. We found that one of the shock processes, d_t^* was particularly persistent, with an autoregressive root of 0.998. While we doubt that this value is distinguishable from unity, we retained the unconstrained estimate of 0.998 in our subsequent calculations. The two shocks are parameterized as second order auto-regressions. We write them as

$$(1 - \phi_1 L)(1 - \phi_2 L)d_t^* = c_{d^*} w_t^{d^*},$$

$$(1 - \alpha_1 L)(1 - \alpha_2 L)\hat{d}_t = c_{\hat{d}} w_t^{\hat{d}}.$$

18. A previous draft specified two stochastic shock processes: an endowment shock, d_t , and a preference shock, b_t . We have chosen to report results for the bivariate endowment process with a constant preference shifter b in response to a comment from one of the anonymous referees. The results from the preference shock version of our model are available in an earlier version of this paper available at <http://riffle.stanford.edu>.

TABLE 1

Likelihood values

Transitory endowment specification	2 × Log Likelihood
AR1	776.78
AR2	779.05
AR3	779.05

For the transitory process \hat{d} we experimented with autoregressive processes of order 1, 2, and 3, which revealed the log likelihood values depicted in Table 1. In the table, ‘‘AR1’’ denotes the first-order autoregression, and so on. The likelihood values show a substantial gain in increasing the order from 1 to 2, but negligible gain in going from 2 to 3. These results led us to specify a second order autoregression for the transitory endowment process.

Thus the forcing processes are governed by seven free parameters: $(\alpha_1, \alpha_2, c_d, \phi_1, \phi_2, c_{d^*}, \mu_d)$. We use the parameter μ_b to set the bliss point. While μ_b alters the marginal utilities, as we noted previously, it does not influence the decision rules for consumption and investment. Consequently, we fixed μ_b at an arbitrary number, namely 32, in our estimation.

The four parameters governing the endogenous dynamics are: $(\gamma, \delta_h, \beta, \lambda)$. We set $\delta_k = 0.975$. We initially did not impose the permanent income restriction, $\beta R = 1$, but the restriction was satisfied by our estimates, so we proceeded to impose it. That is, our estimates confirmed the random walk prediction for both the marginal utility process for consumption goods and the marginal utility process for consumption services. The restrictions that $\beta R = 1$, $\delta_k = 0.975$ pin down γ once β is estimated. We chose to impose $\beta = 0.9971$, which after adjustment for the effects of the geometric growth factor of 1.0033 implies an annual real interest rate of 2.5%.¹⁹

Maintaining the $\beta R = 1$ restriction, we estimated the model for different values of γ (and therefore of β). The likelihood increases moderately as γ rises (and β decreases) over a large range of γ 's. However, over this range other parameters of the model do not change much. Allowing β to decrease below the value 0.9971 would have the primary effect on our results of increasing the risk-free rate above the already excessive value of 2.5% per year. Therefore, we chose to fix β at 0.9971.

In Table 2 we report our estimates for the parameters governing the endogenous and exogenous dynamics. In Figure 3 we report impulse response functions for consumption and investment to innovations in both components of the endowment process. For sake of comparison, we also report estimates from a no habit persistence ($\lambda = 0$) model in Table 2, and the resulting impulse response functions in Figure 4.

Notice that the persistent endowment shock process contributes much more to consumption and investment fluctuations than does the transitory endowment shock process.

To assess the statistical evidence for habit persistence, in Figure 5(a) we graph twice the concentrated log likelihood as a function of the habit persistence parameter. Notice the asymmetry of this function, which has a much steeper descent towards zero. A likelihood-based confidence interval can be deduced by comparing the likelihood deterioration

19. When $\sigma = 0$ (the expected utility, rational expectations case) we can scale the state variables to account for geometric growth without affecting the subsequent analysis. However, when $\sigma < 0$, the same transformation has the effect of imposing a time-varying risk adjustment. This problem does not arise when the single period utility function has a different form, say logarithmic. In order to preserve the tractability of the quadratic specification, we have decided to proceed despite this problem.

TABLE 2
Parameter Estimates

	Habit persistence	No habit persistence
Risk free rate	0.025	0.025
β	0.997	0.997
δ_h	0.682	0
λ	2.443	0
α_1	0.813	0.900
α_2	0.189	0.241
ϕ_1	0.998	0.995
ϕ_2	0.704	0.450
μ_d	13.710	13.594
c_d	0.155	0.173
c_{d^*}	0.108	0.098
2 × Log Likelihood	779.05	762.55

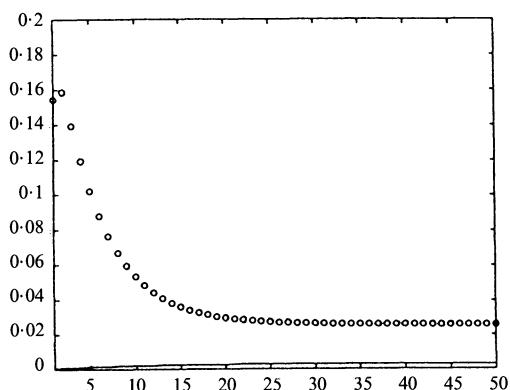


FIGURE 3(a)

Impulse response of investment (circles) and consumption (line) to innovation in transitory endowment process (\hat{d}), at maximum likelihood estimate of habit persistence

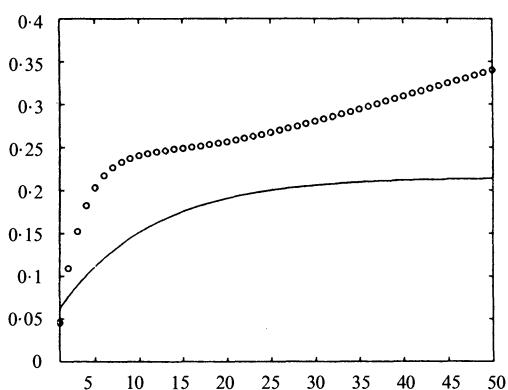


FIGURE 3(b)

Impulse response of investment (circles) and consumption (line) to innovation in persistent shock (d^*), at maximum likelihood estimate of habit persistence

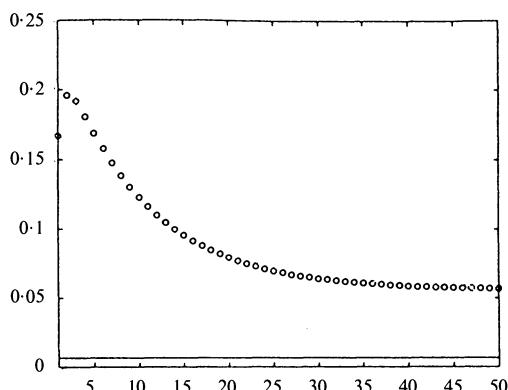


FIGURE 4(a)

Impulse response of investment (circles) and consumption (line) to innovation in transitory endowment (\hat{d}), no habit persistence

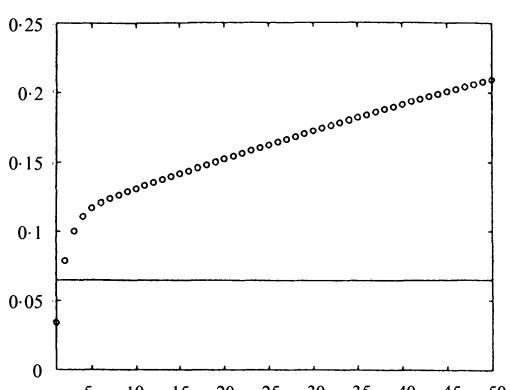


FIGURE 4(b)

Impulse response of investment (circles) and consumption (line) to innovation in persistent shock (d^*), no habit persistence

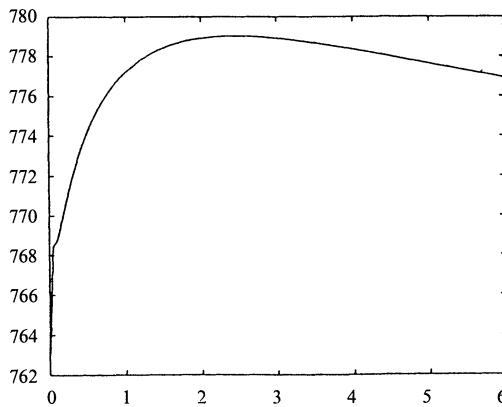


FIGURE 5(a)

Twice log likelihood, the coordinate, as a function of λ , the ordinate (other parameters being concentrated out)

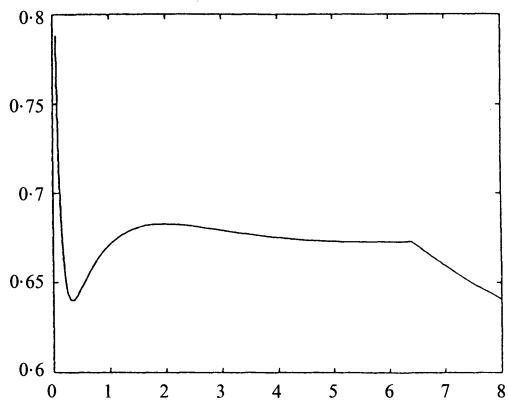


FIGURE 5(b)

Maximum likelihood δ_h , the coordinate, as a function of λ , the ordinate

to critical values obtained from the chi-square one distribution. Thus, while values of λ near zero are implausible, values considerably larger than the maximum likelihood values are harder to dismiss.²⁰ Figure 5(b) shows the values of the depreciation parameter δ_h as a function of the λ obtained after concentrating the likelihood function. Estimates of the depreciation parameter decrease as λ approaches zero, but remain around 0.65, within the more plausible range of λ 's.

We put our estimates of the habit persistence parameters, λ and δ_h , into perspective by comparing them with ones emerging from other empirical studies of aggregate U.S. data. Heaton (1993) finds a comparable value of λ , but a higher depreciation factor δ_h using a permanent income model without preference shocks fit to consumption. Heaton also notes that his δ_h is estimated very imprecisely.²¹ As an extension to this work, Heaton (1995) estimates a power utility, habit persistence model using consumption and asset market data. In this alternative formulation, he provides evidence for larger values of λ and a larger depreciation factor δ_h . Again the estimate of δ_h has a large standard error. From Heaton's work, we see that more pronounced habit persistence is estimated only when it is offset in the short run by local durability, a source of dynamics that we ignore. Recently, Boldrin, Christiano and Fisher (1995) find smaller values of λ and δ_h than ours, although they model production in a different and maybe more interesting way than we do. In contrast to Heaton (1995) and Boldrin, Christiano and Fisher (1995), our estimates of habit persistence embody no asset market implications beyond one for the risk free interest rate.

5. ASSET PRICING

For the purposes of decentralization, we regard the robust (or risk-sensitive) solution to the permanent income model as the solution to an optimal resource allocation problem. This viewpoint permits us to compute the equilibrium stochastic process of quantities

20. The parameter δ_h is not identified when $\lambda = 0$.

21. Like Christiano, Eichenbaum and Marshall (1991), Heaton (1993) also studies the implications of time aggregation, which we abstract from, and at the same time he allows for local durability in a continuous-time formulation of the model.

before deducing the prices that clear a competitive security market. We follow Lucas (1978) in assuming a large number of identical agents who trade in security markets. We can price assets by treating the consumption process that solves the robust permanent income model as though it were an endowment process. Because agents are identical, equilibrium prices become shadow prices that leave consumers content with that “endowment process”. The pricing implications under robustness are slightly different than those under risk-sensitivity. We will proceed in this section by assuming risk-sensitivity and pointing out where the analysis would differ under robustness.

The state for the model is $x_t = [h_{t-1} \ k_{t-1} \ z'_t]'$. The equilibrium consumption and service processes can be represented as $c_t^e = S_c x_t$, $s_t^e = S_s x_t$. Represent the endowment and preference shock processes as $d_t = S_d x_t$, $b_t = S_b x_t$. The equilibrium law of motion for the state has representation

$$x_{t+1} = A^0 x_t + C w_{t+1}. \quad (29)$$

The value function at the optimal allocation can be represented as $U_t^e = x_t' \Omega x_t + \rho$ where

$$\Omega = -(S_s - S_b)'(S_s - S_b)/2 + \beta \hat{\Omega}, \quad (30a)$$

$$\rho = \beta \hat{\rho}, \quad (30b)$$

and $\hat{\Omega}$ satisfies (6a), with A^* evaluated at A^0 .

Key subgradient inequality

We begin our analysis of asset pricing by computing the current time t price of a state-contingent claim to utility U_{t+1} tomorrow. This component of pricing is trivial when preferences are represented as the usual recursive version of the von Neumann–Morgenstern specification, but is nontrivial in the case of risk sensitivity. The pricing of state-contingent utility will be a key ingredient for pricing state-contingent consumption services tomorrow and ultimately for the pricing of multi-period securities that are direct claims on consumption goods. Let s_t be any service process measurable with respect to J_t , and U_t be the associated utility index. For purposes of valuation, Appendix A establishes the following subgradient inequality

$$\mathcal{R}_t(U_{t+1}) - \mathcal{R}_t(U_{t+1}^e) \leq \mathcal{A}_t U_{t+1} - \mathcal{A}_t U_{t+1}^e, \quad (31)$$

where

$$\mathcal{A}_t U_{t+1} \equiv E(V_{t+1} U_{t+1} | J_t) / E(V_{t+1} | J_t), \quad (32)$$

and

$$V_{t+1} \equiv \exp(\sigma U_{t+1}^e / 2). \quad (33)$$

As elaborated further below, the operator \mathcal{A} acts much like a conditional expectation.²² Combining (31) with the familiar gradient inequality for quadratic functions, it follows that

$$U_t - U_t^e \leq (s_t - s_t^e) \mathcal{M}_t^s + \beta \mathcal{A}_t(U_{t+1} - U_{t+1}^e), \quad (34)$$

22. Depicting prices of derivative claims using distorted expectations is a common technique in asset pricing (e.g. see Harrison and Kreps (1979)). In our investigation and in Epstein and Wang (1994), the distortion is also needed to price state-contingent utility.

where

$$\mathcal{M}_t^s \equiv 2(b_t - s_t^e). \quad (35)$$

If we regard the marginal utility of services \mathcal{M}_t^s as the price for time t services, then (34) states that any pair (s_t, U_{t+1}) that is preferred to (s_t^e, U_{t+1}^e) , costs more at time t . This justifies treating \mathcal{M}_t^s as the equilibrium time t price of services, and using $\beta \mathcal{T}_t$ to value time $t+1$ state-contingent utility.

The \mathcal{T}_t operator can be computed as the conditional expectation of the state in the transformed transition equation

$$x_{t+1} = \hat{A}x_t + \hat{C}w_{t+1}, \quad (36)$$

where \hat{C} satisfies

$$\hat{C}\hat{C}' = C(I - \sigma C'\Omega C)^{-1}C', \quad (37)$$

and \hat{A} is given by (12). Given the matrices \hat{A} and \hat{C} , asset prices can be computed using the algorithms described in Hansen and Sargent (1996). Formula (37) shows that when $\sigma < 0$ and Ω is negative semidefinite, the conditional variance associated with the operator \mathcal{T}_t is always greater than or equal to CC' , because an identity matrix is replaced by a larger matrix $(I - \sigma C'\Omega C)^{-1}$. Thus, to interpret \mathcal{T}_t as a conditional expectation operator requires both a pessimistic assignment of the conditional mean for the future state vector and an increase in its conditional variance.²³

We can interchange the risk sensitivity and the uncertainty aversion interpretations of the optimal resource allocation problem. As shown by Epstein and Wang (1994), equilibrium asset prices can be deduced by referring to the “pessimistic beliefs” that implement optimal decisions. For the uncertainty aversion interpretation, the counterpart to the \mathcal{T}_t operator is the distorted conditional expectation operator, call it \hat{E}_t , induced by the state transition equation of formula (11). This transition law distorts the conditional mean, but not the conditional variance.²⁴

Pricing multi-period streams

The valuation of the state-contingent utility can be used to evaluate future consumption services. Construct a family of operators by sequential application of \mathcal{T}_t

$$\mathcal{S}_{t,\tau} = \mathcal{T}_t \mathcal{T}_{t+1} \cdots \mathcal{T}_{t+\tau-1}, \quad (38)$$

where $\mathcal{S}_{t,0}$ is the identity map. Like \mathcal{T}_t , $\mathcal{S}_{t,\tau}$ can be interpreted as a conditional expectation under a transformed conditional probability measure except that $\mathcal{S}_{t,\tau}$ is a time t conditional expectation applied to random variables that are measurable with respect to $J_{t+\tau}$.

In the permanent income model below, the consumption good is a bundle of claims to future consumption services. We can use the equilibrium prices of services to deduce corresponding prices of consumption goods. Thus, consider any process $\{s_t\}$ with components in L_0^2 , and let $\{U_t\}$ denote the associated utility process. Let $\{U_t^e\}$ denote the utility

23. It follows from James (1992) that this covariance correction vanishes in the continuous-time formulation of the problem. Instead the original covariance structure is used.

24. Epstein and Wang (1994) consider different ways of introducing Knightian uncertainty, including ones in which there is an important difference between the game with time zero commitment and the game with sequential choice. Their specification of Knightian uncertainty can result in two-person games in which the “beliefs” are not unique. This leads them to a form of price indeterminacy, which they link to empirical findings of excess volatility. In our setup, the “beliefs” turn out to be unique and price indeterminacy is absent.

process associated with the equilibrium service process $\{s_t^e\}$. Then by iterating on (34), we find

$$U_t - U_t^e \leq \sum_{\tau=0}^{\infty} \beta^\tau \mathcal{S}_{t,\tau}(\mathcal{M}_{t+\tau}^s s_{t+\tau}) - \sum_{\tau=0}^{\infty} \beta^\tau \mathcal{S}_{t,\tau}(\mathcal{M}_{t+\tau}^s s_{t+\tau}^e). \quad (39)$$

Inequality (39) says that whenever $\{s_t\}$ is strictly preferred to $\{s_t^e\}$ as reflected by the associated time zero utility index, ($U_t > U_t^e$), it also costs more. Hence $\{s_t^e\}$ is a solution to the consumer's intertemporal optimization problem when the time t value of $\{s_t\}$ is computed according to the formula $\sum_{\tau=0}^{\infty} \beta^\tau \mathcal{S}_{t,\tau}(\mathcal{M}_{t+\tau}^s s_{t+\tau})$. This justifies regarding this sum as the price of an asset offering a claim to the stream of services $\{s_t\}$.

If services are not traded "unbundled", but only as bundles of state and date contingent claims, via the consumption goods, then what we really want is a consumption goods counterpart to (39), namely

$$U_t - U_t^e \leq \sum_{\tau=0}^{\infty} \beta^\tau \mathcal{S}_{t,\tau}(\mathcal{M}_{t+\tau}^c c_{t+\tau}) - \sum_{\tau=0}^{\infty} \beta^\tau \mathcal{S}_{t,\tau}(\mathcal{M}_{t+\tau}^c c_{t+\tau}^e). \quad (40)$$

A formula for the indirect marginal utility of consumption is deduced by ascertaining the implicit service flow associated with a unit of consumption and then pricing that intertemporal bundle. Using this argument, it follows that $\mathcal{M}_t^c = M_c x_t$ where

$$M_c \equiv 2[(1 + \lambda) + (1 - \delta_h) \sum_{\tau=1}^{\infty} \beta^\tau (\delta_h)^\tau (-\lambda)(\bar{A})^\tau](S_b - S_s). \quad (41)$$

Single-period security pricing

A large body of empirical research has focussed on pricing one-period securities. Imagine purchasing a security at time t at a price q_t , holding it for one time period, then collecting the dividend and selling it at time $t+1$ for a total payoff p_{t+1} of the consumption good. The payoff and price should satisfy

$$q_t = \mathcal{R}\{[\beta \mathcal{M}_{t+1}^c / \mathcal{M}_t^c] p_{t+1}\}, \quad (42)$$

where $\mathcal{M}_t^c = M_c x_t$ is the marginal utility of consumption and the formula for M_c is given in (41). Under robustness, the price-payoff relationship would be given by

$$q_t = \tilde{E}_t\{[\beta \mathcal{M}_{t+1}^c / \mathcal{M}_t^c] p_{t+1}\}, \quad (43)$$

where \tilde{E}_t is the distorted conditional expectations operator described above. A formula for q_t in terms of the original conditional expectation operator is

$$q_t = E(m_{t+1,t} p_{t+1} | J_t), \quad (44)$$

where the exact specification of $m_{t+1,t}$ will depend upon whether the robustness or the risk-sensitivity interpretation is adopted. The two alternatives will be explored in the next section. The random variable $m_{t+1,t}$ has an interpretation as a one-period stochastic discount factor, or alternatively as an equilibrium intertemporal marginal rate of substitution for the consumption good. The next section will show how risk-sensitivity and uncertainty aversion are reflected in the usual measure of the intertemporal marginal rate of substitution being scaled by a random variable (that depends on the interpretation—robustness or risk-sensitivity) with conditional expectation one. We use this multiplicative adjustment to the stochastic discount factor to increase its variability and to enhance risk premia.

From the one-period stochastic discount factor, we can easily deduce the "market price of risk." For simplicity, think of a one period payoff on an asset as a bundle of two

attributes: its conditional mean and its conditional standard deviation. In our environment, these two attributes only partially describe asset payoffs. Furthermore, we cannot extract unique prices of the attributes, in part because one of the attributes, the standard deviation, is a nonlinear function of the asset payoff. Nevertheless, like any stochastic discount factor model, ours conveys information about how these attributes are valued (see Hansen and Jagannathan, 1991). To see this, consider the covariance decomposition of the right-hand side of (42)

$$q_t = E_t(p_{t+1})E_t(m_{t+1,t}) + \text{cov}_t(m_{t+1,t}, p_{t+1}),$$

where cov_t denotes the covariance conditioned on time t information. Applying the Cauchy–Schwarz Inequality, we obtain the price bound

$$q_t \geq E_t(p_{t+1})E_t(m_{t+1,t}) - \text{std}_t(m_{t+1,t}) \text{std}_t(p_{t+1}),$$

where std_t denotes the standard deviation conditioned at time t . Along the so-called “efficient frontier”, the “price of risk” relative to expected return is given by the ratio: $\text{std}_t(m_{t+1,t})/E_t(m_{t+1,t})$ which is commonly referred to as the market price of risk. This ratio is one way to encode information about how risk averse consumers are at the equilibrium consumption process.²⁵ Appendix C describes how to compute the stochastic process for the market price of risk when σ is negative under risk-sensitivity.

6. QUANTIFYING ROBUSTNESS FROM THE MARKET PRICE OF RISK

Because it is not identified from data on consumption and investment, other information must be used to restrict the risk sensitivity parameter. In this section, we study how risk sensitivity alters the predicted market price of risk. We then exploit the connection between risk sensitivity and Knightian uncertainty by computing the magnitude of the specification errors needed to generate implications comparable to various settings of the parameter σ . In particular, we show how allowing for mistakes transmits to the equilibrium market price of risk. We are attracted to the interpretation in terms of robustness as a way of confronting an observation of Weil (1989), who noted how market prices of risk can be enhanced by risk sensitivity, but at the cost of making the implied risk aversion “extreme”. Risk aversion has typically been measured by studying choice problems with unique specifications of the probability laws. That our risk sensitivity parameter has a nearly equivalent interpretation as reflecting aversion to uncertainty raises hopes for reinterpreting implausibly large estimates of risk aversion as coming partly from a “preference for robustness.”

Market price of risk

While the risk-sensitivity parameter σ and the preference curvature parameter μ_b are not identifiable from quantity data, we now show that they affect the market price of risk. In Tables 3(a) and 3(b), we report median market prices of risk as functions of the risk sensitivity parameter for three choices of μ_b . The tables are constructed using the implied state vectors obtained by applying the Kalman filter. Where $y_t = [c_t \ i_t]'$, and $x_t =$

25. Gallant, Hansen and Tauchen (1990), Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) interpret the *equity premium puzzle* as the large market price of risk implied by asset market data. The market price of risk can be expressed as the least upper bound on Sharpe ratios $|E_t r_{t+1} - r_t^f|/\text{std}_t(r_{t+1})$ where r_{t+1} is a one-period return and r_t^f is the one-period riskless return. Thus the Sharpe ratio for the one-period return on equity gives a lower bound on the market price of risk.

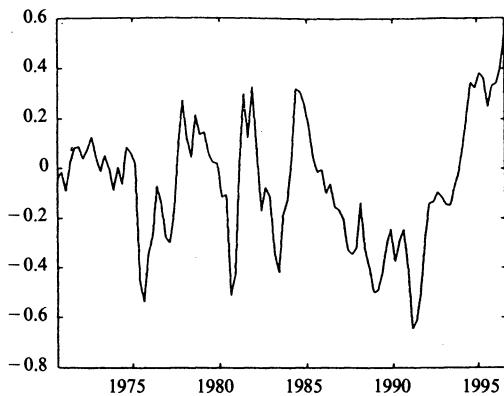


FIGURE 6(a)

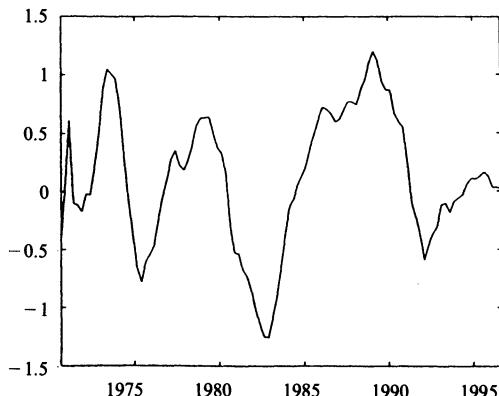
Estimated process or transitory endowment, \hat{d}_t 

FIGURE 6(b)

Estimated process for the permanent endowment process, d^*

$[h_{t-1} \ k_{t-1} \ 1 \ \hat{d}_t \ \hat{d}_{t-1} \ d_t^* \ d_{t-1}^*]'$, we used the Kalman filter to compute $E(x_t | y_t, y_{t-1}, \dots, y_1)$ for each time t in our sample. It can be shown that the conditional covariance of the time t state vector given time t information converges to zero, implying that the "hidden" states should be approximately revealed by the observations. Deviations around the means of the implied endowment processes under habit persistence are graphed in Figure 6. We used these fitted states to calculate the median market price of risk over the sample. In Tables 3(a) and 3(b), we report results for the model estimated with and without habit persistence, respectively. The tables show how we can achieve a "target" market price of risk with alternative (σ, μ_b) pairs.

Given our high value of the risk free rate (2.5% per annum) and sampling error in estimates of the market price of risk, model predictions in the range of 0.075–0.150 seem a reasonable "target."²⁶ Thus in the absence of risk sensitivity, for the μ_b specifications

TABLE 3(a)

Median market price of risk (with habit persistence)

μ_b	$\sigma:$	0	-0.00005	-0.0001	-0.00015
18		0.0610	0.0739	0.0869	0.1000
24		0.0226	0.0575	0.0927	0.1281
30		0.0139	0.0708	0.1283	0.1865
36		0.1000	0.0890	0.1691	0.2509

TABLE 3(b)

Median market price of risk (no habit persistence, $\lambda = 0$)

μ_b	$\sigma:$	0	-0.00005	-0.0001	-0.00015
18		0.0182	0.0221	0.0261	0.0300
24		0.0068	0.0173	0.0279	0.0385
30		0.0042	0.0213	0.0385	0.0557
36		0.0030	0.0268	0.0506	0.0745

26. It is known from the work of Hansen and Jagannathan (1991) that achieving a market price of risk target is weaker than satisfying the consumption Euler equation. For example, we have not enabled the model to explain one of the glaring empirical failures of consumption-based asset pricing models: the observed lack of correlation between the implied intertemporal marginal rates of substitution and stock market returns. For a description of how to build statistical tests based on market price of risk targets, see Burnside (1994), Cecchetti, Lamb and Mark (1994), and Hansen, Heaton and Luttmer (1995).

we consider, the market prices of risk are very small. The market price of risk can be raised by reducing further the parameter μ_b , but at the cost of enhancing the probability of satiation in the quadratic preference ordering. But increasing $|\sigma|$ pushes the model predictions towards more empirically plausible market prices of risk without altering the satiation probabilities.²⁷ Roughly speaking, introducing habit persistence triples (or multiples by $(1 + \lambda)$) the market price of risk across all of the (μ_b, σ) specifications that we study. This conclusion from Table 3(b) emerges from the estimates from the second (No habit persistence) column of Table 2. There the parameters governing the exogenous dynamics are adjusted to match the temporal covariations of consumption and investment as closely as possible.

Holding fixed σ and increasing the preference translation parameter μ_b also enhances the market price of risk except when σ is close to zero. To understand this finding, note that under risk sensitivity, the stochastic discount factor can be represented as the product

$$m_{t+1,t} = m_{t+1,t}^f m_{t+1,t}^r, \quad (45)$$

where

$$m_{t+1,t}^r \equiv \beta \frac{\mathcal{M}_{t+1}^c}{\mathcal{M}_t^c},$$

is the “familiar” intertemporal marginal rate of substitution in the absence of risk sensitivity and

$$m_{t+1,t}^r \equiv \frac{\exp(\sigma U_{t+1}^e / 2)}{E[\exp(\sigma U_{t+1}^e / 2) | J_t]}.$$

(See Appendix C for an explicit formula for $m_{t+1,t}$ in terms of the equilibrium laws of motion.) When $\sigma = 0$ this second term is one, and it always has conditional expectation equal to one. The latter property is what permits us to interpret this second factor as a pessimistic “distortion” of the conditional expectation operator. Finally, recall that the market price of risk is simply the (conditional) standard deviation of $m_{t+1,t}$ divided by its (conditional) mean.

When μ_b is increased and $\sigma = 0$, the single-period utility function is closer to being linear (risk neutral) over the empirically relevant portion of its domain. As a consequence, the market price of risk decreases as μ_b is increased (see the first columns of Tables 3(a) and 3(b)).

Consider next cases in which $\{m_{t+1,t}^f\}$ is much smoother than $\{m_{t+1,t}^r\}$, so that the market price of risk is approximately $\text{std}(m_{t+1,t}^r | J_t)$. The (conditional) standard deviation of $\{m_{t+1,t}^r\}$ will be large when the distortion in the conditional expectation operator is large. As μ_b increases, the representative consumer’s consumption is moved further away from his ideal point and hence the scope for pessimism is more pronounced. Thus increasing μ_b enhances the market price of risk.

More generally, the overall impact of increasing μ_b for a fixed σ is ambiguous except when $\sigma = 0$ and depends on the particular features of the calibrated economy. For the

27. It can be argued that risk sensitivity is simply repairing a defect in quadratic preferences, a criticism to which we certainly are vulnerable in this paper. The usual measure of relative risk aversion in the absence of habit persistence is $-cU''(c)/U'(c)$. In the case of our quadratic preferences, this is given by $c/(b - c)$, which requires that the bliss point process be twice the consumption level to attain a risk aversion coefficient of one. For an investigation of risk sensitive preferences and logarithmic utility, see Tallarini (1999).

calculations reported in Tables 3(a) and 3(b), the median market price of risk increases with μ_b except when σ is near zero.

Market price of risk and robustness

As we have just seen, risk sensitivity introduces an additional (multiplicative) factor $m'_{t+1,t}$ into the stochastic discount factor. This factor changes only slightly when risk sensitivity is reinterpreted as a preference for robustness. When interpreted as a preference for robustness, we can abstract from the covariance enhancement of the shocks. However, relative to those reported in Tables 3(a) and 3(b), the numbers for the market price of risk barely change when computed assuming Knightian uncertainty rather than risk-sensitive preferences.

Let $m''_{t+1,t}$ denote the resulting multiplicative factor, so that the composite stochastic discount factor is

$$m_{t+1,t} = m''_{t+1,t} m'_{t+1,t}.$$

To aid our understanding, suppose initially that $m'_{t+1,t}$ is constant, so the market price of risk is given by

$$\text{mpr}_t = \text{std}(m''_{t+1,t} | J_t).$$

The first columns of Tables 3(a) and 3(b) suggest that the conditional standard deviation of $m'_{t+1,t}$ is indeed close to zero for the preference specification used in our calculations.

Under our particular specification of uncertainty aversion, recall that asset prices are computed using the “pessimistic” view of tomorrow’s shock vector: w_{t+1} is normally distributed with conditional mean \hat{v}_t and covariance matrix I where \hat{v}_t is computed from the solution to the two-person game. It follows that

$$m''_{t+1,t} = \frac{\exp[-(w_{t+1} - \hat{v}_t)'(w_{t+1} - \hat{v}_t)/2]}{\exp(-w_{t+1}'w_{t+1}/2)},$$

which is the density ratio of the “distorted” relative to the “true” probability distribution. By a straightforward calculation, it follows that

$$E_t[(m''_{t+1,t})^2] = \exp(\hat{v}_t'\hat{v}_t),$$

and by construction

$$E_t(m''_{t+1,t}) = 1.$$

Therefore,

$$\text{std}(m''_{t+1,t} | J_t) = [\exp(\hat{v}_t'\hat{v}_t) - 1]^{1/2} \approx |\hat{v}_t|,$$

for small distortions. In other words, the market price of risk is approximately equal to the magnitude of the time t specification error. Our market prices of risk calculated under uncertainty aversion are only slightly smaller than those computed under risk sensitivity due to the small variance adjustment associated with the operator \mathcal{T} .

To understand better this approximate mapping from the permissible specification errors to the market price of risk, consider the following. Under the correct model specification, the shock vector is normally distributed and is normalized to have the identity as its covariance matrix. Suppose a misspecification takes the form of a conditional mean distortion of, say, 10% times a direction vector with Euclidean norm one. This direction

vector has the same dimension as the shock vector and picks the direction of the conditional mean distortion. This 10% distortion would alter a Gaussian log-likelihood function by

$$0.005 = \frac{\hat{v}_t \cdot \hat{v}_t}{2},$$

times the number of time periods in the sample. Thus a distortion of this magnitude would be hard to detect using a sample like ours, which consists of a little more than one hundred time periods. Having economic agents allow for distortions of this magnitude gives a market price of risk of approximately 0.10, assuming that there is *no* variation in the usually constructed stochastic discount factor. The fact that a mistake in forecasting w_{t+1} could lead to a direct enhancement of the market price of risk by the magnitude of the mistake is perhaps not surprising. What is conveyed here is how concern for robustness approximately directs the associated pessimism to returns that are conditionally mean-standard deviation efficient.

More generally, we expect that $|\hat{v}_t|$ is an upper bound on the approximate enhancement to the market price of risk caused by the concern for robustness. Given the ‘‘pessimistic’’ construction of v_t , we expect the two components $m''_{t+1,t}$ and $m^f_{t+1,t}$ of the stochastic discount factor to be positively correlated. This upper bound is closer to being attained when the two terms are highly positively correlated.

Measuring Knightian uncertainty

Let $\hat{v}_{d,t}$ and $\hat{v}_{d^*,t}$ be the two components of \hat{v}_t associated with the innovation to the two endowment shocks. Equation (10) makes these ‘‘worst case’’ specification errors linear functions of the current Markov state. We report measures of the sizes of the $\hat{v}_{d^*,t}$ and $\hat{v}_{d,t}$ processes in Tables 4(a) and 4(b). The tables report the medians in \hat{v}_t as well as minima and maxima over the sample. Like the market prices of risk, these measures are evaluated at the estimated values of the shock processes (\hat{d}_t, d_t^*) over the estimation period.

Recall from our previous discussion that the enhancement of the market price of risk caused by Knightian uncertainty is approximately $|\hat{v}_t|$. The tables show how $|\hat{v}_t|$ is mostly composed of specification errors in the shock for the persistent component of income d^* . Figure 7 displays time series estimates of d_t^* and $\hat{v}_{d^*,t}$ for $\mu_b = 30$, $\sigma = -0.0001$. Relative to the transition density induced by the undistorted model, the distorted mean is a random walk, as shown in the proof of the observational equivalence proposition.

7. INTERTEMPORAL MEAN-RISK TRADEOFFS

The market price of risk reported above conveys information about the one-period trade-off between the mean and standard deviation of asset returns as encoded in shadow prices. We now investigate the implied *intertemporal* tradeoff between means and standard deviations associated with our alternative configurations of μ_b and σ . Specifically, given a proportionate increase in the innovation standard deviation of an endowment shock, we aim to compute what proportionate increase in the conditional mean of that component of the endowment is required to keep the social planner on the same indifference curve. Initially we answer this question ‘‘locally’’ by considering small interventions. This imitates in part local measures of risk aversion. However, local measures of risk aversion are often computed around certainty lines. In our case, we localize around the solution to the

TABLE 4(a)

Median $\hat{v}_{d^,t}^*$ (with habit persistence)*

μ_b	$\sigma:$	0	-0.00005	-0.0001	-0.00015
18	0		-0.0129	-0.0259	-0.0388
	(0, 0)		(-0.0166, -0.0096)	(-0.0331, -0.0191)	(-0.0498, -0.0287)
24	0		-0.0349	-0.0698	-0.1048
	(0, 0)		(-0.0385, -0.0315)	(-0.0771, -0.0631)	(-0.1158, -0.0947)
30	0		-0.0569	-0.1138	-0.1708
	(0, 0)		(-0.0605, -0.0535)	(-0.1211, -0.1071)	(-0.1818, -0.1607)
36	0		-0.0788	-0.1578	-0.2368
	(0, 0)		(-0.0825, -0.0754)	(-0.1650, -0.1510)	(-0.2478, -0.2267)

TABLE 4(b)

Median $\hat{v}_{d,t}^$ (with habit persistence)*

μ_b	$\sigma:$	0	-0.00005	-0.0001	-0.00015
18	0		-0.0002	-0.0004	-0.0005
	(0, 0)		(-0.0002, -0.0001)	(-0.0005, -0.0003)	(-0.0007, -0.0004)
24	0		-0.0005	-0.0010	-0.0015
	(0, 0)		(-0.0005, -0.0004)	(-0.0011, -0.0009)	(-0.0016, -0.0013)
30	0		-0.0008	-0.0016	-0.0024
	(0, 0)		(-0.0009, -0.0008)	(-0.0017, -0.0015)	(-0.0026, -0.0023)
36	0		-0.0011	-0.0022	-0.0033
	(0, 0)		(-0.0012, -0.0011)	(-0.0023, -0.0021)	(-0.0035, -0.0032)

*Note: Minimum and maximum values are in parentheses below each median.

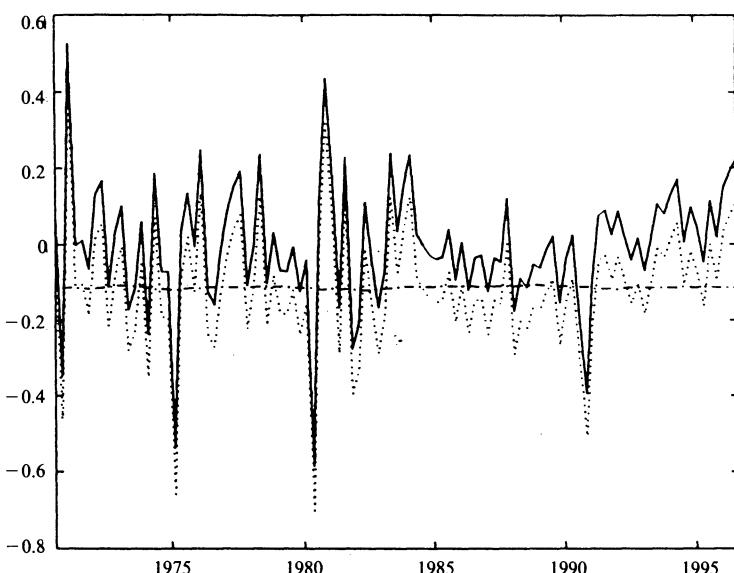


FIGURE 7

Estimated innovation to d_t^* (solid line), distorted mean $\hat{v}_{d^*,t}^*$ (dashed line), and sum of innovation and distorted mean (dotted line) for $\mu_b = 30$, $\sigma = -0.0001$

permanent income optimal resource allocation problem. Our localization permits us to depict risk-aversion as the ratio of two appropriately chosen intertemporal prices. Thus, like the market price of risk, our intertemporal measure of risk aversion also can be interpreted as a price ratio. We supplement this local experiment with a global one in which the standard deviation of the shock is set to zero. The intertemporal vantage point adopted in this section affects the character of the implied measures of risk aversion. The calculations will be conducted using the “risk-sensitive” decentralization. A corresponding “robust” decentralization gives rise to essentially the same numbers.

Local measure of risk aversion

We form a local intertemporal tradeoff between the standard deviation and the mean of the endowment about the equilibrium process for consumption and investment. Specifically, given a proportional enhancement of standard deviation of the endowment innovation in all future time periods, we aim to compute what proportional mean increase in the endowment is required to keep the social planner on the same indifference curve, at least locally. To perform this computation we attain two “value expansions,” both of which we describe below. The first-order terms or “derivatives” in these expansions can be interpreted as prices of appropriately chosen infinitely lived securities.

We implement a “local” modification in the state evolution equation by adopting the parameterization of the law of motion starting for $j \geq 0$ as

$$x_{t+j+1}^\varepsilon = A_0 x_{t+j}^\varepsilon + (C + \varepsilon G) w_{t+1+j},$$

where ε is a small positive scalar. A positive ε initiates a change in the innovation standard deviation starting with date $t+1$. Here the matrix G is designed to select one of the endowment innovations. For example, it can be identical to C except with zeros for entries associated with the other endowment shock. Let $U_t = W^\varepsilon(x_t)$ denote the value function for the resulting control problem indexed by ε ; we take W^0 as the value function for a baseline control problem (say the risk sensitive permanent income model). Let

$$x_{t+1} = A_0 x_t + C w_{t+1},$$

be the corresponding $\varepsilon = 0$ state evolution equation when the optimal control law is imposed. We aim to compute an expansion of the form

$$W^\varepsilon(x) = W^0(x) + \varepsilon W_d(x) + o(\varepsilon),$$

where $o(\varepsilon)$ converges to zero as ε tends to zero uniformly on compact subsets of the state space. We will derive an asset pricing formulation of W_d that, among other uses, facilitates calculations.

A corresponding experiment delivers a “robust control” expansion. Alter the intervention that takes place at time t by introducing “mistakes” in the conditional mean. Now suppose instead that starting for $j \geq 0$ we have

$$x_{t+j+1}^\varepsilon = A_0 x_{t+j} + (C + \varepsilon G)(w_{t+1+j} + v_{t+j}).$$

As before the parameter σ is used to restrain mistakes, rather than to make a risk adjustment in the utility recursion. This perturbed system gives rise to an expansion that, from a quantitative vantage point, is virtually identical to that we report. The subsequent asset pricing interpretation also applies, provided that we use the prices for the “robust” decentralization in place of the prices of the “risk sensitive” decentralization.

Of course, W^ε is a translated quadratic function of the state vector. We write this function as

$$W^\varepsilon(x) = x' \Omega^\varepsilon x + \rho^\varepsilon.$$

The function W_d is quadratic

$$W_d(x) = x' \Omega_d x + \rho_d.$$

In effect, Ω_d is the derivative with respect to ε of the matrix function Ω^ε , evaluated at $\varepsilon = 0$. Similarly, ρ_d is the derivative with respect to ε of the scalar function ρ^ε . Computations of these derivatives are simplified by the fact that we can abstract from the role of optimization of the control vector for small changes in ε . This familiar property follows from the first-order conditions satisfied by the optimal control law, which imply that the contribution to the value function expansion is *second order* in ε . Hence we can compute the derivatives as if we are holding fixed the control law and hence the state evolution matrix A_0 . The matrix Ω_d can be computed easily as the solution of a Sylvester equation.

Measuring risk aversion by asset pricing

Holding fixed the equilibrium law of motion for consumption, $c_t^0 = S_c x_t$, we can use our asset pricing formula to evaluate how utility responds to changes in ε . To compute the desired “derivative” of U_t with respect to ε , we begin by forming a new state vector process

$$x_{t+j}^\varepsilon - x_{t+j}^0 = \varepsilon y_{t+j}, \quad (46)$$

where $\{y_t\}$ evolves according to

$$y_{t+j+1} = A_0 y_{t+j} + G w_{t+1+j},$$

with $y_t = 0$. Notice the linear scaling in ε . A consumption process associated with $\varepsilon > 0$ is

$$c_{t+j}^\varepsilon = c_{t+j}^0 + \varepsilon S_c y_{t+j}.$$

It follows from our subgradient inequality (40) that

$$\frac{W^\varepsilon(x_t) - W^0(x_t)}{\varepsilon} \leq \sum_{j=1}^{\infty} \beta^j \mathcal{S}_{t,j}(\mathcal{M}_{t+j}^c S_c y_{t+j}).$$

It can be verified that as ε declines to zero, this becomes an equality. Therefore, we can evaluate the desired “derivative” by using the following asset pricing formula

$$W_d(x) = \sum_{j=1}^{\infty} \beta^j \mathcal{S}_{t,j}(\mathcal{M}_{t+j}^c S_c y_{t+j}).$$

This is the time t price, scaled in units of marginal utility, of an infinitely-lived security with dividend $\{S_c y_{t+j}\}$.²⁸

To compute the local mean-risk tradeoff, we also estimate the utility change associated with a small change in the conditional mean of the endowment. We capture this

28. To perform the computation, first form the state transition equation for the composite state (x_t^0, y_t') . The transition equation has a block diagonal state matrix with diagonal blocks A_0 . The counterpart to C is constructed by stacking C on top of G . Consumption will be formed by using a matrix $(S_c \ 0)$ and the dividend will be formed by $(0 \ S_c)$. Prices can now be computed recursively using a doubling algorithm.

small change as follows

$$\begin{aligned}x_{t+1}^\delta &= A_0 x_t + \delta D x_t + C w_{t+1}, \\x_{t+j+1}^\delta &= A_0 x_{t+j}^\delta + C w_{t+1+j},\end{aligned}$$

for $j = 1, 2, \dots$. This envisions the change in the conditional expectation as occurring at date $t+1$ continuing into the future and leads us to the time t value-function expansion

$$\hat{W}^\delta(x) = W^0(x) + \delta \hat{W}_d(x) + o(\delta).$$

Here \hat{W}_d is a quadratic function of the state vector, which we represent as $x' \hat{\Omega}_d x$.

Imitating our earlier derivation, we form

$$x_{t+j}^\delta - x_{t+j} = \delta A_0^{j-1} D x_t.$$

Notice the linear scaling in δ . The new consumption process can be expressed as

$$c_{t+j}^\delta = c_{t+j}^0 + \delta S_c A_0^{j-1} D x_t.$$

From our subgradient inequality (39),

$$\frac{W^\delta(x_t) - W^0(x_t)}{\delta} \leq \sum_{j=1}^{\infty} \beta^j \mathcal{S}_{t,j} (\mathcal{M}_{t+j}^c S_c A_0^{j-1} D x_t).$$

Again we can show that this subgradient is actually a gradient by driving δ to zero. Therefore, our target derivative is given by

$$\hat{W}_d(x_t) = \sum_{j=1}^{\infty} \beta^j \mathcal{S}_{t,j} (\mathcal{M}_{t+j}^c S_c A_0^{j-1} D x_t),$$

which is the (time t util) price of an infinitely-lived security with dividend $\{S_c A_0^{j-1} D x_t\}$. Thus, $\{\hat{\Omega}_d\}$ solves a Sylvester equation.

Using our two expansions, the compensation measure is

$$\delta_t = -\frac{x_t' \hat{\Omega}_d x_t + \rho_d}{x_t' \hat{\Omega}_d x_t} = -\frac{W_d(x_t)}{\hat{W}_d(x_t)},$$

which we index by t to accommodate the change in vantage point as time passes.

TABLE 5
Local mean-risk trade-off

μ_b	σ :	0	-0.00005	-0.0001	-0.00015
18		0.3182	0.4347	0.7358	1.9230
24		0.1179	0.3754	0.9482	3.1017
30		0.0723	0.4828	1.3775	4.7178
36		0.0522	0.6175	1.8423	6.4053

In Table 5, we report our (local) intertemporal measures of risk aversion. The effect of increasing (in absolute value) σ has a stronger effect on the mean-risk trade-off than on the market price of risk (compare Table 5 to Table 3(a)). Increases in μ_b also have a slightly greater impact for the trade-off calculation.²⁹

We next verify the local nature of these computations by considering the following experiment. Let $\varepsilon = -1$, which sets to zero the shock variance for the endowment process.

29. Increasing the market price of risk by enlarging μ_b has the virtue of further reducing the probability of satiation. This would appear to increase the intertemporal substitutability of consumption. However, recall that μ_b does not appear in the permanent income decision rule. Thus, by design we have not changed the consumption-savings behaviour of the consumer as we change μ_b . On the other hand, some perverse implications "off the equilibrium path" can occur for large values of μ_b .

By extrapolating the local measures reported in Table 5, the entries in this table should convey what fraction of the endowment the consumer would be willing to forego to achieve this reduction in volatility. Such an inquiry relates to Lucas' (1987) quantification of the welfare costs to fluctuations, except that we are using a permanent income model that permits investment (see also Obstfeld (1994) and Tallarini (1999)). From this vantage point, the numbers in Table 5 look to be enormous, particularly for the larger (in absolute value) specifications of σ . However, that extrapolation of our local measure turns out to be misleading. To see this, in Table 6 we report global numbers for the $\varepsilon = -1$ experiment that holds fixed the permanent income decision rule for the two competing specifications of the endowment process. The global mean-risk tradeoffs are much smaller by a factor ranging from two to four. Nevertheless, the tradeoffs remain quite large, except when σ is close to zero.³⁰

TABLE 6
Global mean-risk trade-off

μ_b	$\sigma:$	0	-0.00005	-0.0001	-0.00015
18		0.1267	0.1635	0.2460	0.4863
24		0.0564	0.1666	0.3709	0.9206
30		0.0355	0.2189	0.5495	1.4281
36		0.0258	0.2818	0.7393	1.9503

8. CONCLUSIONS

Robert E. Lucas, Jr. (1975) warned us about theorists bearing free parameters. Having heard Lucas, we devoted this paper to scrutinizing some of the implications for prices and quantities of a single additional parameter designed to quantify a preference for robustness to specification errors. By exploiting the connection between robustness and the risk-sensitivity specification of Jacobson (1973) and Whittle (1990), we have shown how to decentralize dynamic, stochastic general equilibrium models with a consumer who fears model misspecification. Formulas for consumption, investment, and the risk-free interest rate are identical to ones coming from the usual permanent income models. We presented formulas for the market price of risk, then applied them to account for the market price of risk observed in the U.S. data.

Like Brock and LeBaron (1996), Brock and Hommes (1994), Cochrane (1989), Marcket and Sargent (1989), and Krusell and Smith (1996), we can regard the consumer-investors in our economy as making "mistakes", but as managing them differently than do those in these authors' models.³¹ Our agents are very sophisticated in how they accommodate possible mistakes: they base decisions on worse-case scenarios, following Gilboa and Schmeidler (1989) and Epstein and Wang (1994).

30. The global numbers would be enhanced a little if we reoptimize when setting the endowment shock to zero. The solution to a linear-quadratic problem is unappealing in this context because with less uncertainty, capital ceases to be an attractive way to transform goods from one period to the next. In light of this, it seems crucial to reoptimize subject to a nonnegativity constraint on capital. Our imposition of the suboptimal "permanent income" consumption rule diminishes the impact of this nonnegativity constraint while possibly misstating the global tradeoff.

31. Cochrane's (1989) and Krusell and Smith's (1996) agents use decision rules that are perturbed by small amounts in arbitrary directions from optimal ones. Marcket and Sargent's (1989) agents correctly solve dynamic programming problems, but subject to subtly misspecified constraints: they use estimated transition laws (usually of the correct functional forms) which they mistakenly take as non-random and time-invariant. See Brock and LeBaron (1996), especially their footnote 2, for a lucid explanation of a class of models that mix "adaptation"—to induce local instability near rational expectations equilibria—with enough "rationality" to promote global attraction toward the vicinity of rational expectations. Brock and LeBaron (1996) and Brock and Hommes (1994) balance the tension between adaptation and rationality to mimic some interesting return and volume dynamics.

In contrast to Cochrane (1989) and Krusell and Smith (1996), for our permanent income economy, the quantity allocations are observationally equivalent to those in an economy in which no “mistakes” are contemplated. This situation stems partly from the econometrician’s ignorance of the subjective discount factor. Like Epstein and Wang (1994) and Melino and Epstein (1995), we focus on how aversion to mistakes transmits into security market prices. We find that a conditional mean “mistake” of $x\%$ of a unit norm vector for a multivariate standard normal shock process increases the market price of risk by approximately $x/100$.

We have concentrated on a robust interpretation of the permanent income model of consumption. The permanent income model seemed a natural starting point for exploring the consequences of robust decision theory, partly because of its simplicity. Recent work by Carroll (1992) has emphasized a departure from the permanent income model induced by precautionary savings, low discount factors, and big utility costs to zero consumption.³² As we have emphasized, our reinterpretation of the permanent income model also relies on smaller discount factors and precautionary savings. It does not, however, permit us to explore the ramifications of big utility costs to zero consumption, which is central to the work of Carroll (1992) and others, and which requires nonquadratic objective functions. However, Anderson, Hansen and Sargent (1998) have shown how the connection between risk sensitivity and robustness extends to discounted control problems with nonquadratic criteria and nonlinear, stochastic evolution equations. They formulate a recursive nonlinear robust control problem that applies readily to consumption and savings decisions.

Maybe we take the representative agent paradigm too seriously. We use the representative agent as a convenient starting point to understand the workings of risk sensitivity and robustness in decentralized economies. In other settings, we know how heterogeneity of preferences and incomplete risk sharing affect investment behaviour and the market price of risk. In our model (and Epstein and Wang’s (1994)), agents agree on the amount and location of the Knightian uncertainty. Thus, models like ours can contribute an additional dimension upon which heterogeneity alters equilibrium quantities and prices.

APPENDIX

A: Subgradient inequality

This derives the subgradient inequality used for equilibrium pricing. Let U^e denote the original nonpositive random utility index, U any other nonpositive random utility index and J a sigma algebra of events. We will show that

$$\mathcal{R}(U) - \mathcal{R}(U^e) \leq E[V^e(U - U^e)|J]/E(V^e|J), \quad (47)$$

where

$$\begin{aligned} \mathcal{R}(U) &\equiv (2/\sigma) \log \{E[\exp(\sigma U/2)|J], \\ V^e &\equiv \exp(\sigma U^e/2). \end{aligned} \quad (48)$$

We assume that $E[\exp(\sigma U^e/2)|J]$ and hence $\mathcal{R}(U^e)$ is finite with probability one. Define $h \equiv U - U^e$, and let δ be any real number in $(0, 1)$. Interpret δ as determining the magnitude of a perturbation in direction h . In other words, the perturbation away from U^e under consideration is δh .

By the convexity of the exponential function

$$\exp[\sigma(U^e + h\delta)/2] - \exp(\sigma U^e/2) \geq \delta h(\sigma/2)V^e. \quad (49)$$

32. See Leland (1968) and Miller (1974) for important early contributions to the literature on precautionary saving.

This inequality remains true when computing expectations conditioned on J , although either side may be infinite

$$E\{\exp[\sigma(U^e + h\delta)/2]|J\} - E\{\exp(\sigma U^e/2)|J\} \geq \delta(\sigma/2)E(V^e h|J). \quad (50)$$

Divide each side of (50) by $E(V^e|J)$

$$E\{\exp[\sigma(U^e + h\delta)/2]|J\}/E\{\exp(\sigma U^e/2)|J\} - 1 \geq \delta(\sigma/2)E(V^e h|J)/E(V^e|J). \quad (51)$$

Since $0 < \delta < 1$, $(U^e + h\delta)$ is a convex combination of U^e and U with weights $(1 - \delta)$ and δ respectively. By the conditional version of the Hölder Inequality,

$$\begin{aligned} E\{\exp[\sigma(U^e + h\delta)/2]|J\} &= E\{\exp(\sigma U^e/2)^{1-\delta} [\exp(\sigma U/2)]^\delta | J\} \\ &\leq \{E[\exp(\sigma U^e/2)|J]\}^{1-\delta} \{E[\exp(\sigma U/2)|J]\}^\delta. \end{aligned} \quad (52)$$

Combining (51) and (52) and dividing by δ , we have that

$$(1/\delta)\{E[\exp(\sigma U/2)|J]/E(V^e|J)\}^\delta - 1 \geq (\sigma/2)E(V^e h|J)/E(V^e|J). \quad (53)$$

To complete the derivation, we use the familiar approximation result for logarithms

$$\lim_{\delta \rightarrow 0} (\lambda^\delta - 1)/\delta = \log(\lambda), \quad (54)$$

where the limit is from above. (This limit can be verified by applying L'Hôpital's Rule or by using the series expansion for $\exp[\delta \log(\lambda)]$). Taking limits of the left side of (53) as δ declines to zero yields

$$\log\{E[\exp(\sigma U/2)|J]\} - \log\{E(V^e|J)\} \geq (\sigma/2)E(V^e h|J)/E(V^e|J). \quad (55)$$

The desired inequality (47) is obtained by multiplying both sides of (55) by the negative number $(2/\sigma)$ and reversing the inequality.

B: Computing prices for state-contingent utility

In this Appendix we provide a characterization of the operator $\bar{\mathcal{F}}$ used in pricing state-contingent utility. The characterization relies on a restriction that the utility index U_{t+1}^e . For notational convenience, we will suppress superscripts and subscripts.

Suppose that a utility index is quadratic in a normally distributed random vector $x \in \mathbf{R}^n$

$$U = x' \Omega x + \rho, \quad (56)$$

where Ω is a negative semidefinite matrix and $\rho \leq 0$. In addition, suppose that

$$x = \mu + Cw, \quad (57)$$

where w is normally distributed random vector with mean zero and covariance matrix I . Recall that $\bar{\mathcal{F}}$ can be interpreted as a conditional expectation with a change of probability measure. In terms of the notation just developed, the new probability measure is constructed using V/EV as a Radon–Nikodym derivative where

$$V = \exp(\sigma U/2) \propto \exp(\sigma w' C' \Omega C w / 2 + \sigma w' C' \Omega \mu). \quad (58)$$

We can compute expectations with respect to the transformed measure as follows. Let ϕ be any bounded, Borel measurable function mapping $\mathbf{R}^n \rightarrow \mathbf{R}$. Then

$$E[V\phi]/EV \propto \int \phi(w) \exp(\sigma w' C' \Omega C w / 2 + \sigma w' C' \Omega \mu) \exp(-w' w / 2) dw. \quad (59)$$

Note that

$$\sigma w' C' \Omega C w / 2 + \sigma w' C' \Omega \mu - w' w / 2 = -w'(I - \sigma C' \Omega C)w / 2 + w'(I - \sigma C' \Omega C)(I - \sigma C' \Omega C)^{-1} \sigma C' \Omega \mu. \quad (60)$$

Consequently, the operator on the left side of (59) can be evaluated by integrating ϕ with respect to a normal density with mean vector

$$\tilde{\mu} \equiv (I - \sigma C' \Omega C)^{-1} \sigma C' \Omega \mu, \quad (61)$$

and covariance matrix

$$\tilde{\Sigma} \equiv (I - \sigma C' \Omega C)^{-1}. \quad (62)$$

The corresponding mean vector and covariance matrix for x are $\mu + C\tilde{\mu}$ and $C\tilde{\Sigma}C'$, respectively. The $\bar{\mathcal{F}}$ operator will only be well defined so long as $\sigma C' \Omega C < I$.

C: Computing the conditional variance of the stochastic discount factor

From equation (45), we know that $m_{t+1,t}$, the intertemporal marginal rate of substitution between time t and time $t+1$ can be written as

$$m_{t+1,t} = \frac{\beta \{ \exp(\sigma U_{t+1}^e / 2) v' \mathcal{M}_{t+1}^c \}}{E \{ \exp(\sigma U_{t+1}^e / 2) | J_t \} v' \mathcal{M}_t^c}, \quad (63)$$

or as

$$m_{t+1,t} = \frac{\beta \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho) / 2] v' M_c x_{t+1} \}}{\exp [\sigma(x'_t \hat{\Omega} x_t + \tilde{\rho}) / 2] v' M_c x_t}, \quad (64)$$

where $\hat{\Omega}$ and $\tilde{\rho}$ are given by (6). By applying the results of Appendix B we can compute the mean of $m_{t+1,t}$, conditional on information available at time t . The result can be written as

$$E(m_{t+1,t} | J_t) = \beta (v' M_c \hat{A} x_t) / (v' M_c x_t). \quad (65)$$

Our present goal is to compute the conditional second moment of $m_{t+1,t}$ as a means for computing its conditional variance. We will accomplish this by manipulating $m_{t+1,t}^2$ so that we can transform the probability measure as in Appendix B with a different function V . We have

$$m_{t+1,t}^2 = \frac{\beta^2}{(v' M_c x_t)^2} \frac{\exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)]}{\exp [\sigma(x'_t \hat{\Omega} x_t + \tilde{\rho})]} (v' M_c x_{t+1})^2. \quad (66)$$

Multiply the numerator and denominator by the time t conditional mean of the exponential term in the numerator, $E \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] | J_t \}$. This gives us

$$m_{t+1,t}^2 = \frac{\beta^2 E \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] | J_t \}}{(v' M_c x_t)^2} \frac{\exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] (v' M_c x_{t+1})^2}{E \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] | J_t \}}. \quad (67)$$

This conditional expectation can be computed by using a formula found in Jacobson (1973), only substituting 2σ for σ

$$\begin{aligned} E \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] | J_t \} &= [\det(I - 2\sigma C' \Omega C)]^{-1/2} \exp [\sigma(x'_t \tilde{\Omega} x_t + \tilde{\rho})] \\ &= \exp [\sigma(x'_t \tilde{\Omega} x_t + \tilde{\rho})], \end{aligned} \quad (68)$$

where $\tilde{\Omega} \equiv A'[\Omega + 2\sigma \Omega C(I - 2\sigma C' \Omega C)^{-1} C' \Omega]A$ and $\tilde{\rho} \equiv -(1/2\sigma) \log \det(I - 2\sigma C' \Omega C) + \rho$. So we get that

$$\frac{E \{ \exp [\sigma(x'_{t+1} \Omega x_{t+1} + \rho)] | J_t \}}{\exp [\sigma(x'_t \hat{\Omega} x_t + \tilde{\rho})]} = \exp \{ \sigma[x'_t (\tilde{\Omega} - \hat{\Omega}) x_t + \tilde{\rho} - \tilde{\rho}] \}. \quad (69)$$

This gives us

$$m_{t+1,t}^2 = \frac{\beta^2}{(v' M_c x_t)^2} \exp \{ \sigma[x'_t (\tilde{\Omega} - \hat{\Omega}) x_t + \tilde{\rho} - \tilde{\rho}] \} \frac{\tilde{V}_{t+1} (v' M_c x_{t+1})^2}{E \{ \tilde{V}_{t+1} | J_t \}}, \quad (70)$$

where $\tilde{V}_{t+1} = \exp [(\sigma(x'_{t+1} \Omega x_{t+1} + \rho))]$. So

$$E(m_{t+1,t}^2 | J_t) = \frac{\beta^2}{(v' M_c x_t)^2} \exp \{ \sigma[x'_t (\tilde{\Omega} - \hat{\Omega}) x_t + \tilde{\rho} - \tilde{\rho}] \} \tilde{\mathcal{T}}_t [(v' M_c x_{t+1})^2], \quad (71)$$

where $\tilde{\mathcal{T}}_t$ is the transformed conditional expectation operator for a 2σ economy. We can evaluate the $\tilde{\mathcal{T}}_t$ term in the above expression using results from Appendix B

$$\tilde{\mathcal{T}}_t [(v' M_c x_{t+1})^2] = x'_t \tilde{A}' M'_c v v' M_c \tilde{A} x_t + \text{trace}(v' M_c \tilde{C} \tilde{C}' M'_c v), \quad (72)$$

where

$$\tilde{A} \equiv [I + 2\sigma C(I - 2\sigma C' \Omega C)^{-1} C' \Omega]A, \quad (73)$$

and

$$\tilde{C} \tilde{C}' \equiv C(I - 2\sigma C' \Omega C)^{-1} C'. \quad (74)$$

Finally, we know that the conditional variance of $m_{t+1,t}$ is given by its conditional second moment minus the square of its conditional mean.

Acknowledgements. This research was funded by grants from the National Science Foundation. We thank Andrew Abel, Evan Anderson, John Cochrane, Cristina de Nardi, Martin Eichenbaum, John Heaton, Narayana Kocherlakota, Wen-Fang Liu, Jesus Santos, Kenneth Singleton, Nancy Stokey and Noah Williams for useful criticisms of earlier drafts. We are grateful to Wen-Fang Liu for excellent research assistance. We thank two referees of an earlier draft for comments that prompted an extensive reorientation of our research.

REFERENCES

- ANDERSON, E., HANSEN, L. P. and SARGENT, T. J. (1998), "Risk and Robustness in Equilibrium" (Mimeo, University of Chicago).
- BECKER, G. S. and MURPHY, K. M. (1988), "A Theory of Rational Addiction", *Journal of Political Economy*, **96**, 675–700.
- BEWLEY, T. (1977), "The Permanent Income Hypothesis: A Theoretical Formulation", *Journal of Economic Theory*, **16**, 252–259.
- BOLDRIN, M., CHRISTIANO, L. J. and FISHER, J. D. M. (1995), "Asset Pricing Lessons for Modeling Business Cycles" (NBER Working Paper No. 5262).
- BROCK, W. A. and HOMMES, C. (1994), "Rational Routes to Randomness: Rationality in an Unstable Market with Information Costs Implies Chaos" (Mimeo, Department of Economics, University of Wisconsin at Madison).
- BROCK, W. A. and LEBARON, B. D. (1966), "A Dynamic Structural Model for Stock Return Volatility and Trading Volume", *Review of Economics and Statistics*, **78**, 94–110.
- BURNSIDE, C. (1994), "Hansen–Jagannathan Bounds as Classical Tests of Asset Pricing Models", *Journal of Business and Economic Statistics*, **12**, 57–79.
- BURNSIDE, C., EICHENBAUM, M. and REBELO, S. (1990), "Labor Hoarding and the Real Business Cycle" (NBER Working Paper No. 3556).
- CAMPBELL, J. Y. (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis", *Econometrica*, **55**, 1249–1273.
- CAMPI, M. C. and JAMES, M. R. (1995), "Nonlinear Discrete-Time Risk-Sensitive Control", *International Journal of Robust Nonlinear Control* (forthcoming).
- CARROLL, C. D. (1992), "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence", *Brookings Papers on Economic Activity*, **2**, 61–156.
- CECCHETTI, S. G., LAM, P. and MARK, N. C. (1994), "Testing Volatility Restrictions on Intertemporal Marginal Rates of Substitution Implied by Euler Equations and Asset Returns", *Journal of Finance*, **49**, 123–152.
- CHRISTIANO, L. J. (1987), "Cagan's Model of Hyperinflation under Rational Expectations", *International Economic Review*, **28**, 33–49.
- CHRISTIANO, L. J., EICHENBAUM, M. and MARSHALL, D. (1991), "The Permanent Income Model Revisited", *Econometrica*, **59**, 397–423.
- COCHRANE, J. H. (1989), "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives", *American Economic Review*, **79**, 319–337.
- COCHRANE, J. and HANSEN, L. P. (1992), "Asset Pricing Explorations for Macroeconomics", *NBER Macroeconomics Annual* (Cambridge, MA: MIT Press), 115–165.
- CONSTANTINIDES, G. M. (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, **98**, 519–543.
- DEATON, A. (1991), "Saving and Liquidity Constraints", *Econometrica*, **59**, 1221–1248.
- DESCARTES, R. (1901) *The Method, Meditations, and Philosophy of Descartes*. Translated by John Vietch (Washington and London: M. Walter Dunner).
- EPSTEIN, L. G. (1988), "Risk Aversion and Asset Prices", *Journal of Monetary Economics*, **22**, 179–192.
- EPSTEIN, L. G. and WANG, T. (1994), "Intertemporal Asset Pricing Under Knightian Uncertainty", *Econometrica*, **62**, 283–322.
- EPSTEIN, L. G. and ZIN, S. E. (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, **57**, 937–969.
- EPSTEIN, L. G. and ZIN, S. E. (1990), "First-Order Risk Aversion and the Equity Premium Puzzle", *Journal of Monetary Economics*, **26**, 387–407.
- FRANCIS, B. A. (1987) *A Course in H-infinity Control Theory*. Lecture Notes in Control and Information Sciences, **88**.
- GALLANT, R., HANSEN, L. P. and TAUCHEN, G. (1990), "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution", *Journal of Econometrics*, **45**, 145–179.
- GILBOA, I. and SCHMEIDLER, D. (1989), "Maxmin Expected Utility with Non-Unique Prior", *Journal of Mathematical Economics*, **18**, 141–153.
- GLOVER, K. and DOYLE, J. C. (1988), "State-Space Formulae for All Stabilizing Controllers that Satisfy an H_∞ -norm Bound and Relations to Risk Sensitivity", *Systems Control Letters*, **11**, 167–172.
- HALL, R. E. (1978), "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence", *Journal of Political Economy*, **91**, 249–265.

- HANSEN, L. P. (1987), "Calculating Asset Prices in Three Example Economies", in T. F. Bewley (ed.), *Advances in Econometrics, Fifth World Congress, Vol. 1* (New York: Cambridge University Press).
- HANSEN, L. P., HEATON, J., LUTTNER, E. G. J. (1995), "Econometric Evaluation of Asset Pricing Models", *Review of Financial Studies*, **8**, 237–274.
- HANSEN, L. P. and JAGANNATHAN, R. (1991), "Implications of Security Market Data for Models of Dynamic Economics", *Journal of Political Economy*, **99**, 225–262.
- HANSEN, L. P., ROBERDS, W. and SARGENT, T. J. (1991), "Time Series Implications of Present-Value Budget Balance and of Martingale Models of Consumption and Taxes", *Rational Expectations Econometrics* (Boulder: Westview Press).
- HANSEN, L. P. and SARGENT, T. J. (1983), "Aggregation over Time and the Inverse Optimal Predictor Problem for Adaptive Expectations in Continuous Time", *International Economic Review*, **24**, 1–20.
- HANSEN, L. P. and SARGENT, T. J. (1993), "Seasonality and Approximation Errors in Rational Expectations Models", *Journal of Econometrics*, **55**, 21–55.
- HANSEN, L. P. and SARGENT, T. J. (1995), "Discounted Linear Exponential Quadratic Gaussian Control", *IEEE Transactions on Automatic Control*, **40**, 968–971.
- HANSEN, L. P. and SARGENT, T. J. (1996) *Recursive Linear Models of Dynamic Economies* (forthcoming).
- HANSEN, L. P. and SARGENT, T. J. (1998), "Discounted Robust Filtering and Control in the Frequency Domain" (Mimeo, University of Chicago and Stanford University).
- HARRISON, J. M. and KREPS, D. M. (1979), "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory*, **20**, 381–408.
- HEATON, J. (1993), "The Interaction Between Time-Nonseparable Preferences and Time Aggregation", *Econometrica*, **61**, 353–385.
- HEATON, J. (1995), "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications", *Econometrica*, **63**, 681–718.
- JACOBSON, D. H. (1973), "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games", *IEEE Transactions on Automatic Control*, **18**, 124–131.
- JACOBSON, D. H. (1977) *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory* (New York: Academic Press).
- JAMES, M. R. (1992), "Asymptotic Analysis of Nonlinear Stochastic Risk-Sensitive Control and Differential Games", *Control, Signals, and Systems*, **5**, 401–417.
- JAMES, M. R. (1995), "Recent Developments in Nonlinear H_∞ Control" (Mimeo, Department of Engineering, Australian National University).
- KANDEL, S. and STAMBAUGH, R. F. (1991), "Asset Returns and Intertemporal Preferences", *Journal of Monetary Economics*, **27**, 40–71.
- KREPS, D. M. and PORTEUS, E. L. (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, **46**, 185–200.
- KRUSELL, P. and SMITH, T. (1996), "Rules of Thumb in Macroeconomic Equilibrium: A Quantitative Analysis", *Journal of Economic Dynamics and Control*, **20**, 527–558.
- LELAND, H. (1968), "Savings and Uncertainty: the Precautionary Demand for Savings", *Quarterly Journal of Economics*, **82**, 465–473.
- LJUNG, L. (1978), "Convergence Analysis of Parametric Identification Methods", *IEEE Transactions on Automatic Control*, **23**, 770–783.
- LUCAS, R. E., Jr. (1975), "An Equilibrium Model of the Business Cycle", *Journal of Political Economy*, **83**, 1113–1144.
- LUCAS, R. E., Jr. (1978), "Asset Prices in an Exchange Economy", *Econometrica*, **46**, 1429–1445.
- LUCAS, R. E., Jr. (1987) *Models of Business Cycles* (Oxford: Basil Blackwell).
- MARCET, A. and SARGENT, T. J. (1989), "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, **48**, 337–368.
- MEHRA, R. and PRESCOTT, E. C. (1985), "The Equity Premium Puzzle: A Solution?", *Journal of Monetary Economics*, **22**, 133–136.
- MELINO, A. and EPSTEIN, L. G. (1995), "An Empirical Analysis of Asset Returns Under Non-Bayesian Rational Expectations" (Mimeo, University of Toronto, September).
- MILLER, B. L. (1974), "Optimal Consumption with a Stochastic Income Stream", *Econometrica*, **42**, 253–266.
- MUTH, J. F. (1960), "Optimal Properties of Exponentially Weighted Forecasts", *Journal of the American Statistics Association*, **55**, 299–306.
- OBSTFELD, M. (1994), "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability", *European Economic Review*, **38**, 1471–1486.
- PRESCOTT, E. C. (1986), "Theory Ahead of Business Cycle Measurement", in K. Brunner and A. H. Meltzer (eds.), *Real Business Cycles, Real Exchange Rates and Actual Policies*, Carnegie-Rochester Conference Series on Public Policy, **25**, 11–14.
- RYDER, H. E., Jr. and HEAL, G. M. (1973), "Optimal Growth with Intertemporally Dependent Preferences", *Review of Economic Studies*, **40**, 1–31.
- SUNDARESAN, M. (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth", *Review of Financial Studies*, **2**, 73–89.

- TALLARINI, T. D., Jr. (1999), "Risk Sensitive Real Business Cycles", *Journal of Monetary Economics* (forthcoming).
- VAN DER PLOEG, F. (1993), "A Closed-Form Solution for a Model of Precautionary Saving", *Review of Economic Studies*, **60**, 385–395.
- WEIL, P. (1989), "The Equity Premium Puzzle and the Riskfree Rate Puzzle", *Journal of Monetary Economics*, **24**, 401–421.
- WEIL, P. (1990), "Nonexpected Utility in Macroeconomics", *Quarterly Journal of Economics*, **105**, 29–42.
- WEIL, P. (1993), "Precautionary Savings and the Permanent Income Hypothesis", *Review of Economic Studies*, **60**, 367–383.
- WHITE, H. (1982), "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, **50**, 1–26.
- WHITTLE, P. (1982) *Optimization Over Time, Vol. 1* (Chichester: John Wiley & Sons).
- WHITTLE, P. (1983) *Optimization Over Time, Vol. 2* (Chichester: John Wiley & Sons).
- WHITTLE, P. (1989), "Risk-Sensitive Linear Quadratic Gaussian Control", *Advances in Applied Probability*, **13**, 764–777.
- WHITTLE, P. (1990) *Risk-Sensitive Optimal Control* (New York: John Wiley & Sons).
- ZAMES, G. (1981), "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses", *IEEE Transactions on Automatic Control*, **26**, 301–320.
- ZELDES, S. P. (1989), "Optimal Consumption with Stochastic Income: Deviation from Certainty Equivalence", *Quarterly Journal of Economics*, **104**, 275–298.
- ZHOU, K., DOYLE, J. C. and GLOVER, K. (1996) *Robust and Optimal Control* (Upper Saddle River, Prentice-Hall).