Robust Portfolio Rules and Asset Pricing

Pascal J. Maenhout[1]

November 1999

First draft: March 1999

Comments Appreciated

Abstract: Parameter uncertainty or, more broadly, model uncertainty seems highly relevant in many aspects of financial decision-making. I explore the effects of such uncertainty on dynamic portfolio and consumption decisions, and on equilibrium asset prices. In particular, I use the framework of Anderson, Hansen and Sargent (1999), which attributes a preference for robustness to the decision-maker. Worried that the model she uses is misspecified, a robust agent seeks decision rules that insure against some worst-case misspecification, in accordance with maxmin expected utility. I demonstrate that robustness dramatically decreases the portfolio demand for equities. When modifying the framework of Anderson, Hansen and Sargent to impose homotheticity, I find robustness to be observationally equivalent to recursive preferences: robustness increases risk aversion, without affecting the willingness to substitute intertemporally. When investment opportunity sets are time-varying, robustness leads to an additional hedging-type asset demand, even for logarithmic utility. In an equilibrium exchange economy, robustness increases the equilibrium equity premium. The endogenous worst-case scenario for equity returns supporting the equilibrium is shown to be the equilibrium return generated by a model without robustness. Because of this, matching both the equity premium and the riskfree rate is challenging.

[1]Department of Economics, Littauer Center, Harvard University, Cambridge MA 02138, USA. Email: pmaenhou@kuznets.harvard.edu. I am indebted to John Campbell, David Laibson and Andrew Metrick for guidance and suggestions. I would also like to thank Nick Barberis, Gary Chamberlain, Hanno Lustig, Michael Schwarz, Aaron Tornell, Raman Uppal, Luis Viceira, and workshop participants at Harvard for their comments. I gratefully acknowledge the financial support of the Fund for Scientific Research Flanders.

1 Introduction

In the last few years, many financial economists have focused on the question of how to optimally allocate a portfolio between a risky and a riskless asset. Most of this research was initiated by Samuelson's provocative result that the optimal portfolio share in equities of an investor with power utility facing i.i.d. stock returns is trivially constant over the life-cycle under the assumption of frictionless markets and the absence of labor income (Samuelson, 1969). This result stands in sharp contrast to the advice generally given by practitioners to invest less aggressively in equities as the planning horizon shortens. Recent work has shown how, for instance, the presence of time-varying investment opportunity sets (e.g. Campbell and Viceira, 1999), or of risky labor income (e.g. Cocco, Gomes and Maenhout, 1999), can rationalize these recommendations qualitatively.

A fundamental assumption in the bulk of this work on dynamic portfolio choice is the absence of any parameter uncertainty.[2] Typically, one obtains point estimates for the asset return parameters and assumes these are correct and fixed. With respect to second moments, one can argue that the limit of infinitely fine sampling would remove all estimation risk. However, first moments are notoriously hard to estimate (Merton, 1980, Cochrane, 1998, and Blanchard, 1993). In particular, there exists substantial disagreement about the expected excess return on stocks: for a variety of reasons, many expect lower excess returns in the future (e.g. general equilibrium implications of the recent increase in stock market participation (Heaton and Lucas, 1999); historically low dividend yields (Campbell and Shiller, 1998)). The one solid conclusion from this debate seems to be that there exists substantial uncertainty about the excess return (Cochrane, 1998), or even about the model generating excess returns (see for instance Pastor and Stambaugh (1999) for an analysis of structural breaks). Welch (1999) recently conducted a survey among financial economists exactly with the intent of measuring consensus estimates of the equity premium, and found high dispersion in the reported forecasts. Another reason for scepticism about the reliability of standard historical estimates of the equity premium might be based on the work by Brown, Goetzmann and Ross (1995) and by Jorion and Goetzmann (1999), who argue that historical studies suffer from severe ex-post survival bias.

For those reasons, it is desirable to take parameter uncertainty into account when studying optimal dynamic portfolio and savings decisions. In this paper, I derive consumption and portfolio rules that are robust to a particular type of model misspecification, stemming from parameter uncertainty. An enormous literature in applied mathematics and engineering focuses on robust control.

[2]Notable exceptions are Barberis (1999) in a Bayesian framework, and Brennan (1997) focusing on continuous-time learning or filtering.

The main goal is to design control rules that not only work well when the underlying model for the state variables holds exactly, but also perform reasonably well if there is some form of model misspecification. A key assumption in this line of work is that the decision-maker worries about some worst-case scenario. Admittely, this constitutes a major deviation from standard expected-utility maximization, but it does not imply that the agent is infinitely risk-averse or that any misspecification goes. In particular, the disparity between the 'reference model' that the agent is sceptical about and the worst-case alternative model that he considers is constrained by a (preference) parameter, quantifying the strength of the preference for robustness. In addition, maxmin expected utility has strong foundations in decision theory (Gilboa and Schmeidler, 1989), and can be rationalized from a neobayesian perspective in environments where model uncertainty is so pervasive that the decision-maker entertains an entire family of priors rather than a single prior.

While initial work in applied mathematics was done in a linear-quadratic framework, recently progress has been made on robust control for nonlinear environments. Hansen and Sargent (1995) have extended the framework to allow for discounting, and therefore made the fundamental link to economic applications. Most recently, they show how continuous-time modeling makes it possible to tackle problems involving nonlinear state equations and nonlinear objective functions (Anderson, Hansen and Sargent (1999), henceforth AHS). This is of course exactly the type of set-up that is required for analyzing optimal portfolio decisions and finance problems. In terms of methodology, this paper uses the continuous-time framework built by AHS to solve an otherwise standard portfolio problem. The notion of robustness used by AHS is that of minimum entropy. The reason is partly analytical tractability, yet at the same time it turns out that this is the natural criterion to use if the misspecification concern relates to the expected equity premium.

I first solve the optimal decision problem of an investor facing i.i.d. returns. For power utility, the problem cannot be solved analytically due to a fundamental non-homotheticity. I demonstrate that homotheticity can be imposed by scaling the parameter that measures the strength of the preference for robustness in a natural way. Analytical solutions can be obtained and show that robustness drastically reduces the demand for the risky asset. The endogenous worst-case scenario that the agent hedges against is easily computed and provides guidance in the choice of a reasonable value for the parameter measuring the strength of the preference for robustness. In addition, imposing homotheticity makes robustness observationally equivalent to stochastic differential utility (the continuous-time version of recursive preferences in the sense of Epstein and Zin (1989)). Robustness is then interpreted as increasing risk-aversion without changing the preference for intertemporal substitution.

When deriving the optimal portfolio weights for a robust investor facing time-varying expected returns, I demonstrate that robustness magnifies the importance of hedging demands for stocks. Even logarithmic investors, who simply adopt a myopic investment strategy under expected utility, hold hedging-like portfolios when worrying about model risk. For the empirically plausible case of negative correlation between innovations in expected returns and shocks to actual returns, the hedging demands are always positive for investors who are at least as risk-averse as a log agent, in line with previous results for expected-utility. Negative hedging demands obtain for stochastic volatility, because of the negative correlation between innovations in volatility and innovations in returns.

After these partial-equilibrium exercises, I explore the equilibrium asset pricing implications of robust decision-making in the context of a simple Lucas (1978) exchange economy. The historically observed equity premium and riskfree rate can be matched with log utility. However, because the endogenous worst-case scenario that robust investors guard against turns out to be the equilibrium equity premium that would arise in an expected-utility model, an excessive degree of pessimism or uncertainty-aversion is required. The equity premium puzzle shows up again, albeit in a somewhat different form: the equity premium can only be rationalized when (logarithmic) agents worry about a pessimistic scenario that involves an equity premium of 3 basis points. Power utility does not have to rely on as much pessimism in terms of the worst-case scenario for stock returns. At the same time however, the riskfree rate supported by the equilibrium becomes excessively high, due to the reluctance to substitute intertemporally.

The organization of the paper is as follows. Section 2 lays out the framework used, as developed by Anderson, Hansen and Sargent (1999). It is repeated here for convenience. I also discuss the link between robustness and optimal decision theory. Section 3 studies the basic portfolio choice problem (i.i.d. returns) and its solution. Section 4 then explores the simplest possible time-varying expected return and stochastic volatility processes, in order to analyze hedging demands. Section 5 analyzes the equilibrium model, presents the informal calibration exercise, and relates to the literature. Finally, section 6 concludes. Appendices A, B and C contain the proofs that are omitted from the main text, as well as a brief extension of robustness to pure jump processes.

2 Robustness in Continuous Time

The first part of this section follows closely the exposition in AHS. They carefully develop a general framework for dealing with robustness in economics, and then specialize to the case of continuous time so as to allow for nonlinear objective functions and state equations. The remainder of the section relates to the decision-theoretic literature on maxmin expected utility.

2.1 A State Equation and Preference for Robustness

2.1.1 General Ingredients

Broadly speaking, Anderson, Hansen and Sargent develop a framework where robustness amounts to altering the Bellman equation, the fundamental tool of dynamic programming. As a starting point, AHS find it useful to focus on the Bellman equation while suppressing the optimal control problem. Ignoring optimal controls for the moment, the standard Bellman equation for value function $V(x)$ can then in general be formulated as:

$$V(x) = sU(x) + \exp(-s\delta)T_s(V)(x), \tag{1}$$

where $U(.)$ is the utility function and $\delta > 0$ is the discount rate. This functional equation will first be specialized to discrete-time environments by setting s, the length of the time interval, equal to one. Later on, the continuous-time case is covered by letting s shrink to zero (after dividing by s).

An ingredient of (1) that is of special interest to the robust decision-maker is the expectations operator T_s, defined as:

$$T_s(f)(y) = E\left[f(x_s)|x_0 = y\right]. \tag{2}$$

This expectations operator, used to compute the continuation payoff in the Bellman equation, reflects a particular underlying model or transition equation for the state variable x (in general a vector). The decision-maker accepts the state equation (also called 'reference model') as useful, but suspects it to be misspecified. She therefore wants to consider alternative models or state equations (also called 'candidate models') when computing her continuation payoff. These alternative models, each indexed by w, will be represented by means of the associated conditional expectations operator T_s^w. Loosely speaking, a preference for robustness is then achieved by having the agent guard against the worst alternative model that is reasonably similar or close to the reference model.

Representing alternative models or state equations w in the form of their associated expectations operator T_s^w, proves very useful for the problem of robust decision-making, because it naturally accommodates the introduction of a notion of distance or similarity between the reference model and alternative models, as will be made precise in the next section.

2.1.2 Discrete Time

The main spirit of the framework of AHS is most transparent in discrete time ($s = 1$ in equations (1) and (2), and suppressed for simplicity). AHS show how a preference for robustness amounts to adjusting the expected continuation payoff $T(V)(x)$ as follows. Consider the 'twisted' expectations operator $T^w(f)$ associated with the candidate model indexed by w (a strictly positive function):

$$T^w(f) = \frac{T(wf)}{T(w)}, \tag{3}$$

where the initial or conditioning value for the state variable x is suppressed for notational simplicity. The operator $T^w(f)$ coincides with the expectation $T(f)$ generated by the reference model when w is a constant function (by virtue of the scaling by $T(w)$ in the denominator).

The decision-maker worries that the reference model is misspecified, and that his continuation payoff might therefore have to be evaluated in light of some candidate model w, i.e. using $T^w(V)$ instead of $T(V)$, in the Bellman equation. Of course, not any misspecification goes, and when evaluating different misspecifications, a penalty is incurred for candidate models that are too far away from the reference model. This notion of distance is formally measured by relative entropy $I(w)$, the expectation of the log Radon-Nikodym derivative of the candidate model's measure with respect to the reference model's measure, where the expectation is taken with respect to the candidate model w.[3] AHS show how this can be rewritten in terms of the reference model's expectations operator as:

$$I(w) = \frac{T(w \log(w))}{T(w)} - \log(T(w)). \tag{4}$$

It is easy to see that candidate models w for which w is a constant function have relative entropy equal to zero. For all other perturbations w, $I(w)$ is strictly positive, given the restriction that w be a strictly positive function.

Of course, insisting on using relative entropy to quantify the similarity between different models limits the possible perturbations one can consider. The perturbations have to be absolutely continuous with respect to the reference model, so that the Radon-Nikodym derivative (and hence the

[3] Introducing relative entropy formally on the basis of the concepts formulated so far is complicated by the fact that the operator $T(f)(y)$ is defined as a function of today's state only ($x_0 = y$ is the conditioning vector in T). AHS therefore extend T to be a function of both today's state y and of tomorrow's state z, and denote this concept by $E[.|y]$. Relative entropy is then defined as $I(w)(y) = E^w\left[\log \frac{w(z)}{T(w)(y)}|y\right]$, where $\frac{w(z)}{T(w)(y)}$ is the Radon-Nikodym derivative of the candidate model's measure with respect to the measure of the reference model.

entropy measure) exists. As will be clear from the next section, this restriction will be especially crucial, but at the same time quite natural, in continuous-time diffusion environments.

Formally, robustness is then modeled as follows: to guard against the worst misspecification that is close to the reference model, the agent considers:

$$\inf_w \frac{1}{\theta} I(w) + T^w(V), \tag{5}$$

where $I(w)$ is the relative entropy measure and $\theta \geq 0$ measures the strength of the preference for robustness ($\theta = 0$ corresponds to expected-utility maximization). Therefore the more robust decision-maker (θ larger) has less faith in the reference model and will consider perturbations with larger relative entropy in order to evaluate her continuation payoff.

Given an infimizer w^* the robust Bellman equation is then simply:

$$V(x) = U(x) + \exp(-\delta)\left[T^{w^*}(V)(x) + \frac{1}{\theta} I(w^*)\right]. \tag{6}$$

2.1.3 Continuous Time

The continuous-time equivalent of the Bellman equation is typically the following Hamilton-Jacobi-Bellman (HJB) equation, obtained from (1) as $s \to 0$:

$$0 = U(x) - \delta V(x) + \mathcal{D}V(x), \tag{7}$$

where $\mathcal{D}f$ is defined as $\lim_{s \to 0} \frac{T_s(f) - f}{s}$, i.e. the infinitesimal generator of the family of expectations operators $\{T_s : s \in \mathsf{R}^+\}$, sometimes also called the differential or Dynkin generator (Merton, 1971). Heuristically, $\mathcal{D}f$ is often described as $\frac{1}{dt}E(dV)$ and easily obtained using Ito's Lemma. AHS demonstrate that the continuous-time analog of the robust adjustment described in the previous section, is derived by taking the limit as the time interval s (indexing the operator in (2)) shrinks to zero, while paying attention to some important subtleties.

Following the discrete-time set-up, robustness is introduced by replacing the differential operator $\mathcal{D}V(x)$ by $\mathcal{D}^wV(x)$ for some candidate model indexed by w. Again, only candidate models that are 'similar' to the reference model are of concern to the decision-maker, which is achieved by imposing a penalty (depending of course on θ) for far-away perturbations. Because the continuous-time limit of the relative entropy measure collapses to zero, AHS suggest considering its derivative $I'(w)$. This is exactly equivalent to the continuous-time treatment of the expectations operator, as $TV(x)$ or $T^wV(x)$ in the discrete-time Bellman equation become $\mathcal{D}V(x)$ or $\mathcal{D}^wV(x)$ in the continuous-time HJB equation. Then the analog of equation (6) is:

$$0 = U(x) - \delta V(x) + \mathcal{D}^{w^*} V(x) + \frac{1}{\theta} I'(w^*). \tag{8}$$

AHS model the preference parameter θ that weights the entropy penalty as fixed and state-independent. However, a generalization that will turn out to be crucial for the remainder of this paper replaces θ by a state-dependent (or properly scaled) version of θ, denoted by $\psi(V) > 0$. In particular, scaling θ by the value function V, i.e. having $\psi(V) \equiv \frac{\theta}{f(V)}$ for some function f, is an important modification for the specific environment I analyze, namely portfolio choice with CRRA utility. As will become clear from the next section, this modification is instrumental in assuring the homotheticity or scale-invariance of the resulting preferences. The robust version of the HJB equation then becomes:

$$0 = U(x) - \delta V(x) + \mathcal{D}^{w^*} V(x) + \frac{1}{\psi(V)} I'(w^*). \tag{9}$$

Finally AHS show how this HJB equation simplifies significantly when the state vector x follows a diffusion:

$$dx_t = \mu(x_t)dt + \Lambda(x_t)dB_t, \tag{10}$$

where $\Lambda(x_t)$ is the diffusion coefficient matrix, so that the variance-covariance matrix is $\Sigma(x_t) \equiv \Lambda(x_t)\Lambda(x_t)'$.

In that case the 'twisted' differential continuation payoff $\mathcal{D}^w V$ and the entropy measure $\frac{1}{\psi(V)} I'(w)$ needed for the robust HJB equation can be rewritten as:

$$\mathcal{D}^w V = \mathcal{D}V + u(x)'\Sigma(x)\left(\frac{\partial V}{\partial x}\right), \tag{11}$$

$$\frac{1}{\psi(V)} I'(w) = \frac{1}{2\psi(V)} u(x)'\Sigma(x)u(x), \tag{12}$$

where the perturbation is now indexed by $u(x) \equiv \frac{\partial \log w(x)}{\partial x}$. This simplifies the subsequent analysis significantly. Moreover, as AHS point out, equation (11) also gives insight into the type of misspecification that a minimum-entropy robust decision-maker worries about: because of (11), $u(x)$ can be thought of as a drift added to the diffusion in state equation (10):

$$dx_t = \mu(x_t)dt + \Lambda(x_t)\left[\Lambda(x_t)' u(x)dt + dB_t\right]. \tag{13}$$

The decision-maker acts as if the drift were $\left[\mu + \Lambda\Lambda' u\right] dt$, rather than μdt. Doing so comes at a cost, namely $\frac{1}{2\psi(V)} u'\Sigma u$, effectively restricting the agent's attention to candidate models that are reasonably similar to the reference model.

The fact that general model uncertainty is reduced to uncertainty about the drift of the state variable is a consequence of the assumption of an Ito process for the state variable, in combination with the use of relative entropy. The reason is Girsanov's Theorem: under technical conditions, a given Ito process can always be transformed into an alternative Ito process with a different drift coefficient through a change of measure. The measures have to be equivalent or absolutely continuous. This is exactly what is imposed by the use of relative entropy.

Fortunately, this restriction is entirely natural for the portfolio problem I am interested in, as a preference for robustness was motivated in the introduction by the large degree of uncertainty about the expected equity return, i.e. precisely the drift parameter in equation (13).

A later section in the paper shows how to adapt this concept of robustness developed by AHS to a setting where the state variable is believed to follow a pure jump process, rather than a diffusion process.

2.2 A Robust Control Problem

The final step in the development of the framework for robust decision-making consists of the introduction of a control variable (or vector). The HJB equation of a robust decision-maker choosing control variable i looks as follows:

$$0 = \sup_i \inf_u \left[U(x,i) - \delta V(x) + \mathcal{D}^{(i)} V(x) + u(x,i)' \Sigma(x,i) \left(\frac{\partial V}{\partial x}(x) \right) + \frac{1}{2\psi(V)} u(x,i)' \Sigma(x,i) u(x,i) \right],$$

where $\mathcal{D}^{(i)} V(x)$ is the differential generator, given control i.

The F.O.C. for the 'optimal' misspecification u^* is:

$$u^* = -\psi(V) \left(\frac{\partial V}{\partial x}(x) \right). \tag{14}$$

It is easy to see that this robust optimal control problem nests the standard expected-utility maximization approach ($u^* = 0$ for $\psi(V) = 0$). For strictly positive values of $\psi(V)$ (or of θ), this HJB equation can be interpreted as representing a (differential) zero-sum game between the decision-maker choosing control i and nature who, the agent believes, belligerently deceives her by choosing a perturbation u^* to the state equation.

2.3 Motivation and Link with Decision Theory

In the framework developed above, the decision-maker takes a very conservative and pessimistic perspective and worries about the worst-case misspecification. However, standard decision theory, ubiquitous in economics and finance, typically insists on expected-utility maximization and Bayesian decision-making. Several arguments can be formulated to defend the approach taken here.

First, a Bayesian approach seems conceptually limited to cases of parametric model uncertainty, whereas robustness (risk-sensitivity, minimum-entropy, or H_∞ control) can in principle handle general model uncertainty. But, as was shown above, minimum-entropy robustness combined with the tight structure of continuous-time diffusion processes, limits the type of model uncertainty to parametric uncertainty about the drift of the state variable. A Bayesian approach would therefore be possible. Indeed, a number of papers have investigated the effects of parameter uncertainty on financial decision-making and asset pricing (see for instance Barberis (1999), Brennan (1997) and Gennotte (1986)).

A more fundamental motivation for the maxmin behavior exhibited in the framework above comes from the vast literature on decision theory, started by Wald's classical exploration of maxmin (statistical) decision-making (1950). More recently, maxmin expected utility has been founded axiomatically by Gilboa and Schmeidler (1989). They offer an illuminating example due to Ellsberg (1961) to motivate maxmin decisions. Ellsberg's famous thought experiment is constructed as follows. A decision-maker is asked to rank 4 bets. The bets concern the color of a ball drawn from either urn A or urn B. The decision-maker is told that both urns contain each 100 balls that are either red or black. In addition, urn A is known to contain exactly 50 balls of each color. Bet 1 pays a prize when a ball drawn randomly from urn A is black; bet 2 when a red ball is drawn. Bets 3 and 4 are defined similarly, but for a ball drawn from urn B. Empirically, people seem indifferent between the first 2 bets and between the last 2 bets. But most people do prefer bets 1 and 2 to bets 3 and 4. Ellsberg notes that this behavior fails to satisfy the Savage postulates. Indeed, there is no unique prior that generates these choices when combined with expected utility maximization: indifference between bets 3 and 4 requires a 50/50 prior, which then leads to indifference between 1 and 3 or 2 and 4.

The explanation axiomatized in Gilboa and Schmeidler is that a decision-maker operating in an information vacuum entertains a family of priors rather than a single prior, and computes expected utility under the worst-case prior. They show that this decision-making process satisfies the usual expected-utility axioms, extended to allow for multiple priors, and complemented with certainty-independence (a weakening of the standard independence axiom) and uncertainty aversion. In this sense, the framework developed by Anderson, Hansen and Sargent can be interpreted as consistent with the neobayesian paradigm that allows for multiple priors.

Of course, no effort is made here to formally introduce and characterize the entire set of multiple priors. A rigorous treatment of asset pricing in such a setting can be found in Epstein and Wang (1994).[4] On the other hand, the approach taken here does allow me to calculate the endogenous worst-case scenario that supports the investor's decisions. As was shown above (equation (14)), this scenario is proportional to $\psi(J)$ and hence to θ, as I will use the scaling $\psi(V) \equiv \frac{\theta}{f(V)}$ for some function f. This allows for an interesting interpretation of θ that also relates well to the Ellsberg example. In particular, one can view θ as indexing the set of priors that the decision-maker entertains. For $\theta = 0$, this set collapses to a singleton, which brings us back to expected-utility, or unique-prior decision-making. As θ increases, the cardinality of the set of priors grows, so that the least-favorable prior or worst-case misspecification becomes more unfavorable. How does this relate to the Ellsberg example? The set of possible point-mass priors on the composition of urn B is $\{(0, 100), (1, 99), ..., (100, 0)\}$. Standard expected-utility decision-making corresponds to using $(50, 50)$ as unique prior, or having $\theta = 0$. Maxmin expected utility allows for a wider set of priors, and one could model the cardinality of this set of priors as being governed by a parameter θ measuring the strength of the preference for robustness, or the degree of uncertainty aversion. In other words, a larger θ then implies more uncertainty aversion so that the decision-maker wants to take less favorable compositions of urn B into account. This is important, because while people do tend to prefer bets 1 and 2 to bets 3 and 4, they would normally be willing to pay a positive price for bets 3 and 4, which precludes $(0, 100)$ as least-favorable prior. Indeed, experiments have been designed to elicit the degree of uncertainty aversion as reflected in the difference between the willingness to pay for bets 1 and 2 and bets 3 and 4 (see for instance Camerer, 1997). One could use this experimental evidence to calibrate θ. Instead, I will judge the plausibility of a given value for θ by examining the implied least-favorable prior, along the lines of equation (14).

3 Basic Portfolio Problem: I.i.d. Returns

3.1 General Set-Up

I will consider the simplest possible dynamic portfolio problem, where the agent has to maximize 'expected' life-time utility from consumption of a single good and has access to two financial assets: one

[4] Dow and Werlang (1992) offer an interesting application of uncertainty aversion and non-expected utility in the spirit of Gilboa and Schmeidler to portfolio choice in an atemporal environment.

riskless paying an instantaneous return r, and one risky (equities) paying a constant instantaneous expected excess return of $\mu - r$.

The objective function to maximize (in the absence of a preference for robustness) is:

$$\int_0^T \exp(-\delta t) \frac{C_t^{1-\gamma}}{1-\gamma} dt,$$

where γ denotes the coefficient of relative risk aversion. The limit as γ tends to unity is defined as $\int_0^T \exp(-\delta t) \log(C_t)\, dt$.

The price of the risky asset evolves according to the standard geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{15}$$

Therefore the state equation for wealth is:

$$dW_t = \left[W_t \left(r + \alpha_t(\mu - r)\right) - C_t\right] dt + \alpha_t \sigma W_t dB_t, \tag{16}$$

where α_t is the fraction of wealth invested in the risky asset at time t, i.e. the first control of the agent. The second control variable is consumption C_t. As always, both controls are nonanticipating and suitably adapted to the σ-algebra generated by the underlying Brownian motion.

Denoting the value function by $J(W,t)$ and its partial derivatives with respect to x by J_x, the robust control problem becomes:

$$0 = \sup_{\alpha, C} \inf_{u} \left[\frac{C_t^{1-\gamma}}{1-\gamma} - \delta J(W,t) + \mathcal{D}^{(\alpha,C)} J(W,t) + J_w \alpha^2 \sigma^2 W^2 u - \frac{1}{2\psi(V)} \alpha^2 \sigma^2 W^2 u^2\right], \tag{17}$$

where according to Ito's Lemma:

$$\mathcal{D}^{(\alpha,C)} J = J_w \left[W\left(r + \alpha(\mu - r)\right) - C\right] + J_t + \frac{1}{2} J_{ww} \alpha^2 \sigma^2 W^2,$$

with boundary condition:

$$J(W,T) = 0. \tag{18}$$

Solving first for the infimization part of the problem yields:

$$u_t^* = -\psi(V) J_w. \tag{19}$$

If the agent desires no robustness or has complete faith in the validity of the model ($\psi(V) = 0$), then $u^* = 0$, i.e. there are no perturbations to guard against. Given the fact that $\psi(V) > 0$ in the robust case, this means that the perturbation amounts to a negative drift added to the diffusion component of the state equation if $J_w > 0$ (which must hold in equilibrium due to the first-order condition for consumption).

Substituting for u^* in the HJB equation gives:

$$0 = \sup_{\alpha,C} \left[\frac{C_t^{1-\gamma}}{1-\gamma} - \delta J + J_w \left[W \left(r + \alpha(\mu - r) \right) - C \right] + J_t + \frac{1}{2} \left(J_{ww} - \psi(V) J_w^2 \right) \alpha^2 \sigma^2 W^2 \right], \qquad (20)$$

subject to (18). Then the necessary optimality conditions for consumption and portfolio choice are:

$$\left(C^* \right)^{-\gamma} = J_w, \qquad (21)$$

$$\alpha^* = \frac{-J_w}{\left[J_{ww} - \psi\left(J\right) J_w^2 \right] W} \frac{\mu - r}{\sigma^2}. \qquad (22)$$

It is interesting to note that the *structure* of the optimality condition for consumption is not affected by the introduction of robustness. This is not the case for the portfolio rule. Setting $\theta = 0$ yields the familiar result that the optimal portfolio rule is just the myopic or logarithmic $\alpha = \frac{\mu - r}{\sigma^2}$, adjusted for risk aversion $\frac{-J_{ww} W}{J_w}$. Robustness adds an extra term to this 'risk aversion' adjustment, equal to $\psi\left(J\right) J_w W > 0$. However, the overall quantitative or even qualitative effect of robustness on portfolio choice cannot be evaluated without knowledge of the explicit value function or of its partials.

The next step is usually to substitute conditions (21) and (22) back into the PDE (20) and to try to solve for J. However for the problem at hand, the PDE is quite involved, so that guessing a trial solution is not straightforward. In particular, there exists no analytical solution when $\psi\left(J\right) = \theta$, as formulated in AHS. Although numerical methods could in principle be used (based on natural perturbation methods (see for instance Judd (1996) or Bender and Orszag (1978))), I will instead demonstrate in the next section how the investment problem can be solved analytically. Essentially, I propose a suitable choice of $\psi\left(J\right)$ or scaled version of θ that yields both homotheticity and analytical tractability.

3.2 Homothetic Robust Decision-Making: A Solution for Power Utility

The preferences induced by the robust adjustment for a power-utility investor are not homothetic when the preference parameter governing uncertainty-aversion, $\psi\left(J\right)$, is fixed and state-independent

$(\psi(J) = \theta > 0)$. The simplest way to see this is to go back to the general FOC's, (21) and (22) repeated here for convenience:

$$(C^*)^{-\gamma} = J_w,$$

$$\alpha^* = \frac{-J_w}{\left[J_{ww} - \psi(J) J_w^2\right] W} \frac{\mu - r}{\sigma^2}.$$

For the infinite-horizon case, a constant consumption-wealth ratio then requires that $J(W) = \kappa W^{1-\gamma}$, where κ is some negative constant (for $\gamma > 1$). Substituting this into the optimality condition for the portfolio weight results in $\alpha_t^* = \frac{(\mu-r)/\sigma^2}{\gamma + \theta(1-\gamma)\kappa W^{1-\gamma}}$. This portfolio weight is not independent of wealth. In particular, it increases in wealth and asymptotes to $\frac{\mu-r}{\gamma\sigma^2}$, the standard solution for expected utility. Robustness wears off as the state variable increases, at least for CRRA preferences.

Another way to see this is to remember the general HJB equation for a robust investor:

$$0 = \sup_{\alpha, C} \left[\frac{C^{1-\gamma}}{1-\gamma} - \delta J + J_w \left[W\left(r + \alpha(\mu - r)\right) - C\right] + \frac{1}{2}\left(J_{ww} - \theta J_w^2\right) \alpha^2 \sigma^2 W^2\right]. \tag{23}$$

The preference for robustness stems from the last term, θJ_w^2. If I assume for now that robustness does not change the curvature of the value function, it is clear that this term will vanish as wealth rises: when $J(W) = \kappa W^{1-\gamma}$, then $\theta J_w^2 = o(W^{-2\gamma-1})$, while $J_{ww} = o(W^{-\gamma-2})$.[5] As a consequence, robustness, proportional to a constant parameter θ, would only be expected to imply homothetic preferences for $\gamma = 1$ (log utility) or $\gamma = -1$ (quadratic preferences). The case of log utility is covered in Appendix B[6], and quadratic preferences are used in a discrete-time setting by for instance Hansen, Sargent and Tallarini (1999). Only for log or quadratic utility is J_w^2 of the same asymptotic order as J_{ww}.

While homotheticity is often only a modeling assumption made for convenience, it is important for a number of reasons. Although economies exhibit growth, rates of return are stationary. Second, when the scale of the state variable matters, natural unit-invariance of optimal decisions disappears and calibrations have to take this into account. Finally, homotheticity facilitates aggregation and the construction of a representative agent.

In order to preserve the homotheticity of the preferences, the above analysis suggests adjusting the set-up of Anderson, Hansen and Sargent in the following way. Focusing on power utility with

[5] If anything, robustness would be expected to make the value function more concave, thereby reinforcing the argument.

[6] Although homotheticity obtains for log utility even for a constant θ, it is shown in Appendix B that one should still scale θ appropriately, so that the homothetic results for power utility converge to the solution for logarithmic utility as γ tends to unity. This is achieved by using $\lim_{\gamma \to 1} \psi(J)$ in the HJB for log utility.

$\gamma \neq 1$, homotheticity can be imposed by making the parameter θ state-dependent. One natural way of scaling is to use $\psi(J(W,t))$, a scaled version of θ:[7]

$$\psi(J(W,t)) = \frac{\theta}{(1-\gamma)\,J(W,t)} > 0. \tag{24}$$

Recalling the FOC (19) for the 'optimal' misspecification that the robust investor guards against, we then get that:

$$u^* = \frac{\theta}{(1-\gamma)\,J(W,t)} J_w.$$

This imposes the desired homotheticity property: robustness will no longer wear off as wealth increases. This modified version of AHS minimum-entropy robustness can naturally be labeled 'homothetic robustness'. HJB equation (20) then becomes:

$$0 = \sup_{\alpha, C} \left[U(C) - \delta J + J_w \left[W\left(r + \alpha(\mu - r)\right) - C \right] + J_t + \frac{1}{2}\left(J_{ww} - \frac{\theta J_w^2}{(1-\gamma)\,J} \right) \alpha^2 \sigma^2 W^2 \right]. \tag{25}$$

The following results can then be obtained:

Proposition 3.1: *When $U(C) = \frac{C^{1-\gamma}}{1-\gamma}$, with $\gamma \neq 1$, equation (25) subject to (18) is solved by*

$$J(W,t) = \left[\frac{\left(1 - e^{-a(T-t)}\right)}{a} \right]^{\gamma} \frac{W^{1-\gamma}}{1-\gamma}, \tag{26}$$

where $a \equiv \frac{1}{\gamma}\left[\delta - (1-\gamma)r - \frac{1-\gamma}{2[\gamma+\theta]}\left[\frac{\mu - r}{\sigma}\right]^2\right]$. *The optimal portfolio and consumption rules, valid for* $\gamma > 0$, *are given by:*

$$C_t^* = \frac{a}{1 - e^{-a(T-t)}} W_t, \tag{27}$$

$$\alpha^* = \frac{1}{\gamma + \theta} \frac{\mu - r}{\sigma^2}. \tag{28}$$

Proof. See Appendix A.

This result is remarkably simple, in light of the complexity of the HJB equation (25). The optimal fraction of wealth invested in the risky asset is independent of both wealth and time, by virtue of the scaling of θ by $(1-\gamma)\,J$. The portfolio weight is the standard Merton solution, where the usual risk-aversion adjustment γ is replaced by $\gamma + \theta > \gamma$. Robustness amounts therefore to an increase in risk aversion, at least within the confines of the environment studied here. The consumption rule has the same structure as Merton's solution. The only difference is that the key

[7] I find it useful to divide by $(1-\gamma)\,J$, rather than by J, for two reasons. First, this ensures that $\psi(J)$ is always positive for positive θ, and avoids the use of the absolute value operator. Second, this scaling turns out to be economically meaningful in the sense that its limit as $\gamma \to 1$ is nontrivial once one knows the explicit form of J for $\gamma \neq 1$, and hence yields decision rules for general CRRA utility that nest logarithmic utility. See Appendix B for details.

parameter determining the consumption wealth ratio, a, reflects the different portfolio weight. The infinite horizon limit is straightforward: as $T \to \infty$, $C = aW$. The fact that the consumption strategy is not structurally affected by robustness, unlike the portfolio rule (at least in terms of its 'risk aversion' correction), leads to a more fundamental insight about the effect of robustness that is explored in the next section.

3.3 The Link with Stochastic Differential Utility

If I rewrite the HJB equation (20) obtained after solving the infimization problem, an interesting result appears:

$$0 = \sup_{\alpha, C} \left[U(C) - \delta J + D^{(\alpha,C)} J(W,t) - \frac{\psi(J)}{2} J_w^2 \alpha^2 \sigma^2 W^2 \right]. \tag{29}$$

As was noted by AHS for the general abstract control problem they analyzed, the structure of this PDE exactly coincides with the general PDE defining stochastic differential utility (the continuous-time version of recursive utility (e.g. Epstein and Zin (1989) and Weil (1990)), see e.g. Duffie and Lions, 1992). Following the terminology of Duffie and Lions or of Duffie and Epstein (1992a and 1992b), the aggregator is $U(C) - \delta J$, i.e. as for the standard additive utility case, and the variance multiplier is $\psi(J)$, the robustness parameter. For state-separable (i.e. expected) utility, the variance multiplier is zero, while the typical stochastic differential utility case is characterized by a variance multiplier that depends on the value function itself. Robustness as formulated by AHS, on the other hand, leads to a variance multiplier that is constant (θ).

Despite the very different motivation underlying both preference specifications[8], and the formal difference in the instantaneous aggregators, a strong result pertaining to the relationship between recursive preferences and homothetic robustness can be obtained:

Proposition 3.2: *An investor with a homothetic preference for robustness $\psi(J) = \frac{\theta}{(1-\gamma)J}$ and CRRA utility function $U(C) = \frac{C^{1-\gamma}}{1-\gamma}$, is observationally equivalent to an Epstein-Zin-Weil investor with elasticity of intertemporal substitution $\frac{1}{\gamma}$ and coefficient of relative risk aversion $\gamma + \theta$.*

Proof. See Appendix A.

Several useful implications and insights follow from this finding. A first consequence is technical: the observational equivalence allows me to refer to Schroder and Skiadas (1999) for rigorous proofs of existence and optimality.

[8] Caplin and Leahy (1999) provide yet another framework that relates to Epstein-Zin-Weil preferences, and that is based on psychological considerations. In their framework, Epstein-Zin-Weil preferences show up as a special case satisfying time-consistency.

Second, this result leads to a new interpretation of the effect of (homothetic) robustness. Given that the nonrobust agent has power utility, she is equally willing to substitute over time as across states (as the coefficient of relative risk aversion is γ, which is also the inverse of the elasticity of intertemporal substitution). What robustness does is to make the agent less willing to substitute across states (as the coefficient of relative risk aversion becomes $\gamma + \theta > \gamma$), without altering the willingness to substitute intertemporally (as the elasticity of intertemporal substitution remains $\frac{1}{\gamma}$). This explains the finding of Proposition 3.1, namely that while the portfolio rule is structurally affected by the introduction of robustness, the consumption rule (mainly a function of the elasticity of intertemporal substitution, and only a function of risk aversion through its effect on the portfolio rule) remains unchanged.

Proposition 3.2 also sheds new light on the discussion of homotheticity and on the issue of scaling the parameter θ. In the case of stochastic differential utility, the variance multiplier is interpreted as a coefficient of relative risk aversion when scaled by the value function (see Duffie and Epstein, 1992). When θ is state-independent, it acts so as to induce constant *absolute* risk aversion, thus breaking the homotheticity of the preferences. Homotheticity and constant *relative* risk aversion is restored by using $\psi(J) \equiv \frac{\theta}{(1-\gamma)J}$ rather than θ, as proposed in the previous section.

Some subtle differences between recursive preferences and robustness nevertheless persist and are worth emphasizing. First, although robustness is natural to consider when the investor only derives utility from terminal consumption, this is less obvious for Epstein-Zin-Weil investors. This point will be made in the next section when studying time-varying investment opportunity sets and hedging demands.

Second, the interpretation is entirely different, as will be shown for instance in the section on equilibrium asset pricing. Indeed, while a recursive investor has complete faith in the model she uses to form optimal decision rules, a robust agent acts according to a pessimistically perturbed version of the transition dynamics implied by this model.

Most importantly, robustness can reconcile low estimates of risk aversion obtained from experimental evidence or introspection, with high estimates of 'risk aversion' based on asset pricing data. This is not true for recursive preferences. The argument supporting this claim is based on the observation that both experimental evidence and introspection typically involve situations with well-specified events and probabilities. The preference for robustness is therefore not operational, and one obtains an estimate of just γ. Asset markets, on the other hand, might constitute an environment where events and probabilities are less clear-cut so that decision-makers might insist on robust decision rules. The estimate of 'risk aversion' one then obtains combines genuine risk aversion

(γ) with uncertainty aversion (θ), and is simply $\gamma + \theta$.

3.4 Informal Calibration and Interpretation of θ

To explore the quantitative effect of robustness on the portfolio rule as given by (28), an informal calibration is suggestive. More importantly, this also yields an interpretation of u^*, the endogenous 'optimal' worst-case scenario that the investor is worried about (see equation (19)), along the lines of (13). This in turn offers guidance in selecting a reasonable value for θ, a parameter that was so far left unspecified.

As shown by AHS, a minimum-entropy robust decision-maker solving an equation like HJB (17), can also be viewed as employing a 'twisted' Markov process (13) describing the evolution of her relevant state variable, rather than (10). For the problem at hand this concerns the intertemporal budget constraint. The modification amounts to adding an endogenous drift term to (16):

$$dW_t = \left[W_t \left(r + \alpha^*(\mu - r)\right) - C_t\right] dt + \alpha^* \sigma W_t \left[\alpha^* \sigma W_t u^* dt + dB_t\right].$$

Because all uncertainty in this budget constraint (i.e. the Brownian motion B_t) stems from the return on the risky asset, this implies that under the modified Markov process, the investor worries that the stock price evolves according to:

$$\begin{aligned} \frac{dS_t}{S_t} &= \left[\mu + \alpha^* W \sigma^2 u^*\right] dt + \sigma dB_t \\ &= \left[\mu - (\mu - r)\frac{\theta}{\gamma + \theta}\right] dt + \sigma dB_t. \end{aligned} \tag{30}$$

where the second equality obtains upon substitution of (19) and (28) for the optimal u^* and α^* respectively. Consequently, the investor worries that the expected excess return on the risky asset is not $\mu - r$, but rather μ_P, defined as:

$$\mu_P \equiv E_t^{w^*}\left[\frac{dS_t}{S_t} - rdt\right] = \frac{\gamma}{\gamma + \theta}(\mu - r)\, dt, \tag{31}$$

where $E_t^{w^*}[.]$ naturally denotes the expectation according to the endogenous 'optimally distorted' transition law w^*. Analogously relabeling the 'true' equity premium $\mu - r$ as μ_T, θ is then found to be:

$$\theta = \gamma \frac{\mu_T - \mu_P}{\mu_P}, \tag{32}$$

which is informative about reasonable ranges for θ.

Table 1 reports the optimal portfolio weight α^* allocated to the risky asset (from (28)) and the associated worst-case or pessimistic scenario μ_P supporting this optimal choice (from (31)).

The excess return $\mu - r$ used is 6% and the standard deviation σ of stock returns was set at 16%. Because it seems less natural to explore the effects of a concern for model uncertainty for relatively risk-tolerant agents, I focus on the case of $\gamma \geq 1$, and report results for $\gamma = 1$ and $\gamma = 5$. Repeating the exercise for different values of θ is informative about reasonable values for this free parameter.

Table 1: Portfolio Share Invested in Equities and Implied Pessimistic Scenario.

	$\gamma = 1$		$\gamma = 5$	
θ	α^*	μ_P	α^*	μ_P
0	2.3438	0.0600	0.4688	0.0600
0.1	2.1307	0.0545	0.4596	0.0588
0.5	1.5625	0.0400	0.4261	0.0545
1	1.1719	0.0300	0.3906	0.0500
2	0.7813	0.0200	0.3348	0.0429
5	0.3906	0.0100	0.2344	0.0300
10	0.2131	0.0055	0.1563	0.0200

A first observation, when focusing on columns 2 and 4, is that the preference for robustness dramatically decreases the optimal portfolio weight. Relative to the first row, corresponding to the standard expected-utility or $\theta = 0$ case, the share of wealth invested in equities falls with θ. Also, as is obvious from (31) or (32), the crucial parameter in terms of the implied least-favorable equity premium is θ/γ rather than θ. Naturally, for a given pessimistic equity premium, the more risk-averse agent invests less than the risk-tolerant investor. Although less obvious a priori, one could argue that pessimistic scenarios involving equity premia in the 3% range are quite plausible. Cochrane (1998), for instance, reports this value as the lower bound of a 95% confidence interval based on U.S. data. Recently, a number of interesting papers studied questions that can be brought to bear on this issue (e.g. Brown, Goetzmann and Ross (1995) analyzing survival of equity markets, Jorion and Goetzmann (1999) studying world markets, and Welch (1999) conducting a survey among financial economists, eliciting among other things a pessimistic forward-looking estimate of the equity premium). This discussion is left to the section on equilibrium pricing.

4 Time-Varying Investment Opportunity Sets and Hedging Demands

Recently, an active literature has emerged studying the explicit solution of Merton-type hedging demands resulting from time-variation in investment opportunity sets.[9] At the same time however, some empirical work questions the robustness of the predictability results that these portfolio choice models rely upon (see for instance Goyal and Welch (1999) and the references therein for a discussion of the ability of the dividend yield to forecast the equity premium). It seems therefore natural to analyze the optimal hedging demands for an investor who worries about model misspecification and insists on robustness. The framework used in this paper offers one possible strategy for doing so.

As a first source of hedging demands I analyze a mean-reverting risk premium. A second example considers stochastic volatility.

4.1 Mean-Reverting Risk Premium

A simple model of time-varying investment opportunity sets is the Ornstein-Uhlenbeck specification for the risk premium considered by Kim and Omberg (1996). They study optimal portfolio choice for a finitely-lived investor with general expected-utility HARA-preferences analytically, but without intermediate consumption. Wachter (1999) analyzes a similar problem for power utility and with intermediate consumption, using martingale techniques, thus necessitating the assumption of complete markets. Campbell and Viceira (1999) characterize and calibrate the infinite-horizon Epstein-Zin-Weil case (discrete-time recursive preferences) using a log-linear approximation.

I analyze the incomplete markets case with utility over terminal wealth. Although this ignores the important intertemporal consumption-savings decision, doing so is nonetheless interesting for a variety of reasons. First, except under the assumption of complete markets (Schroder and Skiadas, 1999 and Wachter, 1999), no closed-form solutions are available even without introducing robustness. Indeed, the results pioneered by Kim and Omberg are derived for an investor who only consumes at the terminal date. Therefore, my results are immediate extensions of Kim and Omberg's for investors who worry about model misspecification. Interestingly, robustness generates non-myopic portfolio rules even for logarithmic preferences. Second, focusing on terminal wealth isolates an environment where the distinction between robustness and recursive preferences along the lines of Epstein and

[9] Balduzzi and Lynch (1999), Barberis (1999), Brandt (1999), Campbell and Viceira (1999), Kim and Omberg (1996) and Samuelson (1991) consider time-variation in excess returns. Brennan and Xia (1998) and Campbell and Viceira (1998) look at the impact of time-varying interest rates. Finally, Brennan, Schwartz and Lagnado (1997) and Liu (1999) investigate the effects of both.

Zin (1989) or Weil (1990) is clearest, as it seems less natural to endow an investor with recursive preferences when she only consumes at the terminal date. Finally, the portfolio weights derived here can nevertheless be suggestive about the structure of the portfolio demands for an Epstein-Zin-Weil investor facing a time-varying investment opportunity set and consuming at each instant.

4.1.1 Set-up

The price dynamics for the risky asset are as given in (15), except that now the expected stock return μ_t is also given by a diffusion:

$$dS_t = \mu_t S_t dt + \sigma S_t dB_{1t}.$$

The price of risk or risk premium on this asset is defined as:

$$X_t = \frac{\mu_t - r}{\sigma},$$

following a mean-reverting Ornstein-Uhlenbeck process:

$$dX_t = \lambda(\overline{X} - X_t)dt + \sigma_x dB_{xt},$$

where λ, σ_x and $\overline{X}$ are all positive and constant.[10] The Wiener process B_{xt} is allowed to be correlated with the Wiener process affecting actual equity returns B_{1t}:

$$E\left[dB_{1t}dB_{xt}\right] = \rho dt,\ \ \rho \in \left[-1, 1\right]. \tag{33}$$

Because the robust techniques introduced above require independent Brownian motions, I rewrite the processes as follows. Given (33), we have that:

$$dB_{xt} = \rho dB_{1t} + \sqrt{1-\rho^2}dB_{2t},$$

where B_{2t} is a standard Wiener process independent of B_{1t}.

The HJB equation for a homothetically robust investor with utility over terminal wealth becomes:

$$0 = \sup_{\alpha}\inf_{u} \left[\mathcal{D}^{(\alpha)} J(W, X, t) + u'\Sigma\partial J + \frac{1}{2\psi\left(J\right)} u'\Sigma u\right], \tag{34}$$

with $\Sigma \equiv \Lambda\Lambda^{'}$, $\Lambda \equiv \begin{bmatrix} \alpha\sigma W & 0 \\ \sigma_x \rho & \sigma_x\sqrt{1-\rho^2} \end{bmatrix}$ and ∂J denoting the row vector of partial derivatives of $J(W, X, t), \partial J \equiv \begin{bmatrix} J_w \\ J_x \end{bmatrix}$. The infinitesimal generator is standard and given by $\mathcal{D}^{(\alpha)} J(W, X, t) =$

[10] Kim and Omberg accomodate in addition time-variation in σ. They show that the only relevant state variable is X_t. Because the next section deals explicitly with stochastic volatility, I choose to focus on time-variation in risk premia that only stems from time-variation in μ.

$J_w\left[W\left(r+\alpha(\mu-r)\right)\right]+J_x\lambda(\overline{X}-X)+J_t+\frac{1}{2}trace\left[\Sigma\partial^2 J\right]$. Homotheticity requires $\psi\left(J\right)=\frac{\theta}{(1-\gamma)J}$. Finally, the usual boundary condition applies:

$$J(W,X,T)=\frac{W^{1-\gamma}}{1-\gamma},\gamma>1. \tag{35}$$

I will focus on the case of $\gamma\geq 1$. Not only is this the empirically relevant part of the parameter space, it also ensures that the solutions are well-behaved.[11] As argued before, a preference for robustness seems moreover most plausible for relatively risk-averse investors. I do consider the limiting case of $\gamma=1$ (logarithmic utility): although the standard Merton hedging demands vanish for logarithmic investors, robustness reintroduces hedging. Finally, as we learned from (the proof of) proposition 3.2 and Schroder and Skiadas (1999), one cannot have 'too much' robustness when $\gamma<1$, because of the restriction that θ be strictly smaller than $1-\gamma$.

While the optimization problem is set up for $\gamma>1$, the optimal portfolio rule is valid for $\gamma=1$. The complete derivation for $\gamma=1$ is relegated to Appendix B.

The first-order conditions with respect to u are:

$$u^*=-\psi(J)\partial J,$$

which can be substituted back into the HJB equation (34) to give:

$$0=\sup_{\alpha}\left[\mathcal{D}^{(\alpha)}J(W,X,t)-\frac{\psi\left(J\right)}{2}\left[\alpha^2\sigma^2W^2J_w^2+2\rho\sigma\sigma_x\alpha WJ_wJ_x+\sigma_x^2J_x^2\right]\right]. \tag{36}$$

Again, robustness introduces a concern about the quadratic variation in the partial derivatives of the value function (in the spirit of recursive preferences or stochastic differential utility), weighted by $\psi\left(J\right)$, the importance of the preference for robustness. The first-order condition for portfolio choice are:

$$\alpha^*=\frac{-J_w}{\left[J_{ww}-\psi\left(J\right)J_w^2\right]W}\frac{X}{\sigma}+\frac{-J_{wx}}{\left[J_{ww}-\psi\left(J\right)J_w^2\right]W}\frac{\sigma_x\rho}{\sigma}+\frac{\psi\left(J\right)J_wJ_x}{\left[J_{ww}-\psi\left(J\right)J_w^2\right]W}\frac{\sigma_x\rho}{\sigma}. \tag{37}$$

This portfolio rule has a notable structure and consists of three components. The first term is the myopic demand, based only on the current investment opportunity set X. It is identical to the portfolio demand derived before in (22): the risk-aversion adjustment is complicated by the introduction of the preference for robustness. The second term in (37) represents the standard hedging demand. It is zero whenever there is no scope for hedging ($\rho=0$), no need for hedging ($\sigma_x=0$), or no preference for hedging ($J_{wx}=0$). Again, as for the myopic demand, the risk-aversion

[11] Kim and Omberg (1996) carefully characterize the expected-utility solutions for $\gamma<1$ and distinguish well-behaved solutions from nirvana solutions, where the investor's expected utility approaches infinity.

adjustment differs from the standard $\frac{-J_w}{J_{ww}W}$ factor. The last term in (37) is interesting and novel. Its structure resembles the one of the hedging demand, in that it disappears whenever hedging is infeasible or unnecessary. However, it does not vanish if the investor is uninterested in hedging ($J_{wx} = 0$), as it is driven by $J_w J_x$, rather than by the usual cross-partial J_{wx}. The source of this extra term is the concern about the quadratric variation in ∂J, which is present as long as $\theta > 0$, i.e. as long as the agent worries about model risk.

4.1.2 Solution

These components of the demand for the risky asset can be further analyzed using the explicit analytical solution to this investment problem.

Proposition 4.1: *Equation (36), subject to the boundary condition (35), is solved by*

$$J(W, X, t) = \frac{W^{1-\gamma}}{1-\gamma} \exp\left[(1-\gamma)\left(A(t) + B(t)X + \frac{C(t)}{2}X^2\right)\right]. \tag{38}$$

For $\gamma \geq 1$, the optimal portfolio rule is:

$$\alpha_t = \frac{1}{\gamma+\theta}\left[\frac{X}{\sigma} + (1-\gamma-\theta)\left[B(t) + C(t)X\right]\frac{\sigma_x \rho}{\sigma}\right], \tag{39}$$

where the functions $A(t)$, $B(t)$ and $C(t)$ are the solutions to the system of differential equations given in Appendix A.

Proof. See Appendix A.

This solution to the dynamic portfolio problem is a direct extension of Kim and Omberg, allowing for robustness. It is easy to show that the portfolio weight collapses to their solution for power utility when is θ set to zero: the hedging component of the portfolio is then proportional to $(1-\gamma)$. Otherwise, robustness increases the relative importance of the hedging demand, as it is now weighted by $(1-\gamma-\theta)$. At the same time, the 'risk-aversion' correction applied to the total asset demand is enhanced by robustness. The following lemma provides further insight.

Lemma 4.1: *Under the assumptions of Proposition 4.1, it can be shown that:*

$$C(t) > 0 \text{ and } B(t) > 0, \ \forall t < T,$$
$$C'(t) < 0 \text{ and } B'(t) < 0, \ \forall t < T.$$

Proof. See Appendix A.

We then easily obtain the following results:

Corollary 4.1: *A homothetically robust investor with CRRA preferences ($\gamma \geq 1$) invests more in the risky asset than the myopic investor if and only if $\rho < 0$.*

The fact that the robust investor always invests more for the empirically plausible case of negative correlation between innovations to expected returns and innovations to actual returns, confirms the results derived in Kim-Omberg and Campbell-Viceira for $\gamma > 1$. The negative correlation between shocks to returns and shocks to risk premia implies that the risky asset tends to pay off well when future investment opportunities are expected to worsen and therefore provides a hedge against this deterioration in investment opportunities. Relatively risk-averse investors on the one hand, value this hedge and increase their demand for equities. On the other hand, more aggressive and relatively risk-tolerant investors pay no attention to this intertemporal hedging aspect of the risky asset. Instead, they value assets that pay off well when wealth is most productive ($\rho > 0$). This behavior is termed 'speculative'. As a result, when $\rho < 0$, they invest less in equities than a myopic investor, and have a negative hedging demand.

A first finding is therefore that the hedging behavior analyzed in Kim-Omberg and Campbell-Viceira generalizes to investors with a preference for robustness. Secondly, it is shown here that a robust investor with log preferences also hedges. Indeed, the results above apply for $\gamma = 1$. In that case, the hedging demand is proportional to θ. These results are not surprising: the investor examined here has coefficient of relative risk aversion equal to or larger than unity, and a preference for robustness. Loosely speaking, the latter effectively makes the investor more risk-averse: the result then ensues immediately.

Another important aspect of these hedging demands, is that they induce a horizon effect into the investment strategy. In particular:

Corollary 4.2: *Abstracting from time-variation in the risk premium, a homothetically robust investor with CRRA preferences ($\gamma \geq 1$) invests more in the risky asset at long horizons than at short horizons if and only if $\rho < 0$.*

This horizon effect is entirely due to the hedging demand. The intuition is straightforward: hedging against adverse changes in investment opportunity sets is most relevant for investors with long horizons. Put differently, stocks are less useful as a hedging device at short horizons. Positive hedging demands (i.e. when $\rho < 0$) will therefore shrink as the investor ages. This pulls the portfolio rule towards the myopic strategy. The opposite reasoning applies when $\rho > 0$: in that case the negative hedging demand loses strength and leads the investor to increase her relative equity position as she ages.

Finally, it is of interest to investigate the effect of robustness on the portfolio rule. The intuition that robustness acts to make the investor effectively more risk-averse suggests that, while the overall equity demand would be expected to decline as θ rises, the importance of the hedging demand should increase (Campbell and Viceira, 1999). Analytical results are difficult to obtain, as they depend on the parameters governing the return dynamics. Instead, I conduct a simple calibration exercise.

4.1.3 Informal Calibration

Barberis (1999) estimates the discrete-time analogue of the mean-reverting process for the risk premium assumed here, using the dividend yield as a predictor for the expected excess return. Based on monthly data for 1986 to 1995, his estimates correspond to $\sigma = 0.0436$, $\sigma_x = 0.037$, $\lambda = 0.0423$, $\rho = -0.93$ and $\overline{X} = 0.0942$ (at a monthly frequency).

Table 2 reports the optimal allocation to equities (α) according to (39) for an investor with a 5-year horizon ($T - t = 60$) and for $X = \overline{X}$. I also compute the fraction of this asset demand due to hedging (in percentage terms), $\frac{\alpha_{hedging}}{\alpha}$, where $\alpha_{hedging} \equiv \frac{1-\gamma-\theta}{\gamma+\theta}\left[B(t) + C(t)X\right]\frac{\sigma_x \rho}{\sigma}$.

Table 2: Mean-Reverting Risk Premium: Total and Percentage Hedging Demand.

	$\gamma = 1$		$\gamma = 5$	
θ	α	% hedging	α	% hedging
0	2.1606	0.0000%	1.0348	58.2409%
0.1	2.1073	6.7955%	1.0212	58.5146%
0.5	1.9139	24.7410%	0.9701	59.5085%
1	1.7118	36.8916%	0.9131	60.5622%
2	1.4079	48.8470%	0.8168	62.2132%
5	0.9131	60.5622%	0.6203	65.1694%
10	0.5742	65.7938%	0.4426	67.4541%

Naturally, total asset demand declines as the preference for robustness gains strength. Simultaneously, the composition of the portfolio weight tilts away from the myopic component and towards the hedging component. For $\theta = 10$, the hedging demands make up two thirds of the total asset.

4.2 Stochastic Volatility

Another important source of hedging demands is stochastic volatility. Unlike the study of mean-reverting excess returns, the analysis of stochastic volatility has received little attention in the literature on portfolio choice. Two important exceptions are Chacko and Viceira (1999) and Liu (1999). Chacko and Viceira model the inverse of volatility as a square-root Ornstein-Uhlenbeck process and solve for the optimal consumption and portfolio decisions of an infinitely-lived investor with stochastic differential utility. They obtain closed-form solutions for log stochastic differential utility (unit elasticity of intertemporal substitution) and approximate analytical solutions for the general case. Liu considers more complicated return dynamics (stochastic volatility and stochastic short-term interest rates), but restricts attention to expected utility over terminal wealth. In this section, I study the portfolio decision of a robust investor with power utility facing stochastic volatility. This is important for several reasons. First, it was argued before that minimum-entropy robustness as developed by Anderson, Hansen and Sargent amounts to uncertainty about the expected equity return. This section introduces model uncertainty about second moments as well by allowing for stochastic volatility. Second, while ignoring short-rate risk for simplicity, I extend Liu's results and leave the realm of expected utility. Finally, because I focus on the case of utility over terminal wealth, exact closed-forms are obtained for all parametrizations. Robustness can then again be interpreted as a form of stochastic differential utility, and the results can be compared with the approximate results of Chacko and Viceira for intermediate consumption. This latter strategy proves useful as it is more natural to consider utility over terminal wealth from a robust perspective than from the viewpoint of stochastic differential utility, and given the fact that moreover closed-forms are not available for recursive preferences (or even expected utility) with intermediate consumption unless one resorts to the (unrealistic) assumption of complete markets.

4.2.1 Set-up

I consider the parametrization for return dynamics proposed in Chacko and Viceira (1999). They assume a square-root mean-reverting process for the inverse of volatility. This process is parsimonious and analytically tractable, and yields dynamics that are empirically plausible.

As before, the price of equity is assumed to follow a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dB_{1t}, \tag{40}$$

where the mean return μ is assumed to be fixed and the volatility v_t is now stochastic. In particular, $y_t \equiv \frac{1}{v_t}$, the precision of the equity return, is assumed to follow a square-root Ornstein-Uhlenbeck process:

$$dy_t = \kappa(\overline{y} - y)dt + \sigma\sqrt{y_t} dB_{yt}, \tag{41}$$

where κ, $\overline{y}$ and σ are all positive and constant and it is assumed that $2\kappa\overline{y} > \sigma^2$. The Wiener process B_{yt} is allowed to be imperfectly correlated with the Wiener process affecting equity returns B_{1t}:

$$E\left[dB_{1t} dB_{yt}\right] = \rho dt, \;\; \rho \in [-1, 1]. \tag{42}$$

Again, the process can be rewritten to involve only independent Wiener increments:

$$dy_t = \kappa(\overline{y} - y)dt + \sigma\rho\sqrt{y_t}dB_{1t} + \sigma\sqrt{y_t(1-\rho^2)}dB_{2t}, \tag{43}$$

where the processes B_{1t} and B_{2t} are now independent.

A simple application of Ito's Lemma gives the dynamics of the volatility v_t:

$$\frac{dv_t}{v_t} = (\kappa - \kappa\overline{y}v_t + \sigma^2 v_t)dt - \sigma\rho\sqrt{v_t}dB_{1t} - \sigma\sqrt{v_t(1-\rho^2)}dB_{2t}. \tag{44}$$

Thus, volatility is mean-reverting, and is subject to larger proportional changes when volatility is high than when it is low. Finally, the instantaneous correlation between proportional changes in volatility and equity returns is $-\rho$.

4.2.2 Solution

Given these return dynamics for equity and a constant riskless rate r, the Hamilton-Jacobi-Bellman equation for a (homothetically) robust investor with power utility over terminal wealth can be written as:

$$0 = \sup_{\alpha}\left[\mathcal{D}^{(\alpha)}J(W,y,t) - \frac{\psi(J)}{2}\left[\alpha^2 vW^2J_w^2 + 2\rho\sigma\alpha WJ_wJ_y + \sigma^2 yJ_y^2\right]\right], \tag{45}$$

with

$$\mathcal{D}^{(\alpha)}J(W,y,t) = J_wWr + \alpha J_wW(\mu - r) + J_t + J_y\kappa(\overline{y} - y) + \frac{1}{2}\left[\alpha^2 vW^2J_{ww} + 2\rho\sigma\alpha WJ_{wy} + \sigma^2 yJ_{yy}\right]$$

and subject to:

$$J(W,y,T) = \frac{W^{1-\gamma}}{1-\gamma},\ \gamma > 1. \tag{46}$$

The F.O.C. for optimality is:

$$\alpha_t^* = \frac{-J_w}{J_{ww}W - \psi(J)J_w^2W}\frac{\mu - r}{v} + \frac{-J_{wy} + \psi(J)J_wJ_y}{J_{ww}W - \psi(J)J_w^2W}\frac{\sigma\rho}{v}. \tag{47}$$

The general structure of this portfolio weight is similar to the one obtained for the mean-reverting risk premium. It consists of a myopic component, a hedging component (proportional to J_{wy}) and a component that stems from robustness (proportional to $\psi(J)J_wJ_y$). As before, the risk-aversion correction applied to these asset demands also has a component that arises because of the concern for model uncertainty, namely $\psi(J)J_w^2W$.

Proposition 4.2: *Equation (45), subject to the boundary condition (46), is solved by*

$$J(W,y,t) = \frac{W^{1-\gamma}}{1-\gamma}\exp\left[(1-\gamma)(A(t) + B(t)y)\right]. \tag{48}$$

For $\gamma \geq 1$, the optimal portfolio rule is:

$$\alpha_t = \frac{1}{\gamma + \theta} \left[\frac{\mu - r}{v} + (1 - \gamma - \theta) \, B(t) \frac{\sigma \rho}{v} \right], \tag{49}$$

where the functions $A(t)$ and $B(t)$ are the solutions to the system of differential equations given in Appendix A.

Proof. See Appendix A.

To analyze the portfolio share further, it is again necessary to solve the Riccati equation for $B(t)$. The structure is identical to the one encountered in the previous section. Because I focus on the case of $\gamma > 1$, the solution is well-behaved, with or without robustness.

Lemma 4.2: *Under the assumptions of Proposition 4.2, it can be shown that:*

$$B(t) > 0 \text{ and } B'(t) < 0, \ \forall t < T.$$

Proof. See Appendix A.

This leads to the following result:

Corollary 4.3: *A homothetically robust investor with CRRA preferences ($\gamma \geq 1$) invests more in the risky asset than the myopic investor if and only if $\rho < 0$.*

Chacko and Viceira estimate ρ to be large and positive. This implies that innovations to volatility are negatively correlated with innovations to equity returns. Because the expected excess return is constant, an investor who is at least as risk averse as a logarithmic investor will then have a negative hedging demand: equities are not useful as a hedge against adverse changes in investment opportunities given that they tend to pay off well when investment opportunities are advantageous. This is true whether or not the agent is robust: all the above results hold for $\theta \leq 0$ (except of course for logarithmic utility, where $\theta = 0$ results in the myopic investment strategy).

It is noteworthy that the long-term mean of volatility or of precision is not a determinant of the portfolio weight, as can be seen from the formula for $B(t)$ given in Appendix A. This result was also obtained by Liu, who considers $\theta = 0$, but for more complicated return dynamics.

In terms of horizon effects, Lemma 4.2 also implies:

Corollary 4.4: *Abstracting from time-variation in the risk premium, a homothetically robust investor with CRRA preferences ($\gamma \geq 1$) invests less in the risky asset at long horizons than at short horizons if and only if $\rho > 0$.*

The importance of the hedging demand decreases as the horizon shrinks. Combining this with the fact that the hedging demand is negative for $\rho > 0$ then yields that investors allocate a larger fraction of the portfolio to equities as they age, in contrast with the results obtained above for the time-varying risk premium.

Before investigating the quantitative effect of robustness on the demand for stocks, it can be noted that the portfolio weights obtained here also prove useful when solving the more general problem involving intermediate consumption, as analyzed in Chacko and Viceira for stochastic differential utility. They find the problem analytically intractable and resort to approximate analytical solutions. Their portfolio share can be compared with the one found here, by interpreting my results from the viewpoint of stochastic differential utility (taking $\frac{1}{\gamma}$ to be the elasticity of intertemporal substitution and $\gamma + \theta$ to be the coefficient of relative risk-aversion) and by taking the limit as the terminal date tends to infinity.

4.2.3 Informal Calibration

Chacko and Viceira (1999) estimate the process for stochastic volatility used here. Based on annual data, they obtain: $\overline{y} = 24.7718$, $\sigma = 1.1786$, $\kappa = 0.0426$, $\rho = 0.3708$ and $\mu - r = 0.0841$. As for the mean-reverting risk premium, I report the total asset demand and percentage hedging demand for an investor with a 5-year horizon ($T - t = 5$) and for $y = \overline{y}$ (where $y \equiv \frac{1}{v}$).

Table 3: Stochastic Volatility: Total and Percentage Hedging Demand.

	$\gamma = 1$		$\gamma = 5$	
θ	α	% hedging	α	% hedging
0	2.0833	0.0000%	0.3912	-6.4951%
0.1	1.8798	-0.7507%	0.3835	-6.5263%
0.5	1.3519	-2.7365%	0.3552	-6.6398%
1	1.0007	-4.0884%	0.3252	-6.7604%
2	0.6587	-5.4297%	0.2783	-6.9496.%
5	0.3252	-6.7604%	0.1942	-7.2897%
10	0.1764	-7.3618%	0.1291	-7.5538%

Compared to Table 2, Table 3 shows that the total asset demand decreases very rapidly as the investor has a more pronounced preference for robustness. As θ rises, the myopic demand (28) shrinks. In addition, the hedging demand becomes more important. Unlike for the mean-reverting risk premium however, the hedging demand is now negative (because $\rho > 0$) and therefore helps to lower the total asset demand.

It can be noted that time-variation in excess returns leads to larger hedging demands than stochastic volatility (in absolute value). This might suggest that a more general modeling effort nesting both aspects of dynamic investment opportunity sets (time-variation in both μ and σ) would share most of the predictions of the previous section, namely positive hedging demands and optimal portfolio shares that increase as the investment horizon lengthens.

5 Equilibrium Asset Pricing

In order to explore the general equilibrium implications of the robust decision rules derived in the previous sections, I consider a simple representative-agent exchange economy in the style of Lucas (1978). The representative agent receives an endowment that he has to consume in equilibrium. Two assets can be traded in this economy: the risky asset, entitling the owner to the risky endowment (interpreted as the dividend), and the riskless asset. The returns or prices of these assets adjust to support a no-trade equilibrium. This approach therefore not only yields the equilibrium equity premium, but in addition provides insight into the determinants of the equilibrium riskfree rate. The latter is important: we know that a standard CRRA expected-utility model is capable of generating arbitrarily large equity premia, simply by increasing the coefficient of relative risk aversion γ. It is equally well known that this parameter choice simultaneously predicts a counterfactually high riskfree rate, forcing the rate of time preference to be negative (Weil, 1989). From the informal calibration in the previous partial-equilibrium section, it is easy to see that the demand for risky assets can be made arbitrarily small by assuming a strong preference for robustness (high θ), very much like a high γ would decrease equity holdings for expected-utility investors. I will show that an economy with robust agents is different from an economy with highly risk-averse expected-utility consumers, because it can generate both a high equilibrium equity premium and a low riskfree rate. However, the model relies on a worst-case scenario that seems excessively pessimistic. Increasing risk aversion in response, one encounters the standard difficulties of consumption-based asset pricing models. Robustness fails to simultaneously explain the high equity premium and low riskfree rate in a simple exchange economy.

5.1 The Equilibrium Model

The primitives of the model are as follows. For simplicity I assume that the dividend or endowment process is i.i.d. and characterized by a geometric Brownian motion:

$$dD_t = \mu_D D_t dt + \sigma_D D_t dB_t, \tag{50}$$

where the expected instantaneous growth rate μ_D and the instantaneous standard deviation σ_D are strictly positive parameters. It is natural to conjecture that the price S_t of the risky asset representing a claim on the dividend stream follows an Ito process as well:

$$dS_t = \left(S_t \mu_S - \frac{D_t}{S_t} \right) dt + \sigma_S S_t dB_t.$$

The coefficients μ_S and σ_S are to be determined from equilibrium conditions. The conjecture implies that the total return on the risky asset, consisting of both the dividend yield and the capital gain, is simply:

$$\frac{dS_t + D_t dt}{S_t} = \mu_S dt + \sigma_S dB_t.$$

Denoting as before the riskfree rate by r, and the fraction of wealth allocated to the risky asset by α, the representative agent's wealth dynamics are:

$$dW_t = \left[W_t \left(r + \alpha_t(\mu_S - r) \right) - C_t \right] dt + \alpha_t \sigma_S W_t dB_t. \tag{51}$$

Analyzing the infinite-horizon case and using the results from the previous section, the HJB for a robust investor becomes:

$$0 = \sup_{\alpha, C} \left[\frac{C^{1-\gamma}}{1-\gamma} - \delta J + J_w \left[W \left(r + \alpha(\mu_S - r) \right) - C \right] + \frac{1}{2} \left(J_{ww} - \frac{\theta}{(1-\gamma)J} J_w^2 \right) \alpha^2 \sigma_S^2 W^2 \right], \tag{52}$$

Definition: *A robust equilibrium*[12] *consists of a consumption rule C^*, an investment rule α^*, and prices S and r, such that simultaneously:*

1. *Markets clear continuously ($C^* = D$ and $\alpha^* = 1$),*

2. *The agent solves (52) subject to transversality condition (55).*

Repeated here for convenience, the optimality conditions are:

$$\alpha^* = \frac{1}{\gamma + \theta} \frac{\mu_S - r}{\sigma_S^2}, \tag{53}$$

$$C^* = aW, \tag{54}$$

$$\lim_{t \to \infty} E\left[e^{-\delta t} J(W_t) \right] = 0, \tag{55}$$

where $a \equiv \frac{1}{\gamma} \left[\delta - (1-\gamma) r - \frac{1-\gamma}{2[\gamma + \theta]} \left[\frac{\mu_S - r}{\sigma_S} \right]^2 \right]$.

[12] The standard term denoting the equilibrium in a Lucas pure exchange model is of course the Rational Expectations Equilibrium (REE). I do not use this term because robust decision-makers, although highly sophisticated, lack faith in the model and therefore do not form expectations rationally, i.e. using the model. See Hansen and Sargent (1999) for a discussion and an exploration of the Lucas critique in a robust decision-making framework.

5.2 Results

These optimality conditions obtained in the partial-equilibrium framework, combined with clearing in the security market, immediately yield the following:

$$\begin{aligned} \mu_S - r &= [\gamma + \theta]\, \sigma_S^2 \\ &= [\gamma + \theta]\, cov\left(\frac{dC}{C}, \frac{dS}{S}\right). \end{aligned} \tag{56}$$

This CCAPM result follows directly from the fact that consumption growth and equity return are by construction perfectly correlated in this simple model. The price of risk is given by $\gamma + \theta$. The first component is standard (as the coefficient of relative risk aversion is γ), while the second stems from robustness.

As suggested before, it is clear that for a given covariance between consumption growth and asset returns σ_{CS}, the equity premium can be made arbitrarily large by simply assuming a sufficiently large preference for robustness ($\theta >> 0$). This is of course also true in the standard CCAPM for CRRA expected-utility, where γ is substituted for $\gamma + \theta$ in (56). Therefore, assuming high values for θ in this model seems tantamount to relying on high risk aversion in the expected-utility environment. Fortunately however, this equivalence does not carry over to the equilibrium riskfree rate:

Proposition 5.1: *In equilibrium, the price of the risky asset is given by* $S_t = \frac{1}{a} D_t$. *The excess return on the risky asset follows:*

$$\frac{dS_t + D_t dt}{S_t} - rdt = [\gamma + \theta]\, \sigma_{CS} dt + \sigma_D dB_t, \tag{57}$$

and the equilibrium riskfree rate is given by:

$$r = \delta + \gamma \mu_D - \frac{1}{2} [1 + \gamma] [\gamma + \theta]\, \sigma_D^2. \tag{58}$$

Proof. See Appendix A.

First, as a useful consistency check, one can easily verify that (57) and (58) collapse to the standard formulae when the agent has complete faith in the model ($\theta = 0$). More importantly, the equilibrium condition for the riskfree rate reveals that the same force that produces a high equity premium ($\theta >> 0$), is also instrumental in keeping the equilibrium riskfree rate low through increased precautionary savings. These precautionary savings operate through the preference for robustness. However, unlike in the CRRA expected-utility paradigm, this effect is not mitigated by any upward pressure on the interest rate resulting from reluctance to substitute intertemporally. Indeed, the agents in this model have an elasticity of intertemporal substitution equal to $\frac{1}{\gamma}$, independent of their preference for robustness.

5.3 Informal Calibration and Interpretation: Quantifying the Pessimistic Scenario

Given values for the parameters μ_D $(= \mu_C)$, σ_{CS} and σ_D^2 $(= \sigma_C^2)$, the free preference parameters δ, γ and θ can be chosen to match the observed riskfree rate and equity premium according to (57) and (58). Table 4 reports the results using the parameter values in Campbell (1998).

Table 4: Preference Parameters Required to Match Riskfree Rate and Equity Premium.

μ_C	0.0176
σ_{CS}	0.0003
σ_C	0.0111
r	0.01
$\mu_S - r$	0.06
δ	0.015
γ	1
θ	188
γ_{EU}	189
δ_{EU}	-1.10
μ_P^*	0.0003

At first sight, the robust economy fares well: in order to produce the historically observed riskfree rate and equity premium, the model with robust *logarithmic* investors requires a discount rate of 1.5% and a preference for robustness θ of 188. In contrast, both the equity premium puzzle and the riskfree rate puzzle show up for standard expected utility: $\theta = 0$ requires a coefficient of relative risk aversion (denoted γ_{EU}) of 189 and a negative discount rate ($\delta_{EU} = -1.10$). The rate of time preference obtained in the robust setting is not implausible at all. The value obtained for θ is less easy to interpret. Table 1 in the partial-equilibrium section might be informative.

A yet better test involves computing the implied endogenous worst-case scenario that investors worry about. Towards such an interpretation, equation (31) is useful. Combining with the equilibrium condition that $\alpha = 1$, one obtains an interesting result:

$$\begin{aligned} \mu_P^* &= \gamma \sigma_S^2 \\ &= \gamma cov\left(\frac{dC}{C}, \frac{dS}{S}\right). \end{aligned} \tag{59}$$

Of course, this is the equilibrium equity premium in a model with expected-utility agents. The worst-case μ_P^* no longer depends on the preference parameter θ. Mechanically speaking, this is due to the equilibrium condition: one can generally (i.e. also in partial equilibrium) rewrite (31) as $\mu_P^* = \alpha\sigma_D^2$. Imposing $\alpha = 1$ eliminates θ from the equation. Intuitively, what's happening is that while μ_P^* no longer depends on θ, the endogenous equilibrium equity premium does. All θ does is to index the distance between the true and the pessimistic equity premium. In the partial equilibrium setting studied before, the true equity premium was exogenously given, so that the least-favorable excess return adjusted. Now the opposite happens.

The interpretation of the robust equilibrium is then as follows. Investors worry that the equilibrium equity premium is low and invest cautiously. This conservative behavior generates a high equity premium. But robust investors worry that this premium is too good to be true and keep on investing pessimistically, thus supporting the equilibrium. At the same time, precautionary savings keep the equilibrium riskfree rate low.

Unfortunately, this strong restriction has powerful implications: given the parameters of Table 4, the endogenous pessimistic equity premium μ_P^* that drives down the equilibrium demand for stocks and thereby increases the equilibrium equity premium is around 0.03%! Robustness thus does not quite solve the equity premium puzzle: the implied pessimistic scenario is simply too pessimistic. This perspective can be added to the list of different ways of looking at the equity premium and of noting how puzzling its historical generosity is.

How much pessimism seems reasonable? In his survey among financial economists, Welch reports 2% as the average pessimistic answer (defined as the lowerbound of a 95% confidence interval). Interestingly, he adds that "(...) It is remarkable that the typical pessimistic-case scenario foreseen by financial economists is roughly the return suggested by theory (...)" (Welch (1999), p. 23). This is exactly what happens in my model, at least if γ is sufficiently high.

Another useful source of information to judge the plausibility of the pessimistic scenario required to match the riskfree rate and the excess return on equities, comes from the work by Brown, Goetzmann, Ross (1995) and Jorion and Goetzmann (1999). They argue that by calibrating models using historical data for the U.S., those studies implicitly condition on the ex-post survival of the world's most successful equity market. By carefully collecting data on world stock markets, Jorion and Goetzmann report a median real equity return of only 0.8% for those markets. This stands in sharp contrast to the 4.3% observed for the U.S. stock market.[13] If agents are aware of this and

[13]This number might seem quite low. The reason is that the return is based on capital gains only, and excludes dividend yields. The authors argue that the results are not sensitive to this issue.

realize that the generous U.S. equity premium is at least partially the outcome of 'good luck', they might make savings and portfolio decisions as described here. Interpreted in this way, the equilibrium model presented here, formalizes the idea that the equity premium puzzle is due to the cautious behavior of investors who worry that the historically observed equity premium is not genuine. Their cautiousness can be rationalized by the findings of Brown, Goetzmann and Ross, and of Jorion and Goetzmann.

An obvious answer to the equity premium puzzle might be to make the agent more risk-averse: calibrating an economy where the representative agent has $\gamma >> 1$ and insists on robust decision rules might be more successful. The intuition from (59) and the discussion of the riskfree rate in this framework gives some hope. Unfortunately, the results are not encouraging as can be understood immediately from (58): as in expected-utility economies, a high γ also increases the equilibrium riskfree rate. For the parameters of Table 4, δ becomes negative, unless γ is kept low. This leads us back to the finding for log-utility: to support a low riskfree rate and a high equity premium, agents have to be excessively pessimistic in their perception of the worst-case equity premium.

It is noteworthy that the results above can be immediately recast as results for an economy with a stochastic differential utility representative agent (again, using the mapping that $\frac{1}{\gamma}$ corresponds to the elasticity of intertemporal substitution and $\gamma + \theta$ to the coefficient of relative risk-aversion). This is important because the pricing relations obtained by Epstein and Zin (1991) and Duffie and Epstein (1992b) are quite different: they prove that the risk premium on an asset is a linear combination of the CCAPM and the standard CAPM. This result has been critized because the return on wealth is itself an endogenous variable (Campbell, 1996 and 1998). I obtain a different result because I use the explicit portfolio weights and consider a single risky asset.

Finally, the failure of robustness to generate high equity premia and low riskfree rates does not imply that robustness is uninteresting for understanding asset prices. The lack of success is more likely due to the simple nature of the exchange economy analyzed here.

Another major caveat of the robust decision-making framework, is that agents do not engage in any type of learning. Admittedly, the omission of learning in this approach is quite fundamental. Incorporating learning, however, is a nontrivial challenge. A good starting point might be Gilboa and Schmeidler (1993).

5.4 Related Literature

Recent papers by Abel (1997) and Cecchetti, Lam and Mark (1997), although using very different techniques, address the equity premium and riskfree rate puzzle by deviating from rationality in a very similar spirit. Abel considers pessimism and doubt in a Lucas exchange economy. Pessimism makes the subjective distribution of dividend growth be first-order stochastically dominated by the objective distribution function, whereas doubt adds a mean-preserving spread to the objective distribution to obtain the subjective one. Abel demonstrates how pessimism and doubt reduce the riskfree rate. Under certain conditions (uniformity), both concepts also act to increase the equilibrium equity premium. Loosely speaking, robustness in my model is equivalent to Abel's pessimism. Doubt however, is absent here: as argued before, the tight structure imposed by the diffusion setting forces the agent's subjective perception of second moments to be correct. Abel shows how both deviations from rationality allow plausible parameters to explain observed asset prices. Given the results presented here, it is not surprising that the effect of doubt is stronger than the effect of pessimism. Unfortunately, it is not straightforward to interpret the amount of pessimism and doubt required for the calibration exercise.

Cecchetti, Lam and Mark (1997) analyze the pricing implications of subjective beliefs about (serially correlated) dividend growth that are distorted in a more complicated manner. Agents are assumed to believe that expansions and contractions exhibit less persistence than implied by the correct transition probabilities. Moreover, these beliefs vary stochastically. This is achieved by having agents follow an ad-hoc, but intuitive, rule-of-thumb when computing the transition probabilities. Cecchetti, Lam and Mark match both the equity premium and the riskfree rate using conventional preference parameter values. However, this requires that agents believe that consumption growth is negative and almost 50% more volatile than historically observed.

Hansen, Sargent and Tallarini (1999, henceforth HST) and Tallarini (1999) obtain results very similar to the ones presented here, albeit in seemingly different and discrete-time environments. HST develop a permanent income model populated by agents with habit persistence and risk-sensitized quadratic preferences. They derive a remarkable observational equivalence result: the quantity implications of their model are the same as in a model with agents that are not risk-sensitive, but more patient. Exactly as in (58), the intuition is that risk-sensitivity enhances the precautionary-savings motive. This can be offset by endowing agents with a higher rate of time preference.[14] Subsequently, HST demonstrate how their model can generate a stochastic discount factor that is

[14] HST focus on equilibrium savings rather than on the equilibrium riskfree rate, since the riskfree rate in their model is fixed by a linear production technology.

sufficiently volatile to obtain a high market price of risk.

Tallarini studies both endowment and stochastic growth economies, and increases risk aversion while fixing the elasticity of intertemporal substitution at unity. The resulting preferences can be interpreted as the discrete-time equivalent of the ones used here for log utility. Tallarini argues that a real business cycle model with this preference specification preserves the stochastic growth model's implications in terms of aggregate fluctuations, while dramatically improving its ability to match asset prices.

My paper can viewed as complementary to this work: HST and Tallarini demonstrate how risk-sensitivity (or robustness) combined with quadratic or log utility preserves the quantity and business cycle predictions of the model while improving its asset pricing implications. I show that, in addition, important insights from financial economics and portfolio choice theory are preserved, given that even logarithmic utility generates the type of hedging demands analyzed in the standard expected-utility models. I also present solutions for general power utility.

Anderson, Hansen and Sargent, who develop the tools that I use here, derive asset pricing results using the same approach as in HST. They assume the existence of a stochastic discount factor and characterize equilibrium asset prices in terms of the moments of this stochastic discount factor. My results complement AHS's by solving the investor's decision problem explicitly (both for constant and time-varying investment opportunity sets) and by presenting the resulting equilibrium asset prices analytically.

6 Conclusion

Parameter uncertainty, or even model uncertainty, seems pervasive in many aspects of financial decision-making. Even a parameter as crucial as the expected return on equities is the subject of major disagreement and dispute, on the basis of both theoretical and empirical considerations (see for instance Cochrane (1998) and Welch (1999)). This paper takes such uncertainty into account and explores its effects on both dynamic portfolio rules and equilibrium asset pricing. To incorporate a preference for robustness, I use the framework developed by Anderson, Hansen and Sargent (1999). Although partly motivated by tractability, this choice is very natural. Indeed, the type of model uncertainty that an agent worries about in their set-up concerns the drift of the diffusion process, precisely the expected return on equities in finance models. In particular, AHS derive a framework where an agent faces a dynamic decision problem. She uses a particular model as guidance for this decision problem, but suspects that some aspect of this reference model is misspecified. In response, the decision-maker considers alternative models that are relatively similar or close (in a precise sense). How much these alternative models are allowed to differ from the reference model depends on a parameter θ that measures the strength of the preference for robustness ($\theta = 0$ nests standard expected utility). Robust decisions are then designed to insure against the worst-case misspecification or alternative model. Such pessimism can be motivated using the decision-theoretic literature on maxmin expected utility (initiated by Wald (1950) and axiomatized by Gilboa and Schmeidler (1989)), which is closely connected to the neobayesian paradigm where a decision-maker operating in an information vacuum entertains an entire family of priors, rather than a single prior.

Modifying the framework of AHS to impose homotheticity of the preferences, I present analytical solutions for the optimal consumption and portfolio rules of a robust investor with power utility facing a constant investment opportunity set. A preference for robustness dramatically decreases the optimal share of the portfolio allocated to equities. I calculate the endogenous worst-case scenario for stock returns that supports this cautious behavior, and indicate how it can be used to select a reasonable value for θ, the parameter measuring the strength of the preference for robustness. This homothetic version of robustness is shown to be observationally equivalent to recursive preferences in the sense of Epstein-Zin or Weil. Essentially, robustness can be interpreted as increasing the investor's risk aversion (from γ to $\gamma+\theta$) without affecting her willingness to substitute intertemporally (which remains $\frac{1}{\gamma}$). Nonetheless, a subtle difference between recursive preferences and robustness persists: robustness, unlike recursive preferences, can be useful in reconciling relatively low degrees of risk aversion, based on introspection and experimental evidence, with the low risk tolerance inferred from asset prices. Introspection or experiments based on choice problems in environments with well-specified probabilities and events elicit the degree of risk aversion γ, because the preference for robustness is not operational. At the same time, asset prices in addition reflect a preference for robustness (θ), so that 'risk aversion' can be perceived to be $\gamma + \theta$.

Considering next the dynamic investment problem of an investor facing a time-varying investment opportunity set (due to a mean-reverting risk premium or due to stochastic volatility), I find that a preference for robustness gives rise to an additional hedging-type demand for the risky asset, even for logarithmic investors. Hedging therefore mitigates the reduction in the demand for equities found in the i.i.d setting when considering time-variation in expected excess returns (in which case hedging demands are positive), while it reinforces the downward pressure on the portfolio weight for stochastic volatility (as hedging demands are negative).

The cautiousness of the investment strategies chosen by robust investors in the partial-equilibrium setting gives some hope with respect to the ability of robustness to generate high equilibrium risk premia. To explore the strength of these effects, I consider a simple Lucas-style exchange economy. In principle, robustness can match the historical equity premium and riskfree rate with logarithmic utility. An interesting result is that the endogenous worst-case scenario for equity returns supporting this equilibrium is precisely the prediction of the expected-utility equilibrium model. In a way, a sophisticated robust investor is aware of the low equity premium generated by standard equilibrium models, and is therefore wary of the generosity of the historical equity premium. Guided by this suspicion, he invests cautiously. Of course, this drives up the equilibrium equity return. Simultaneously, increased precautionary savings due to the preference for robustness keep the equilibrium riskfree rate low. However, because the endogenous worst-case scenario for equity returns supporting the equilibrium coincides with the prediction of the expected-utility equilibrium model, an excessive amount of pessimism is required to obtain a high price of risk. Power utility demands a less pessimistic perspective, but only at the cost of increasing the riskfree rate, due to the reluctance to substitute intertemporally. Nonetheless, robustness increases the equilibrium price of risk and might be able to achieve more success when leaving the confines of the simple CRRA exchange economy.

A number of extensions and applications of the framework considered here seem promising. First, the equilibrium exercise reported here can be extended to the case of mean-reverting dividend (and consumption) processes. Most promising would be the extension of robustness to recursive preferences. In that case, one would be able to completely disentangle three distinct determinants of behavior, namely the elasticity of intertemporal substitution, the coefficient of relative risk aversion and the strength of the preference for robustness.

Second, it would be interesting to explore the equilibrium effects of heterogeneity in the preference for robustness along the lines of Anderson (1998), Dumas (1989) or Dumas, Uppal and Wang (1999).

A strong assumption in the work on robustness is that decision-makers do not engage in any type of learning. Relaxing this is challenging, but important. A useful starting point might be the theory in Gilboa and Schmeidler (1993).

International portfolio decisions naturally induce a concern for model uncertainty, stemming from for instance exchange rate regime uncertainty or deviations from Purchasing Power Parity. The effects of robustness on international portfolio allocation and diversification are investigated in Lustig and Maenhout (1999).

Finally, although I illustrate how to adapt the robust framework of AHS to a setting of pure jump processes (Appendix C), it would be interesting to extend these results to mixed jump-diffusions. Indeed, if model uncertainty is a concern, this is a particularly relevant environment. Possible applications of these findings might be of interest to the pricing of options and derivatives in the face of model uncertainty. Exploring the relationship between robust decision-making and the risk management and Value-at-Risk techniques used in practice (Boudoukh, Richardson and Whitelaw (1997)) might also be fruitful in this respect.

References

[1] Abel, A., 1997, "An Exploration of the Effects of Pessimism and Doubt on Asset Returns," unpublished paper, University of Pennsylvania.

[2] Anderson, E., 1998, "Uncertainty and the Dynamics of Pareto Optimal Allocations," unpublished paper, University of North Carolina.

[3] Anderson, E., L. Hansen, and T. Sargent, 1999, "Risk and Robustness in Equilibrium," unpublished paper, Stanford University.

[4] Balduzzi, P. and A. Lynch, 1999, "Transaction Costs and Predictability: Some Utility Costs Calculations," *Journal of Financial Economics* 52, 47-78.

[5] Barberis, N., 1999, "Investing for the Long Run When Returns are Predictable," forthcoming *Journal of Finance*.

[6] Bender, C. and S. Orszag, 1978, "*Advanced Mathematical Methods for Scientists and Engineers,*" McGraw Hill: New York.

[7] Blanchard, O., 1993, "Movements in the Equity Premium," *Brookings Papers on Economic Activity* 2, 75–118.

[8] Boudoukh, J., M. Richardson and R. Whitelaw, 1997, "Expect the Worst," *VAR: Understanding and Applying Value-at-Risk*, Risk Publications: London, 79-82.

[9] Brandt, M., 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," *Journal of Finance* 54, 1609-1646.

[10] Brennan, M., 1997, "The Role of Learning in Dynamic Portfolio Decisions," *European Finance Review* 1, 295-306.

[11] Brennan, M., E. Schwartz and R. Lagnado, 1997, "Strategic Asset Allocation," *Journal of Economic Dynamics and Control* 21, 1377-1403.

[12] Brennan, M. and Y. Xia, 1998, "Resolution of a Financial Puzzle," unpublished paper, UCLA.

[13] Brown, S., W. Goetzmann and S. Ross, 1995, "Survival," *Journal of Finance* 50, 853-873.

[14] Camerer, C., 1997, "Individual Decision Making," in J. Kagel and A. Roth eds., *The Handbook of Experimental Economics*, Princeton University Press, 587-703.

[15] Campbell, J. Y., 1996, "Understanding Risk and Return," *Journal of Political Economy* 104, 298–345.

[16] Campbell, J. Y., 1998, "Asset Prices, Consumption, and the Business Cycle," NBER Working Paper No. 6485, forthcoming in J. Taylor and M. Woodford eds., *Handbook of Macroeconomics*, North-Holland: Amsterdam, 1999.

[17] Campbell, J. Y. and R. Shiller, 1998, "Valuation Ratios and the Long-Run Stock Market Outlook," *Journal of Portfolio Management* 24, 11–26.

[18] Campbell, J.Y. and L. Viceira, 1998, "Who Should Buy Long-Term Bonds?," NBER Working Paper No. 6801.

[19] Campbell, J.Y. and L. Viceira, 1999, "Consumption and Portfolio Decisions When Expected Returns are Time Varying," *Quarterly Journal of Economics* 114, 433-495.

[20] Caplin, A. and J. Leahy, 1999, "Psychological Expected Utility Theory and Anticipatory Feelings," unpublished paper, New York University.

[21] Cecchetti, S., P. Lam and N. Mark, 1997, "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to be True?," unpublished paper, Ohio State University.

[22] Chacko, G. and L. Viceira, 1999, "Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets," unpublished paper, Harvard University.

[23] Cocco, J., F. Gomes, and P. Maenhout, 1999, "Consumption and Portfolio Choice over the Life-Cycle," unpublished paper, Harvard University.

[24] Cochrane, J., 1998, "Where is the Market Going? Uncertain Facts and Novel Theories," NBER Working Paper No. 6207.

[25] Cochrane, J. and L. Hansen, 1992, "Asset Pricing Explorations for Macroeconomics," in *NBER Macroeconomics Annual 1992*, MIT Press: Cambridge, 115-165.

[26] Cox, J. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics* 3, 145-166.

[27] Dow, J. and S. Werlang, 1992, "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica* 60, p. 197-204.

[28] Duffie, D. and L. Epstein, 1992a, "Stochastic Differential Utility," *Econometrica* 60, p. 353-394.

[29] Duffie, D. and L. Epstein, 1992b, "Asset Pricing with Stochastic Differential Utility," *Review of Financial Studies* 5, 411-436.

[30] Duffie, D. and P.-L. Lions, 1992, "PDE Solutions of Stochastic Differential Utility," *Journal of Mathematical Economics* 21, 577-606.

[31] Duffie, D., 1996, "*Dynamic Asset Pricing Theory*," Second Edition, Princeton University Press: Princeton.

[32] Dumas, B., 1989, "Two-Person Dynamic Equilibrium in the Capital Market," *Review of Financial Studies* 2, 157-188.

[33] Dumas, B., R. Uppal, and T. Wang, 1999, "Efficient Intertemporal Allocations with Recursive Utility," NBER Working Paper No. T231.

[34] Ellsberg, D., 1961, "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics* 25, 643-669.

[35] Epstein, L. and S. Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 937-69.

[36] Epstein, L. and S. Zin, 1991, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation," *Journal of Political Economy* 99, 263-286.

[37] Epstein, L. and T. Wang, 1994, "Intertemporal Asset Pricing Under Knightian Uncertainty," *Econometrica* 62, 283-322.

[38] Gennotte, G., 1986, "Optimal Portfolio Choice Under Incomplete Information," *Journal of Finance* 41, 733-746.

[39] Gilboa, I. and D. Schmeidler, 1989, "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics* 18, 141-153.

[40] Gilboa, I. and D. Schmeidler, 1993, "Updating Ambiguous Beliefs," *Journal of Economic Theory* 59, 33-49.

[41] Goetzmann, W. and P. Jorion, 1999, "Global Stock Markets in the Twentieth Century," *Journal of Finance* 54.

[42] Goyal, A. and I. Welch, 1999, "The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium?," unpublished paper, UCLA.

[43] Hansen, L. and T. Sargent, 1995, "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control* 40, 968-971.

[44] Hansen, L., T. Sargent, and T. Tallarini, 1999, "Robust Permanent Income and Pricing," forthcoming *Review of Economic Studies*.

[45] Heaton, J. and D. Lucas, 1999, "Stock Prices and Fundamentals," unpublished paper, Northwestern University.

[46] Judd, K., 1996, "*Numerical Methods in Economics*," MIT Press: Cambridge.

[47] Kim, T. S. and E. Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior," *Review of Financial Studies* 9, 141-161.

[48] Liu, J., 1999, "Portfolio Selection in Stochastic Environments," unpublished paper, Stanford University.

[49] Lucas, R. E., 1978, "Asset Prices in an Exchange Economy," *Econometrica* 46, 1426-1445.

[50] Lustig, H. and P. Maenhout, 1999, "Robustness and Home Bias," work in progress, Stanford University.

[51] Merton, R. C., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics* 51, 247–257.

[52] Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory* 3, 373–413.

[53] Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics* 8, 323–361.

[54] Pastor, L. and R. Stambaugh, 1999, "The Equity Premium and Structural Breaks," unpublished paper, University of Pennsylvania.

[55] Samuelson, P. A., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics* 51, 239–246.

[56] Samuelson, P. A., 1991, "Long-Run Risk Tolerance when Equity Returns are Mean Regressing: Pseudoparadoxes and Vindication of a 'Businessman's Risk'," in W. Brainard, W. Nordhaus and H. Watts eds., *Money, macroeconomics, and economic policy: Essays in honor of James Tobin*, MIT Press: Cambridge and London, 181-200.

[57] Schroder, M. and C. Skiadas, 1999, "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," forthcoming *Journal of Economic Theory.*

[58] Tallarini, T. Jr., 1999, "Risk Sensitive Real Business Cycles," forthcoming *Journal of Monetary Economics.*

[59] Wachter, J., 1998, "Portfolio and Consumption Decisions under Mean-Reverting Returns: an Exact Solution for Complete Markets," unpublished paper, Harvard University.

[60] Wald, A., 1950, "*Statistical Decision Functions,"* Wiley: New York.

[61] Weil, P., 1989, "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics* 24, 401-421.

[62] Weil, P., 1990, "Nonexpected Utility in Macroeconomics," *Quarterly Journal of Economics* 105, 29-42.

[63] Welch, I., 1999, "Views of Financial Economists On The Equity Premium And Other Issues," unpublished paper, UCLA.

Appendix A: Proofs

This appendix contains the proofs that were omitted from the main text.

Proof of Proposition 3.1: Immediate by plugging (26), (27) and (28) into (25). The boundary condition $J(W,T) = \frac{W^{1-\gamma}}{1-\gamma}$ is satisfied. Finally, nonnegativity of consumption requires $a > 0$. The proof for the limiting case of $\gamma = 1$ is given in Appendix B. ■

Proof of Proposition 3.2: It suffices to prove that the preferences of a homothetically robust agent with $U(C) = \frac{C^{1-\gamma}}{1-\gamma}$ and $\psi(J) = \frac{\theta}{(1-\gamma)J}$ are ordinally equivalent to those of an agent with stochastic differential utility with coefficient of relative risk aversion $\gamma+\theta$ and elasticity of intertemporal substitution $\frac{1}{\gamma}$. Using the terminology of Duffie and Epstein (1992a and 1992b), the aggregator and variance multiplier of the robust agent are:

$$\begin{aligned} f(c,v) &= \frac{c^{1-\gamma}}{1-\gamma} - \delta v, \\ A(v) &= \frac{-\theta}{(1-\gamma)v}. \end{aligned}$$

When $\gamma > 1$, both $(1-\gamma)$ and v are negative. Hence, $A(v) = \frac{-\theta}{(1-\gamma)v} = \frac{\theta}{(1-\gamma)|v|}$. Using the transformation $\varphi(v) = -|v|^{\frac{1-\gamma-\theta}{1-\gamma}}$, the normalized aggregator is then:

$$\overline{f}(c,v) = \frac{1-\gamma-\theta}{1-\gamma}\left[\frac{c^{1-\gamma}}{1-\gamma}|v|^{\frac{-\theta}{1-\gamma-\theta}} - \delta v\right].$$

Denoting, as in Duffie and Epstein (1992a and 1992b), the elasticity of intertemporal substitution by $\frac{1}{1-\rho}$ and the coefficient of relative risk aversion by $1-\alpha$, the mapping for the preference parameters given in proposition 3.2 implies that $\gamma = 1-\rho$ and $\theta = \frac{\rho-\alpha}{\rho}$. Thus, the normalized aggregator can be rewritten in terms of the parameters of Duffie and Epstein as[15]:

$$\overline{f}(c,v) = \frac{\alpha}{\rho}\left[\frac{c^{\rho}}{\rho}|v|^{\frac{\alpha-\rho}{\alpha}} - \delta v\right].$$

Finally, using the transformation $\varphi(v) = \frac{(\delta|\rho|)^{\alpha/\rho}}{|\alpha|}v$, we obtain:

$$\overline{f}(c,v) = \frac{\delta}{\rho}\left[\frac{c^{\rho} - (\alpha v)^{\rho/\alpha}}{(\alpha v)^{(\rho/\alpha)-1}}\right],$$

which is precisely the normalized aggregator for stochastic differential utility as defined in Duffie and Epstein.

For $\gamma < 1$ (and thus $v > 0$), the transformation required to obtain the normalized aggregator is $\varphi(v) = v^{\frac{1-\gamma-\theta}{1-\gamma}}$, which is strictly increasing if $\theta < 1-\gamma$,[16] and which yields:

$$\overline{f}(c,v) = \frac{1-\gamma-\theta}{1-\gamma}\left[\frac{c^{1-\gamma}}{1-\gamma}v^{\frac{-\theta}{1-\gamma-\theta}} - \delta v\right].$$

[15] This is also the parametrization used in Schroder and Skiadas (1999).

[16] Schroder and Skiadas impose the same restriction in their proofs of existence and optimality.

Again, using the mapping from the proposition, and the transformation $\varphi(v) = \frac{(\delta\rho)^{\alpha/\rho}}{\alpha} v$, one gets:

$$\overline{f}(c, v) = \frac{\delta}{\rho} \left[\frac{c^\rho - (\alpha v)^{\rho/\alpha}}{(\alpha v)^{(\rho/\alpha)-1}} \right],$$

and the same conclusion. ■

Proof of Proposition 4.1: Substituting the value function (38) into the first-order condition (37) immediately gives (39). Substitution of both the value function (38) and the optimal portfolio rule (39) into HJB equation (36) results in a quadratic equation in X. Collecting terms in X^2, X and the constant gives the following system of first-order nonlinear ordinary differential equations:

$$C'(t) = -\beta_0 - \beta_1 C(t) - \beta_2 (C(t))^2, \tag{60}$$

$$B'(t) = -\lambda \overline{X} C(t) - \frac{\beta_1}{2} B(t) - \beta_2 B(t) C(t),$$

$$A'(t) = -r - \lambda \overline{X} B(t) - \frac{\sigma_x^2}{2} C(t) - \frac{\beta_2}{2} (B(t))^2,$$

with constant coefficients:

$$\begin{aligned} \beta_0 &= \frac{1}{\gamma + \theta} > 0, \\ \beta_1 &= 2 \left[\rho \sigma_x \frac{1 - \gamma - \theta}{\gamma + \theta} - \lambda \right], \\ \beta_2 &= \sigma_x^2 (1 - \gamma - \theta) \left[1 + \frac{1 - \gamma - \theta}{\gamma + \theta} \rho^2 \right] < 0, \end{aligned}$$

and subject to:

$$C(T) = B(T) = A(T) = 0.$$

The conditions $C(T) = B(T) = A(T) = 0$, follow from the boundary condition of the HJB equation.

Towards a solution of the differential equations for $C(t)$ and $B(t)$, define $q \equiv \beta_1^2 - 4\beta_0\beta_2 > 0$ and $\eta \equiv \sqrt{q}$. As is shown for instance in Kim and Omberg (1996), the solutions for $C(t)$ and $B(T)$ are then respectively:

$$C(t) = \frac{2\beta_0 \left(1 - e^{-\eta(T-t)}\right)}{2\eta - (\beta_1 + \eta)\left(1 - e^{-\eta(T-t)}\right)},$$

$$B(t) = \frac{4\beta_0 \lambda \overline{X} \left[1 - e^{-\frac{\eta}{2}(T-t)}\right]^2}{\eta \left[2\eta - (\beta_1 + \eta)\left(1 - e^{-\eta(T-t)}\right)\right]}.$$

Proof of Lemma 4.1: To sign $C(t)$, it suffices to show that the denominator $2\eta - (\beta_1 + \eta)\left(1 - e^{-\eta(T-t)}\right)$ is strictly positive, given that $2\beta_0 \left(1 - e^{-\eta(T-t)}\right) > 0, \forall t < T$. Consider first the case where $\rho > 0$, so that $\beta_1 < 0$, and $\eta + \beta_1 > 0$ from the definition of η. The denominator is then positive because $2\eta - (\beta_1 + \eta)\left(1 - e^{-\eta(T-t)}\right) = \eta - \beta_1 + (\eta + \beta_1)e^{-\eta(T-t)}$. If instead $\rho < 0$, we can distinguish two cases. If β_1 is still negative, the previous argument applies. Otherwise, it suffices to show that $\eta > \beta_1$, which is again immediate from the definition of η.

The sign of $A'(t)$ is easily determined by simply computing the derivative of $A(t)$. Finally, signing $B(t)$ and $B'(t)$ can be done in a similar fashion. ■

Proof of Proposition 4.2: Analogous to the proof of Proposition 4.1, except that now a linear equation in the state variable obtains, rather than a quadratic equation. Collecting terms in y and the constant yields the following system of first-order ordinary differential equations:

$$B'(t) = -\beta_0 - \beta_1 B(t) - \beta_2 (B(t))^2, \tag{61}$$

$$A'(t) = -r - \kappa \overline{y} B(t),$$

with constant coefficients:

$$\begin{aligned}
\beta_0 &= \frac{1}{2}\frac{[\mu - r]^2}{\gamma + \theta} > 0, \\
\beta_1 &= \frac{1 - \gamma - \theta}{\gamma + \theta}(\mu - r)\rho\sigma - \kappa, \\
\beta_2 &= \frac{\sigma^2}{2}(1 - \gamma - \theta)\left[1 + \frac{1 - \gamma - \theta}{\gamma + \theta}\rho^2\right] < 0,
\end{aligned}$$

and subject to:

$$A(T) = B(T) = 0.$$

Defining

$$q \equiv \beta_1^2 - 4\beta_0\beta_2 > 0, \;\; \eta \equiv \sqrt{q},$$

the solution is

$$B(t) = \frac{2\beta_0\left(1 - e^{-\eta(T-t)}\right)}{2\eta - (\beta_1 + \eta)\left(1 - e^{-\eta(T-t)}\right)}.$$

Proof of Lemma 4.2: The proof is identical to the proof of Lemma 4.1.

Proof of Proposition 5.1: The equilibrium can be constructed as follows. Note first that $S_t = \frac{1}{a} D_t$ implies immediately that $\frac{dS_t}{S_t} = \mu_D dt + \sigma_D dB_t$. Substituting the optimality conditions (53) and (54) into the stochastic differential equation for wealth (51) gives:

$$\frac{dW_t}{W_t} = \left[\frac{r}{\gamma} - \frac{\delta}{\gamma} + \frac{1 + \gamma}{2\gamma[\gamma + \theta]}\left(\frac{\mu_S - r}{\sigma_S}\right)^2\right] dt + \frac{1}{\gamma + \theta}\frac{\mu_S - r}{\sigma_S} dB_t.$$

Using the CCAPM result from (56):

$$\frac{dW_t}{W_t} = \left[\frac{r}{\gamma} - \frac{\delta}{\gamma} + \frac{1}{2\gamma}[1 + \gamma][\gamma + \theta]\sigma_S^2\right] dt + \sigma_S dB_t. \tag{62}$$

Also, market clearing in the goods market implies that:

$$C_t = D_t = aW_t.$$

Moreover $S_t = \frac{1}{a}D_t$, so that in equilibrium $S_t = W_t$. Combining $\frac{dS_t}{S_t} = \mu_D dt + \sigma_D dB_t$ with the implied wealth dynamics in (62) yields the following equilibrium condition:

$$S_0 \exp\left\{\left(\mu_D - \frac{1}{2}\sigma_D^2\right)t + \sigma_D \int_0^t dB_s\right\} = W_0 \exp\left\{\left(\frac{r}{\gamma} - \frac{\delta}{\gamma} + \frac{1}{2\gamma}\left[1+\gamma\right]\left[\gamma+\theta\right]\sigma_S^2 - \frac{1}{2}\sigma_S^2\right)t + \sigma_S \int_0^t dB_s\right\}.$$

This results in:

$$\begin{aligned} S_0 &= W_0, \\ \mu_D - \frac{1}{2}\sigma_D^2 &= \frac{r}{\gamma} - \frac{\delta}{\gamma} + \frac{1}{2\gamma}\left[1+\gamma\right]\left[\gamma+\theta\right]\sigma_S^2 - \frac{1}{2}\sigma_S^2, \\ \sigma_D &= \sigma_S. \end{aligned}$$

Rearranging the last two equations immediately produces the desired results in (57) and (58). Finally, it is straightforward to verify that the optimal consumption and investment strategies, (53) and (54) respectively, are consistent with this equilibrium. Similarly, the transversality condition is satisfied as long as $a > 0$, which holds by assumption. ■

Appendix B: Logarithmic Utility

The main idea of this appendix is to show that the results presented in the paper for $\gamma \neq 1$ nest logarithmic utility. To do so, it is crucial to scale θ appropriately even for log utility, despite the fact that logarithmic utility leads to homothetic preferences even for $\psi(J) = \theta$. This is important given that the results presented for log stochastic differential utility in Schroder and Skiadas (1999) do not correspond to the limits of their results for general SDU as the elasticity of intertemporal substitution tends to unity (their solution for the portfolio rule α exhibits a horizon-effect for constant investment opportunity sets when $\gamma = 1$, while the solution for $\gamma \neq 1$ (but arbitrarily close) does not). The same discontinuity at $\gamma = 1$ would obtain here if I used $\psi(J) = \theta$. The source of this horizon effect can be understood from equation (19): when θ is constant, the worst-case misspecification is proportional to the marginal utility of wealth, which shrinks as the horizon shortens. As the worst-case becomes less adverse, the agent invests a higher fraction in equities as the terminal date nears. Therefore, while scaling θ is necessary for general power utility lest robustness wears off as wealth increases, the same holds for log utility as the effect of robustness vanishes as t increases, unless one scales appropriately.

For those reasons, I use $\psi_{\log}(J) = \lim_{\gamma \to 1} \psi(J)$, where the limit is computed using the explicit solution for J.

Proof of Proposition 3.1 for $\gamma = 1$: Based on the solution given in Proposition 3.1 for J when $\gamma \neq 1$, $\psi_{\log}(J) = \lim_{\gamma \to 1} \psi(J) = \frac{\delta\theta}{1-e^{-\delta(T-t)}}$. The appropriate limit of HJB equation (25) for $\gamma \to 1$ is then:

$$0 = \sup_{\alpha,C} \left[\log(C) - \delta J + J_w \left[W \left(r + \alpha(\mu - r) \right) - C \right] + J_t + \frac{1}{2} \left(J_{ww} - \frac{\delta\theta}{1 - e^{-\delta(T-t)}} J_w^2 \right) \alpha^2 \sigma^2 W^2 \right],$$

subject to (18). It is straightforward to verify that this PDE is solved by $J(W,t) = \frac{1-e^{-\delta(T-t)}}{\delta} \log W + \phi(t)$, where $\phi(t)$ solves

$$\phi'(t) = \delta\phi(t) - \log\delta + \log\left[1 - e^{-\delta(T-t)}\right] - \frac{1 - e^{-\delta(T-t)}}{\delta} \left[r + \frac{1}{2\left[1+\theta\right]} \left(\frac{\mu - r}{\sigma} \right)^2 \right],$$

subject to $\phi(T) = 0$. The resulting optimal consumption and portfolio rules are:

$$C^* = \frac{\delta}{1 - e^{-\delta(T-t)}} W,$$

$$\alpha^* = \frac{1}{1+\theta} \frac{\mu - r}{\sigma^2},$$

i.e. the special cases of (27) and (28) for $\gamma = 1$. The portfolio rule contrasts with the solution that would have obtained when using $\psi_{\log}(J) = \theta$, namely $\alpha^* = \frac{\delta}{\delta + \theta\left(1 - e^{-\delta(T-t)}\right)} \frac{\mu - r}{\sigma^2}$ as in Schroder and Skiadas. ■

Proof of Proposition 4.1 for $\gamma = 1$: Based on the solution given in Proposition 4.1 for J when $\gamma \neq 1$, $\psi_{\log}(J) = \lim_{\gamma \to 1} \frac{\theta}{W^{1-\gamma} \exp\left[(1-\gamma)\left(A(t) + B(t)X + \frac{C(t)}{2} X^2\right)\right]} = \theta$. The HJB (36) subject to $J(W,X,T) = \log W$ is then solved by $J(W,t) = \log W + A(t) + B(t)X + \frac{C(t)}{2} X^2$, where $A(t)$, $B(t)$ and $C(t)$ solve system (60) with coefficients:

$$\begin{aligned} \beta_0 &= \frac{1}{1+\theta} > 0, \\ \beta_1 &= -2\left[\rho\sigma_x \frac{\theta}{\gamma+\theta} + \lambda\right], \\ \beta_2 &= -\theta\sigma_x^2 \left[1 - \frac{\theta}{\gamma+\theta}\rho^2\right] < 0, \end{aligned}$$

and subject to:

$$C(T) = B(T) = A(T) = 0.$$

This system is precisely (60) specialized to $\gamma = 1$. Hence, all results obtained before for $\gamma \neq 1$ apply for $\gamma = 1$ as well. ■

Proof of Proposition 4.2 for $\gamma = 1$: Identical to the proof of proposition 4.1 for $\gamma = 1$.

Appendix C: Robustness and Jump Processes

AHS derive their main results under the assumption of a diffusion process for the state variable. It is possible however, to translate their methodology to a setting of pure jump processes. For illustrative purposes, I focus on the simplest possible jump process, along the lines of the first example considered in Merton (1971). This might be of interest in the context of defaultable fixed-income securities or of option pricing (Cox and Ross (1976)).

Given a state variable x, let the event that x drops to ξx, $\xi \in (0,1)$, be an independent Poisson process (denoted by $q(t)$) with fixed intensity λ. The dynamics of x are assumed to follow:

$$dx_t = \mu x_t dt - (1-\xi) x_t dq_t. \tag{63}$$

For example, x_t could be the price of a defaultable bond, with instantaneous rate of interest μ, which in the event of default loses a fraction $(1-\xi)$ of its value. The infinitesimal generator of this process is:

$$\mathcal{D}f = \mu \frac{\partial f}{\partial x} + \lambda \left[f(\xi x) - f(x) \right]. \tag{64}$$

From the definition in (3), the twisted generator, $\mathcal{D}^w f$, corresponding to candidate model w can then be written as:

$$\begin{aligned} \mathcal{D}^w f &= \mu \frac{\partial f}{\partial x} + \lambda \frac{w(\xi x)}{w(x)} \left[f(\xi x) - f(x) \right] \\ &= \mathcal{D}f + \lambda \frac{w(\xi x) - w(x)}{w(x)} \left[f(\xi x) - f(x) \right]. \end{aligned} \tag{65}$$

Interestingly, the robust agent acts as if the jump intensity were $\lambda \frac{w(\xi x)}{w(x)}$, rather than λ as in (64). This corresponds to the result obtained by AHS that a robust agent in a diffusion setting acts as if the drift of the state equation is distorted (equation (13)). This comes at a cost to the decision-maker. The penalty for considering alternative specifications is given by their relative entropy:

$$\begin{aligned} I'(w) &= \lambda \left[\frac{w(\xi x)}{w(x)} \left[\log \frac{w(\xi x)}{w(x)} - 1 \right] + 1 \right] \\ &= \lambda \left[u \left[\log u - 1 \right] + 1 \right], \end{aligned} \tag{66}$$

where $u \equiv \frac{w(\xi x)}{w(x)}$.

Ignoring for simplicity the choice of a control variable, the HJB for a robust decision-maker facing a pure jump process is then (using (65) and (66)):

$$0 = \inf_u \left[U(x) - \delta V(x) + \mu \frac{\partial V(x)}{\partial x} + \lambda u \left[V(\xi x) - V(x) \right] + \frac{\lambda}{\theta} \left[u \left[\log u - 1 \right] + 1 \right] \right].$$

The FOC for u is:

$$u^* = \exp \left[\theta \left(V(x) - V(\xi x) \right) \right] > 1.$$

This gives more insight into the kind of distortion that the robust agent applies to the jump intensity. The agent acts as if the jump were more likely to occur ($\lambda u^* > \lambda$), depending on the strength of the preference for robustness θ, and depending on the relative severity of the outcome when the jump occurs.