Robust Portfolio Rules and Asset Pricing

Pascal J. Maenhout

INSEAD

*I am indebted to John Campbell, David Laibson and Andrew Metrick for guidance and suggestions during my Ph.D. at Harvard University. An anonymous referee and John Heaton (the editor) provided very useful comments and suggestions. I would also like to thank Andrew Ang, Nick Barberis, Gary Chamberlain, Bernard Dumas, Larry Epstein, Lars Hansen, Hanno Lustig, Thomas Sargent, Michael Schwarz, Costis Skiadas, Lucie Teplá, Aaron Tornell, Raman Uppal, Luis Viceira, and seminar participants at Harvard, Columbia Business School, INSEAD, London Business School, Kellogg, Chicago Business School, NYU Stern, Stanford Business School, the Wharton School, IIEC, the University of Cyprus, the 2000 WFA Meetings, the 2000 SED Meetings and the 2000 EFA Meetings for their comments. I gratefully acknowledge the financial support of the Fund for Scientific Research Flanders. Address correspondence to Pascal Maenhout, Finance Department, INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France. Email: pascal.maenhout@insead.edu. Tel. (33) 1 60 72 44 83.
Robust Portfolio Rules and Asset Pricing

Abstract: I present a new approach to the dynamic portfolio and consumption problem of an investor who worries about model uncertainty (in addition to market risk) and seeks robust decisions along the lines of Anderson, Hansen and Sargent (2002). In accordance with maxmin expected utility, a robust investor insures against some endogenous worst-case. I first show that robustness dramatically decreases the demand for equities and is observationally equivalent to recursive preferences when removing wealth effects. Unlike standard recursive preferences however, robustness leads to environment-specific ‘effective’ risk aversion. As an extension, I present a closed-form solution for the portfolio problem of a robust Duffie-Epstein-Zin investor. Finally, robustness increases the equilibrium equity premium and lowers the riskfree rate. Reasonable parameters generate a 4 to 6% equity premium.
A very active line of research in financial economics focuses on the question of how to optimally allocate a portfolio between a risky and a riskless asset in a dynamic environment. Most of this research was initiated by Merton’s and Samuelson’s provocative result that the optimal portfolio share in equities of an investor with power utility facing i.i.d. stock returns is constant over the life-cycle under the assumption of frictionless markets and the absence of labor income (Merton (1969) and 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. A large literature has been devoted to relaxing the assumptions of Merton and Samuelson to rationalize these recommendations qualitatively.

A fundamental assumption in the bulk of this work on dynamic portfolio choice is the absence of any uncertainty about the return process. Typically, one obtains point estimates for the asset return parameters and subsequently assumes these are known and fixed. With respect to second moments, one can argue that the limit of infinitely fine sampling would remove all estimation risk. First moments however are notoriously hard to estimate (Merton (1980), Cochrane (1998) and Blanchard (1993)). Most recently, some authors examine arguments that could support lower expected returns in the future (e.g. equilibrium implications of the recent increase in stock market participation (Heaton and Lucas (1999)), or historically low dividend yields (Campbell and Shiller (1998))). The debate is still open and reflects the lack of consensus concerning the expected risk premium (Cochrane (1998)), or even concerning the model generating excess returns (see for instance Pastor and Stambaugh (2001) for an analysis of structural breaks). Welch (2000) 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 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 uncertainty about the return process into account when studying optimal dynamic portfolio decisions. In this paper, I derive consumption and portfolio rules that are robust to a particular type of model misspecification, stemming from uncertainty about the return process. Robustness of decision rules to model misspecification has been studied extensively in applied mathematics and engineering. The main goal is to design decision 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. This literature is based on the key assumption that the decision-maker worries about some worst-case scenario. Admittedly, this constitutes a major deviation from standard expected-utility maximization, but it does not imply 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.

Robustness was pioneered in economics by Hansen and Sargent (1995). In this paper I use the continuous-time methodology developed in Anderson, Hansen and Sargent (2002, henceforth AHS). Their framework is natural for portfolio choice problems involving uncertainty about the return process for equities.

As a first contribution I propose an important modification of their methodology,
namely homothetic robustness, that preserves wealth independence and analytical tractability. All results are then obtained in closed-form. Robustness drastically reduces the demand for the risky asset: reasonable parameter values lower the optimal equity demand by 50%. The endogenous worst-case scenario that the agent hedges against is easily computed analytically and provides guidance in the choice of a reasonable value for the uncertainty aversion parameter. In addition, I show that imposing homotheticity makes robustness observationally equivalent to stochastic differential utility (Duffie and Epstein (1992a)). Robustness can then be interpreted as increasing risk aversion without changing the willingness to substitute intertemporally. An additional theoretical contribution of this paper is a closed-form solution for the case of a Duffie-Epstein investor with robustness. This is of particular interest, as it distinguishes between three distinct behavioral parameters: risk aversion, intertemporal substitution and uncertainty aversion. Importantly, even though observationally equivalent in terms of portfolio choice, the robust preference structure generates environment-specific (effective) risk aversion, unlike stochastic differential utility. The analysis can be interpreted as formalizing the intuition that investors are very conservative or pessimistic when forming portfolios and suggests that at least part of this pessimism might be due to robustness or uncertainty aversion rather than risk aversion. This allows me to explain cautious portfolios and high equity premia while maintaining risk aversion levels that are reasonable and in line with the low estimates of risk aversion obtained through experiments or introspection. In fact, a key prediction of the model is precisely that ‘risk aversion’ estimates based on asset prices are substantially higher than the estimates of ‘pure’ risk aversion based on stylized (thought) experiments.
In a Lucas-style equilibrium asset pricing model, robustness simultaneously increases the equilibrium equity premium and lowers the riskfree rate. The equity premium puzzle requires high uncertainty and risk aversion. Breeden’s CCAPM shows up as the equilibrium worst-case scenario for the equity premium and is driven by risk aversion only. Robustness drives a wedge between this worst-case and the actual risk premium. Empirically, a 3 to 5% wedge is difficult to detect given the usual length of available time-series. Given plausible values of risk aversion and uncertainty aversion, an equilibrium equity premium between 4% and 6% can then be sustained.

**Related Literature**

Hansen, Sargent and Tallarini (1999) develop a discrete-time permanent income model for quadratic agents with habit formation who fear model misspecification and seek robust decisions. They derive a remarkable observational equivalence result: the quantity implications of their model are the same as in a model without robustness, but where consumers are more patient. Their model can generate a volatile stochastic discount factor and therefore a high market price of risk. Tallarini (1999) risk-sensitizes logarithmic utility and increases risk aversion while fixing the elasticity of intertemporal substitution at unity. He demonstrates 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. AHS, who develop the tools that I use here, characterize equilibrium asset prices in terms of the moments of the stochastic discount factor. My results complement these papers by solving the investor’s problem explicitly and by obtaining equilibrium asset prices analytically for the more general case of
stochastic differential utility.

Cagetti et al. (2002) extend the methodology in AHS to leave the realm of pure diffusions and to consider mixed jump-diffusions. In addition they model robustness in the context of a filtering problem, which results in interesting endogenous time-variation in asset prices.

This paper is also related to the work in financial economics on Knightian uncertainty and ambiguity. An early application of maxmin expected utility to portfolio choice in a static setting can be found in Krasker (1982). Dow and Werlang (1992) study the static portfolio choice problem of an investor with expected utility under a nonadditive probability measure (or Choquet capacity), an alternative formalization of Knightian uncertainty that is closely related to maxmin expected utility. Epstein and Wang (1994) rigorously analyze the asset pricing implications of maxmin expected utility in a discrete-time infinite-horizon model, focusing mainly on price indeterminacy and volatility. Liu (1998) captures Knightian uncertainty using \( \varepsilon \)-contamination beliefs in discrete time and studies its effect on savings and portfolio choice. In a two-period model Knightian uncertainty is shown to reduce stock market participation. Most recently, Chen and Epstein (2002) provide a rigorous continuous-time extension of the multiple-prior framework of Gilboa and Schmeidler (1989) and of Epstein and Wang (1994).

Finally, the work of Abel (2002) and Cecchetti, Lam and Mark (2000) addresses the equity premium and riskfree rate puzzle by exogenously distorting subjective beliefs.

The organization of the paper is as follows. Section 1 presents the basic portfolio choice problem with robustness. Section 2 shows its solution based on the homothetic extension of
AHS. This establishes the isomorphy between robustness and recursive preferences. Section 3 analyzes the equilibrium model and presents the calibration exercise. Finally, section 4 concludes. Appendices A and B contain the proofs that are omitted from the main text.

1 The Basic Portfolio Problem

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

$$E \left[ \int_0^T \exp(-\delta t) \frac{C_t^{1-\gamma} \gamma}{1-\gamma} dt \right],$$

where $\gamma$ denotes the coefficient of relative risk aversion and $\delta > 0$ is the discount factor. The limit of $\frac{C_t^{1-\gamma}}{1-\gamma}$ as $\gamma$ tends to unity is defined as $\log(C_t)$.

The price of the risky asset evolves according to the standard geometric Brownian motion with constant coefficients driven by a standard Wiener process $B_t$:

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{1}$$

Therefore the state equation for wealth is:

$$dW_t = [W_t (r + \alpha_t (\mu - r)) - C_t] dt + \alpha_t \sigma W_t dB_t, \tag{2}$$

where $\alpha_t$ 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 $C_t$. Both controls are nonanticipating and suitably adapted to the $\sigma$-algebra generated by the underlying Brownian motion.
Denoting the value function by \( V(W, t) \) and its partial derivatives with respect to \( x \) by \( V_x \), Merton (1971) showed that the Hamilton-Jacobi-Bellman equation (henceforth HJB) for optimality is:

\[
0 = \sup_{\alpha, C} \left[ \frac{C^{1-\gamma}}{1-\gamma} - \delta V(W, t) + \mathcal{D}^{(\alpha, C)} V(W, t) \right],
\]

where

\[
\mathcal{D}^{(\alpha, C)} V(W, t) = \mathcal{W} \left( W (r + \alpha (\mu - r)) - C \right) + V_t + \frac{1}{2} W \mathcal{W} \sigma^2 W^2,
\]

with boundary condition:

\[
V(W, T) = 0.
\]

### 1.1 Minimum-Entropy Robustness

The object \( \mathcal{D} V \) in (3), the infinitesimal generator applied to the value function, plays a crucial role in the work of AHS and is essentially the mechanism through which robustness is introduced. Heuristically, \( \mathcal{D} V \) can be thought of as \( \frac{1}{dt} E(dV) \) and is easily obtained using Ito’s Lemma. A key insight of AHS is that this (differential) expectations operator, used to compute the (differential) continuation payoff in the Bellman equation, reflects a particular underlying model for the state variable \( W_t \). The decision-maker accepts this ‘reference model’ as useful, but suspects it to be misspecified. She therefore wants to consider alternative models when computing her continuation payoff. Loosely speaking, a preference for robustness is then achieved by having the agent guard against an adverse alternative model that is reasonably similar to the reference model.

In a pure diffusion setting, AHS show that this adverse alternative model simply adds
an endogenous drift $u(W_t)$ to the law of motion of the state variable $W_t$:

$$dW_t = \mu(W_t)dt + \sigma(W_t)[\sigma(W_t)u(W_t)dt + dB_t].$$

(6)

where $\mu(W_t)$ and $\sigma(W_t)$ are short-hand notation for the drift and diffusion in (2). The drift adjustment $u(W_t)$ is chosen endogenously to minimize the sum of the expected (differential) continuation payoff (4), but adjusted to reflect the additional drift component in (6), and of an entropy penalty:

$$\inf_{u} \mathcal{D}V + u(W_t)\sigma(W_t)^2V_w + \frac{1}{2\theta}u(W_t)^2\sigma(W_t)^2$$

(7)

The first two terms in the objective are the expected continuation payoff when the state variable follows (6), i.e. the alternative model based on drift distortion $u$. Of course, not any alternative model goes, and in the infinimization problem a penalty is incurred for alternative models that are too far away from the reference model. This notion of distance is formally measured by (the derivative of) relative entropy, the third term in (7). Intuitively, entropy can be thought of as a log-likelihood ratio. Alternative models with low entropy are statistically hard to differentiate from the reference model for the decision maker and hence worth considering. Technically, the use of entropy imposes that alternative models have to be absolutely continuous with respect to the reference model, so that the entropy measure (the expected log Radon-Nikodym derivative) exists. In the current diffusion setting absolute continuity focuses attention on the subclass of alternative models that only differ in terms of drift function (by Girsanov’s Theorem), so that general model uncertainty is reduced to uncertainty about the drift function of the state variable. 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 substantial uncertainty about the expected return process, i.e. precisely the drift in equation (6).4

The entropy penalty incurred when selecting adverse drift distortions in (7) and moving away from the reference model, is weighted by \( \frac{1}{\theta} \). The parameter \( \hat{\theta} \geq 0 \) measures the strength of the preference for robustness (\( \hat{\theta} = 0 \) corresponds to expected-utility maximization). Therefore the more robust decision-maker (\( \hat{\theta} \) larger) has less faith in the reference model and will consider larger drift distortions (alternative models with larger entropy) when evaluating her continuation payoff. It is important to point out that AHS model the parameter \( \hat{\theta} \) as fixed and state-independent. However, a modification that will be crucial for the remainder of this paper replaces \( \hat{\theta} \) by a state-dependent version of \( \theta \), denoted by \( \Psi(W,t) > 0 \). As will be shown in the next section, this modification is instrumental in assuring the homotheticity or scale-invariance of the decision problem and has therefore a natural economic justification.

Applying these results to the Merton portfolio problem:

\[
0 = \sup_{\alpha,C} \inf_u \left[ \frac{C_{1-\gamma}}{1-\gamma} - \delta V(W,t) + D^{(\alpha,C)} V(W,t) + V_w \alpha^2 \sigma^2 W^2 u + \frac{1}{2\Psi(W,t)} \alpha^2 \sigma^2 W^2 u^2 \right],
\]

where \( D^{(\alpha,C)} V(W,t) \) is given by (4), subject to (5). Solving first for the infimization part of the problem yields:

\[
u^* = -\Psi V_w.
\]

If the agent desires no robustness or has complete faith in the validity of the model (\( \Psi = 0 \)), then \( u^* = 0 \), i.e. there are no perturbations to guard against. Given that \( \Psi > 0 \) in the robust case, the perturbation amounts to adding a negative drift to the state equation if 11
Substituting for $u^*$ in the HJB equation gives:

$$0 = \sup_{\alpha,C} \left[ \frac{C_t^{1-\gamma}}{1-\gamma} - \delta V + \mathcal{D}^{(\alpha,C)} V - \frac{\Psi}{2} V_w^2 \sigma^2 W^2 \right],$$

subject to (5). Then the necessary optimality conditions for consumption and portfolio choice are:

$$(C^*)^{-\gamma} = V_w,$$  \hspace{1cm} (11)

$$\alpha^* = \frac{-V_w}{[V_{ww} - \Psi V_w^2] W} \frac{\mu - \tau}{\sigma^2}. \hspace{1cm} (12)$$

It is interesting to note that the form of the optimality condition for consumption is not affected by the introduction of robustness. This is not the case for the portfolio rule. Setting $\Psi = 0$ yields the familiar result that the optimal portfolio rule is just the myopic or mean-variance efficient $\alpha = \frac{\mu - \tau}{\sigma^2}$, adjusted for risk aversion $-\frac{V_w W}{V_w}$. Robustness adds an extra term to this ‘risk aversion’ adjustment, equal to $\Psi V_w W > 0$. 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 (11) and (12) back into the PDE (10) and to try to solve for $V$. However for the problem at hand, there exists no analytical solution when $\Psi = \tilde{\theta}$, as formulated in AHS. Although numerical methods could in principle be used (based on natural perturbation methods as in for instance Judd (1996) and Kogan and Uppal (2000)), I will instead demonstrate in the next section how the investment problem can be solved analytically. Essentially, I propose a suitable choice of $\Psi (W,t)$ or state-dependent version of $\tilde{\theta}$ that yields both homotheticity and analytical tractability.
1.2 Decision-theoretic Background

In the framework described above the investor takes a very conservative and pessimistic perspective and worries about some worst-case misspecification. However, standard decision theory, ubiquitous in economics and finance, typically insists on expected-utility maximization and Bayesian decision-making.

First, it can be noted that Bayesian approaches are typically limited to parametric versions of model uncertainty. Here, even when imposing the tight diffusion structure, the uncertainty concerns the entire drift function.

Because of the infimization in (8), the preferences of a robust investor can be interpreted as a form of maxmin expected utility (Gilboa and Schmeidler (1989)), where a decision-maker operating in an information vacuum (e.g. Ellsberg’s (1961) urn of unknown composition) entertains a family of priors rather than a single prior and, being uncertainty-averse, computes expected utility under the worst-case prior. In recent work, Hansen et al. (2002) rigorously establish the connection between the multiplier formulation of robustness (as used here) and the constraint formulation. The latter involves a minimization over alternative models subject to a formal entropy constraint and is therefore directly related to the Gilboa-Schmeidler multiple-prior set-up. An alternative treatment that rigorously introduces the set of multiple priors can be found in Epstein and Wang (1994) and Chen and Epstein (2002). While not formally modeling the set of priors, the approach taken here allows me to explicitly calculate the endogenous worst-case scenario that supports the investor’s decisions (from (9)). This means one can view $\Psi$ (and thus $\theta$) as indexing the set of priors that the investor entertains and calibrate accordingly.
Ellsberg-style experiments have been designed to elicit the degree of uncertainty aversion (see Camerer (1997)). One could use these to calibrate $\theta$. It is questionable however how informative this would be for applications of maxmin expected utility to portfolio choice and asset pricing, as financial markets are presumably much more complex environments than stylized experiments concerning Ellsberg urns. In fact, it is fundamental to realize that the $\theta$ parameter is environment-specific and not transferable from one setting to another. I will therefore calibrate $\theta$ by examining the implied least-favorable prior, using (9).

Perhaps most importantly, the robust approach results in a model that captures what may be essential to explaining the equity premium puzzle, namely that investors can be perceived to be more risk averse when acting in financial markets than when participating in experiments or introspective analysis.

2 Homothetic Robust Decision-Making

The preferences induced by the robust adjustment for a power-utility investor are not homothetic when the preference parameter governing uncertainty-aversion is fixed and state-independent as in AHS ($\Psi (W, t) = \hat{\theta}$). The simplest way to see this is to go back to the general FOC’s, (11) and (12). For the infinite-horizon case, a constant consumption-wealth ratio requires that $V(W) = \kappa W^{1-\gamma}$, where $\kappa$ is some negative constant (for $\gamma > 1$). Substituting this into the optimality condition for the portfolio weight results in $\alpha^* = \frac{(\mu - r)/\sigma^2}{\gamma + \theta(1-\gamma)\kappa W^{1-\gamma}}$. This portfolio weight is not independent of wealth. In particular, it increases in wealth and asymptotes to $\frac{\mu - r}{\gamma \sigma^2}$, the solution for expected utility. Robustness wears off as the state variable increases, at least for general CRRA preferences.
Another way to see this is to impose $\Psi(W, t) = \hat{\theta} > 0$ in the HJB equation for a robust investor, (10). The preference for robustness stems from the last term, $\hat{\theta}V_w^2$. 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 $V(W) = \kappa W^{1-\gamma}$, then $\hat{\theta}V_w^2 = o(W^{-2\gamma-1})$, while $V_{ww} = o(W^{1-\gamma})$. As a consequence, robustness proportional to a constant parameter $\hat{\theta}$ 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, 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 $V_w^2$ of the same asymptotic order as $V_{ww}$.

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.

2.1 Explicit Solutions

In order to preserve the homotheticity of the preferences, the above analysis suggests adjusting the set-up of AHS in the following way. Focusing on power utility with $\gamma \neq 1$, I impose homotheticity by making the parameter $\hat{\theta}$ state-dependent. In particular, I propose to scale the entropy-penalty parameter by a function of wealth (the relevant state variable) that is nonlinear in a way that reflects the homotheticity of the initial problem, i.e. without robustness. To motivate this, I (informally) invoke the analysis of Hansen et al. (2002) who
formally relate the multiplier formulation to the constraint formulation. The parameter $\theta$ in the former is then linked to the Lagrange multiplier on the entropy constraint in the latter formulation. Rewriting the multiplier problem as $\inf_{u} \mathcal{D}^{u}V(W) + \frac{1}{\pi} I'(u)$ where $\mathcal{D}^{u}V(W)$ represents the first two terms in (7), i.e. the expected continuation payoff according to the alternative model based on drift distortion $u$, and $I'(u)$ is the relative entropy of the alternative model, the constraint formulation is $\inf_{u} \mathcal{D}^{u}V(W)$ subject to $I'(u) \leq \eta$, where $\eta$ is a parameter indicating the tolerated entropy deviation for alternative models. When $V(W)$ is of the form $\kappa W^{1-\gamma}$ and hence homothetic in $W$, we know that $V(bW) = b^{1-\gamma}V(W)$ for a strictly positive scalar $b$. Homotheticity in the robust problem requires that the solution to the constraint problem is not affected by the presence of $b$. This will indeed be the case through an endogenous rescaling of the Lagrange multiplier associated with the entropy constraint: if $\lambda$ is the multiplier when $b = 1$, then $\lambda_{b \neq 1} = b^{1-\gamma}\lambda$. Moving back to the multiplier problem (7) in an isoelastic setting, this suggests that one can obtain homotheticity by multiplying the entropy penalty by $\kappa W^{1-\gamma}$ or, equivalently, by dividing $\theta$ by $\kappa W^{1-\gamma}$. Problem (7) then becomes invariant to the scale of wealth (i.e. to $b$), just like the constraint problem is by virtue of the endogenous Lagrange multiplier.

Notice that $\mathcal{D}^{u}(W)$ happens to be of the same functional form as the required scaling function $\kappa W^{1-\gamma}$. Therefore, $V(W)$ is a possible candidate to scale the entropy penalty. It turns out that this choice leads to particularly appealing analytical results.\(^8\)

I therefore use $\Psi(W, t)$, which scales $\theta$ by the value function:\(^9\)

$$
\Psi(W, t) = \frac{\theta}{(1-\gamma) V(W, t)} > 0. 
$$

(13)

This imposes the desired homotheticity property as robustness will no longer wear off as
wealth rises. Again, although scaling by some alternative function of wealth could work as well (if it has the form $\kappa W^{1-\gamma}$), the particular form chosen here is especially convenient analytically, as will be clear in the remainder of this paper. This modified version of AHS minimum-entropy robustness can naturally be labeled ‘homothetic robustness’. HJB equation (10) then becomes:

$$0 = \sup_{\alpha, C} \left[ U(C) - \delta V + V_w[W(r + \alpha(\mu - r)) - C] + V_t + \frac{1}{2} \left( V_{ww} - \frac{\theta W^2}{(1-\gamma)\sigma^2} \right) \alpha^2 \sigma^2 W^2 \right].$$

(14)

The following results can then be obtained:

**Proposition 1:** When $U(C) = \frac{C^{1-\gamma}}{1-\gamma}$, with $\gamma \neq 1$, equation (14) subject to (5) is solved by

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

(15)

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:

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

(16)

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

(17)

This result is remarkably simple, in light of the complexity of the HJB equation (14).

The optimal fraction of wealth invested in the risky asset is independent of both wealth and time, by virtue of the homotheticity. The portfolio weight is the standard Merton solution, where the usual risk-aversion adjustment $\gamma$ is replaced by $\gamma + \theta > \gamma$. Robustness amounts therefore to an increase in effective 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 parameter determining the consumption wealth

17
ratio, \( a \), reflects the different portfolio weight. This observation leads to a more fundamental insight about the effect of robustness that is explored in the next subsection.

### 2.2 The Link with Stochastic Differential Utility

If I rewrite the HJB equation (14) obtained after solving the infimization problem, an interesting result appears:

\[
0 = \sup_{\alpha, \gamma} \left[ U(C) - \delta V + D^{(\alpha, \gamma)} V(W, t) - \frac{\Psi(W, t)}{2} V_w^2 \sigma^2 W^{-2} \right].
\]  

(18)

As was noted by AHS for the general abstract control problem they analyzed, the structure of this PDE coincides with the general PDE defining stochastic differential utility (SDU, the continuous-time version of recursive utility (e.g. Epstein and Zin (1989) and Weil (1990))). Following the terminology of Duffie and Epstein (1992a and 1992b) or of Duffie and Lions (1992), the aggregator is \( U(C) - \delta V \), i.e. as for the standard additive utility case, and the variance multiplier is \( \Psi(W, t) \), the scaled robustness parameter. For time-additive utility, the variance multiplier is zero, while the typical SDU 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 (\( \hat{\theta} \)).

Despite the very different motivation underlying both preference specifications a strong result pertaining to the relationship between recursive preferences and homothetic robustness, including a precise parameter mapping, can be obtained:

**Proposition 2:** An investor with a homothetic preference for robustness \( \Psi = \frac{\theta}{(1 - \gamma)W} \) and CRRA utility function \( U(C) = \frac{C^{1-\gamma}}{1-\gamma} \), is observationally equivalent to a Duffie-Epstein-Zin investor with elasticity of intertemporal substitution \( \frac{1}{\gamma} \) and coefficient of relative risk
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.

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 $\gamma$, 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 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.

It is worth noting that the exact isomorphy obtains when scaling $\theta$ by the value function to impose homotheticity: having some function $\kappa W^{1-\gamma}$ multiply the entropy penalty in (7) is instrumental in expressing the unit-free entropy penalty in units of $W^{1-\gamma}$, i.e. in utils. This guarantees homotheticity. Having $\kappa$ of the same form as in the value function itself simplifies the analytics. Simultaneously however, it constructs the preference ordering in a recursive fashion, which turns out to exactly match Duffie-Epstein-Zin preferences.

Before discussing the empirical implications of this result, and especially a crucial difference between SDU and robustness that allows for the empirical distinction between both
preference structures, an important extension can be obtained. Given that risk aversion, uncertainty aversion and willingness to substitute intertemporally are behaviorally distinct, I extend the analysis above to investors with Duffie-Epstein-Zin preferences that also worry about model uncertainty. This can be done by taking (14) and replacing 
\[ \frac{1}{\psi} \left( (1 - \gamma) V \right)^{\frac{1-\psi}{1 - \psi}} - \delta (1 - \gamma) V \right) + D^{(\alpha,\gamma)} V(W,t) - \frac{\theta}{2(1 - \gamma) V} V^{\alpha^2 \sigma^2 W^2} \right]. 
\]

where \( \psi^{-1} \) denotes the elasticity of intertemporal substitution, \( \gamma \) is risk aversion and \( \theta \) the robustness parameter. When \( \theta = 0 \), Schroder and Skiadas show that a solution to this problem exists for \( \psi \neq 1 \) if \( \frac{\gamma - \psi}{1 - \psi} < 1 \) and \( \psi > \max(0, \frac{\gamma - \psi}{1 - \psi}) \). The empirically relevant cases of \( \psi > 1 \) and \( \gamma > 1 \) always satisfy this restriction. Adding robustness does not change this as long as \( \theta > 0 \), except when \( \psi < 1 \), in which case \( \theta < 1 - \gamma \) is needed. I will focus on \( \psi > 1 \) and \( \gamma > 1 \). \( \psi < 1 \) and \( \gamma < 1 \) are empirically less plausible and the special cases of \( \gamma = 1 \) and/or \( \psi = 1 \) can be covered along the lines of Appendix B.

**Proposition 3:** Equation (19) subject to (5) is solved by

\[
V(W,t) = \left[ \frac{(1 - e^{-\alpha(T-t)})}{\alpha} \right]^{\psi(1-\gamma)} \frac{W^{1-\gamma}}{1 - \gamma},
\]

where \( \alpha \equiv \frac{1}{\psi} \left[ \delta - (1 - \psi) r - \frac{1-\psi}{2|\gamma|+\theta} \left( \frac{\mu - r}{\sigma} \right)^2 \right] \). The optimal portfolio and consumption rules are given by:

\[
C^*_t = \frac{a}{1 - e^{-a(T-t)}} W_t, \quad \alpha^* = \frac{1}{\gamma + \theta} \frac{\mu - r}{\sigma^2},
\]

Not surprisingly, the only change relative to Proposition 1 concerns the consumption
rule, which reflects the fact that the elasticity of intertemporal substitution has now been disentangled from the coefficient of relative risk aversion. This will be important for the general equilibrium analysis later on.

Given this explicit solution, the generalization of Proposition 2 is readily obtained.

Proposition 4: An investor with a homothetic preference for robustness \( \Psi = \frac{\theta}{(1-\gamma)^{\psi}} \) and Duffie-Epstein-Zin utility with elasticity of intertemporal substitution \( \psi^{-1} \) and risk aversion \( \gamma \), is observationally equivalent to a Duffie-Epstein-Zin investor with elasticity of intertemporal substitution \( \psi^{-1} \) and coefficient of relative risk aversion \( \gamma + \theta \).}

2.3 Empirical Content and Calibration

Robustness essentially induces a shift in parameter space, where risk aversion is increased while keeping intertemporal substitution constant. There is however a subtle difference between SDU with high risk aversion, and robust preferences with high ‘effective risk aversion’. Because of it, the empirical predictions of a model with robustness are different.

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. In fact the model predicts this difference in estimates. This is not true for recursive preferences. Both experiments and introspection typically involve situations with well-specified events and probabilities. In these situations, the preference for robustness is therefore not operational, and one obtains an estimate of just \( \gamma \). 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 robustness. The estimate of ‘risk aversion’ one then ob-
tains combines genuine risk aversion (γ) with uncertainty aversion (θ), and is simply γ + θ. The analysis conducted here therefore allows for an intuitive distinction between risk and (Knightian) uncertainty in the context of financial markets.

Another way to see this is to keep in mind that θ is an environment-specific parameter, which renders effective risk aversion (γ + θ) environment-specific. Environments where model misspecification is potentially more problematic generate higher perceived or effective risk aversion. On the contrary, recursive preferences predict that risk aversion is constant across environments.

Finally, it can be noted that the observational equivalence only obtains after one finds the endogenous worst-case and plugs it into the HJB equation (i.e. moving from (8) to (10)) while simultaneously imposing homotheticity. Doing so replaces the set of priors with a single prior which yields the equivalence to SDU. It should be emphasized that since the worst-case prior is endogenous and changes with the environment, robust preferences would violate the Savage axioms and generate Ellsberg-type behavior. In summary: the observational equivalence is obtained within the context of the asset pricing model. Having external evidence that risk aversion is relatively low in more stylized environments allows one to break the isomorphy.

To explore the quantitative effect of robustness on the optimal portfolio, 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 (9)), along the lines of (6). This in turn offers guidance in selecting a reasonable value for θ, a parameter that was so far left unspecified.
The robust investor can be viewed as using an alternative model (instead of the benchmark model) which adds an endogenous drift term to (2):

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

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:

\[
\frac{dS_t}{S_t} = \left[\mu + \alpha^* W \sigma^2 u^*\right] dt + \sigma dB_t
\]

\[
= \left[\mu - (\mu - r) \frac{\theta}{\gamma + \theta}\right] dt + \sigma dB_t.
\]

where the second equality obtains upon substitution of (9) and (17) for the optimal \(u^*\) and \(\alpha^*\) respectively. Consequently, the investor worries that the expected excess return on the risky asset is not \(\mu - r\), but rather \(EP_P\) defined as:

\[ EP_P \equiv E_t^{u^*} \left[\frac{dS_t}{S_t} - r dt\right] = \frac{\gamma}{\gamma + \theta} (\mu - r) dt, \]

where \(E_t^{u^*}[.\] naturally denotes the expectation according to the alternative model that includes the ‘optimal drift distortion’ \(u^*\). Analogously, relabeling the ‘true’ equity premium \(\mu - r\) as \(EP_T\), \(\theta\) is then found to be:

\[ \theta = \frac{\gamma}{\gamma + \theta} \frac{EP_T - EP_P}{EP_P}, \]

which is informative about reasonable ranges for \(\theta\).

Table 1 reports the optimal portfolio weight \(\alpha^*\) allocated to the risky asset (from (17)) and the associated worst-case or pessimistic scenario \(EP_P\) supporting this portfolio (from
(24)). I use $\mu - r = 6\%$ and $\sigma = 0.16$. I report results for low, moderate and high risk aversion, and for different values of $\theta$. A first observation is that the preference for robustness dramatically decreases the optimal portfolio weight $\alpha^*$ relative to the first row, which is the expected-utility case. Also, as is obvious from (24) or (25), the crucial parameter in terms of the implied least-favorable equity premium is $\theta/\gamma$ rather than $\theta$. 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. Also, as will be shown in detail in the equilibrium calibration in section 3, a 3% worst-case is difficult to distinguish from a 6% true equity premium, even when the estimate is based on a century-long time-series. Having $EP_p = 3\%$ reduces $\alpha^*$ to half its $\theta = 0$ value. Very conservative portfolios can then be obtained: for example, with moderate risk aversion ($\gamma = 5$), $EP_p = 3\%$ results in a portfolio share in equities of only 23.44%, while for high risk aversion ($\gamma = 10$), the portfolio share drops below 12%.

3 Equilibrium Asset Pricing

To explore the equilibrium implications of the robust decision rules derived in the previous section, I now consider a simple exchange economy in the style of Lucas (1978). The representative agent receives an endowment, which he has to consume in equilibrium, and can trade two assets in this economy: a risky asset, entitling the owner to the risky endowment (the dividend), and a riskless asset. The returns of these assets adjust to support a no-trade
equilibrium. Using the explicit partial-equilibrium results for SDU in the presence of robustness, I show in closed-form how the different determinants of behavior (intertemporal substitution $\psi^{-1}$, risk aversion $\gamma$ and uncertainty aversion $\theta$) affect the equilibrium equity premium and riskfree rate. The model is calibrated showing first how much robustness is needed to fully explain the equity premium and riskfree rate puzzle. After that, I explore which equity premium can be generated for plausible parameter values (in particular for $\theta$).

3.1 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:

$$dD_t = \mu_D D_t dt + \sigma_D D_t dB_t,$$

(26)

where the expected instantaneous growth rate $\mu_D$ and the instantaneous standard deviation $\sigma_D$ are strictly positive parameters. I 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 $\mu_S$ and $\sigma_S$ 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 $\alpha$, the representative agent’s wealth dynamics are:

$$dW_t = [W_t (r + \alpha_t (\mu_S - r)) - C_t] dt + \alpha_t \sigma_S W_t dB_t. \quad (27)$$

Analyzing the infinite-horizon case and using the results from section 2, the HJB for a robust investor with intertemporal substitution elasticity $\psi^{-1}$, risk aversion $\gamma$ and preference for robustness $\theta$ is:

$$0 = \sup_{\alpha, C} \left[ \frac{1}{1-\psi} \left( \frac{e^{1-\psi}}{((1-\psi)V)} - \delta (1 - \gamma) V \right) + V_w \left[ W (r + \alpha (\mu_S - r)) - C \right] \right. 
+ \frac{1}{2} \left( V_{ww} - \frac{\theta}{(1-\psi)V^2} V_w^2 \right) \alpha^2 \sigma_S^2 W^2 \right]. \quad (28)$$

**Definition:** A robust equilibrium consists of a consumption rule $C^*$, an investment rule $\alpha^*$, and prices $S$ and $r$, such that simultaneously:

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

2. *The agent solves* (28) *subject to the transversality condition*

$$\lim_{t \to \infty} E \left[ e^{-\delta t} V(W_t) \right] = 0. \quad (29)$$

### 3.2 Closed-Form Solution

Given the closed-form solutions for partial-equilibrium consumption and portfolio decisions, the equilibrium riskfree rate and equity premium are obtained explicitly. In addition, the worst-case scenario for the equity premium supporting the equilibrium is derived. This will prove useful in the subsequent calibration.

**Proposition 5:** In equilibrium, the price of the risky asset is given by $S_t = \frac{1}{\theta} D_t$. The excess return on the risky asset follows:

$$\frac{dS_t + D_t dt}{S_t} - r dt = [\gamma + \theta] \sigma_C S_t dt + \sigma_D dB_t. \quad (30)$$
with $\sigma_{CS} \equiv \text{cov} \left( \frac{dC}{c}, \frac{dS}{S} \right)$. The equilibrium riskfree rate is given by:

$$r = \delta + \psi \mu_D - \frac{1}{2} [1 + \psi] [\gamma + \theta] \sigma_D^2. \quad (31)$$

The pessimistic scenario for the expected equity premium supporting the equilibrium is

$$EP^*_p = \gamma \sigma_{CS} \quad (32)$$

The equilibrium equity premium is given by a CCAPM result (Breeden (1979)) and 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 $\gamma + \theta$: both market risk and model uncertainty are priced in equilibrium. A key empirical prediction of the model is therefore that the price of risk is higher than what would be expected based on genuine risk aversion ($\gamma$) alone. Put differently, the model predicts that risk aversion measures obtained from experiments or introspection in stylized and simple environments (where there is no scope for robustness) are lower than the ‘risk aversion’ inferred from risk premia in asset markets (as these are truly driven by $\gamma + \theta$). As mentioned before, this feature distinguishes the model from recursive preferences without robustness and breaks the observational equivalence that occurs when only examining asset pricing evidence.

The equilibrium riskfree rate depends on the three fundamental determinants of savings in the economy: time preference, intertemporal substitution and growth, and precautionary savings in response to consumption uncertainty. Importantly, robustness drives down the equilibrium riskfree rate through the precautionary savings channel. Having intertemporal substitution decoupled from risk aversion ($\gamma \neq \psi$) will prove important in the calibration later on: it allows high (but reasonable) values of $\gamma$ without counterfactually producing a
high riskfree rate (which would occur with time-additive utility where $\gamma = \psi$). In other words, it is instrumental in avoiding Weil’s riskfree rate puzzle (1989). Having moderately high risk aversion matters because of the worst-case equity premium supporting the equilibrium, given by (32). Of course, this is the equilibrium equity premium in a model with expected-utility agents ($\theta = 0$). The worst-case $EP_p^*$ no longer depends on the preference parameter $\theta$. Mechanically speaking, this is due to the equilibrium condition: one can generally (i.e., also in partial equilibrium) rewrite (24) as $EP_p^* = \alpha \gamma \sigma^2$. Imposing $\alpha = 1$ eliminates $\theta$ from the equation. Intuitively, what’s happening is that while $EP_p^*$ no longer depends on $\theta$, the endogenous equilibrium equity premium does. All $\theta$ 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 given exogenously, so that the least-favorable excess return adjusted. Now the opposite happens.

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

Before turning to the calibration, notice that the model nests the following special cases previously analyzed in the literature. First, time-additive utility without robustness obtains when $\gamma = \psi$ and $\theta = 0$. In that case, the equity premium puzzle and riskfree rate puzzle show up. Second, recursive preferences without robustness involve $\gamma \neq \psi$ and $\theta = 0$. The riskfree rate puzzle can be avoided, but generating a high risk premium requires unreasonably high risk aversion. Finally, introducing robustness in the standard time-additive framework
(\(\gamma = \psi, \theta > 0\)) helps qualitatively. Quantitatively however, the model can only explain part of the equity premium. Relying on moderately high risk aversion (\(\gamma > 2\)) to obtain a high enough \(EP^b\) pushes up the riskfree rate, forcing \(\delta\) to be negative. On the other hand, keeping \(\gamma\) low requires too much robustness: AHS show under time-additive utility that solving the equity premium puzzle requires a very high market price of uncertainty (high \(\theta\)) or a very large gap between the actual and pessimistic equity premium. The gap is so large that it could easily be detected given available time-series. Put differently, the equity premium puzzle shows up again, albeit in the form of excessive uncertainty aversion rather than excessive risk aversion.

Having \(\gamma \neq \psi\) and \(\theta > 0\) extends this previous work and avoids most of these problems. In a nutshell, \(\gamma\) matters for the worst-case, but does not push up the riskfree rate as intertemporal substitution is controlled separately by \(\psi\) (actually in this model, increasing \(\gamma\) lowers the riskfree rate through precautionary savings as is evident from (31)). The gap between the actual and worst-case equity premium is driven by \(\theta\) and can be calibrated analyzing detection-error probabilities as in AHS. Plausible parameter values are consistent with a relatively high equity premium and low riskfree rate.

### 3.3 Calibration and Empirical Implications

The calibration strategy consists of the following steps. Given values for the parameters \(\mu_D (= \mu_C), \sigma_{CS} = \sigma_C \sigma_S p\) and \(\sigma_D^2 (= \sigma_C^2)\), the preference parameters \(\delta, \gamma, \psi\) and \(\theta\) are chosen to match the observed riskfree rate and equity premium according to (30) and (31). I impose strict positivity of the discount rate \(\delta\) and consider only \(\gamma \leq 10\) as imposed by
Mehra and Prescott (1985). In line with recent empirical work using cohort-level (Attanasio and Weber (1995)) and individual-level (Vissing-Jorgensen (2002)) consumption data $\psi^{-1}$ is constrained to be less than one. Initially, $\theta$ is set to match the historical equity premium. As a next step, the worst-case scenario $EP^*_p$ supporting this equilibrium is analyzed. Finally, the plausibility of $\theta$ needed to generate historical equity premia given this worst-case scenario is assessed using detection-error probabilities described in the next subsection. Calibrating the model using time-series of different length will provide insight, given the link between the span of the time-series and the robustness parameter $\theta$ described there. This will illustrate the claim made earlier that $\theta$ is an environment-specific parameter.

### 3.3.1 Implied pessimistic scenario

The data is taken from Campbell (1999): a long annual sample from 1891 to 1994 and a quarterly post-war sample from 1947.2 to 1996.3. Table 2 reports the results without and with robustness for both samples. The century-long sample is easier to match given the high covariance of consumption growth and stock returns (stemming from high consumption variability, high return volatility and especially relatively high correlation between stock returns and consumption growth), combined with a lower equity premium and higher risk-free rate as compared to the post-war sample. Expected utility requires a risk aversion coefficient of 21 to explain the equity premium. Even for this sample the risk-free rate puzzle shows up for expected utility as $\delta$ becomes negative. Especially the post-war sample is problematic when $\theta = 0$.

Introducing robustness makes the equity premium puzzle disappear. Risk aversion
can be kept relatively low at 7 and 10. Plausible values for both $\psi$ and $\delta$ explain the low riskfree rate, owing to the precautionary savings motive associated with the values for $\gamma$ and $\theta$. These calibration results are useful in understanding how non-robust recursive preferences and robustness are observationally equivalent in asset markets, yet can easily be distinguished outside. Applying the shift in parameter space establishes in Proposition 4 it is immediate to see that Duffie-Epstein-Zin preferences resolve the equity premium and riskfree rate puzzle with the same low $\delta$ and reasonable $\psi$. However, risk aversion becomes $\gamma + \theta$ and needs to be as high $247$ in the postwar sample. Casual introspection and experimental evidence suffice to rule this out and therefore to break the isomorphy.

The key question remaining is the plausibility of the required $\theta$ parameter. At a first and informal level, this question can be addressed by examining the worst-case scenario for the equity premium. The worst-case is given by (32) and satisfies a strong restriction. For the century-long annual sample, $EP^*_P = 2.1\%$. 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 “(...) the typical pessimistic 1-in-20 case 30-year scenario foreseen by financial economists is about the equity premium which Mehra and Prescott (1985) consider to be consistent with reasonable risk aversion (...)” (Welch (2000), p. 15). This is exactly what happens in my model theoretically and, if $\sigma_{CS}$ is sufficiently high, also empirically. Indeed in the postwar sample, the equity premium puzzle is especially strong given that $\gamma \sigma_{SC} = 0.32\%$. This is still the equity premium generated by an expected-utility model, even with high risk aversion ($\gamma = 10$). It is however substantially below the average survey response.\textsuperscript{16}
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., we 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.\textsuperscript{17} If agents are aware of this and 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.

### 3.3.2 Calibrating $\theta$ using detection-error probabilities

As a final step in the calibration, I address the question of a reasonable $\theta$ given the worst-case implied by the model. The gap between the reference and worst-case model, and therefore here the expected equity premium $EP^* \equiv \mu_S - r = EP^*_e + \theta \sigma_{CS}$, is dictated by $\theta$. AHS link the calibration of $\theta$ elegantly to a Bayesian model selection problem faced by the robust decision-maker (see also Chapter 8 of Hansen and Sargent (2002)). When choosing between two potential models or data generating processes (say the model associated with
$EP_T^*$ versus the one associated with $EP_P^*$), a decision-maker can perform likelihood ratio tests under both models based on available data. The models are difficult to distinguish if the probability of rejecting one model mistakenly in favor of the other is high. This probability is given by the probability that the log likelihood ratio is negative even though the rejected model is correct (denoted by $p_i$ when model $i$ holds). Given initial priors of 0.5 on each model and a sample of length $N$, the detection-error probability $\varepsilon_N$ is defined as:

$$\varepsilon_N = 0.5p_1 + 0.5p_2$$

When $\varepsilon_N$ is high, the two models are difficult to distinguish empirically given sample $N$. Because the alternative model is indexed by $\theta$, $\varepsilon_N$ will depend on $\theta$. A degree of robustness is then reasonable if the associated alternative model is difficult to distinguish from the benchmark as revealed by the high detection-error probability $\varepsilon_N$. Hansen and Sargent suggest 10% as the lowest $\varepsilon_N$ that a robust decision maker should reasonably worry about.\(^{18}\)

In the context of the present Gaussian i.i.d. set-up, $\varepsilon_N$ can easily be calculated. Because the benchmark and alternative model are both geometric Brownian motions with constant coefficients, the log-likelihood ratios (or Radon-Nikodym derivatives) turn out to be Brownian motions and are hence normally distributed random variables. It is straightforward to derive that the detection-error probability is given by

$$\varepsilon_N = \text{prob} \left[ x < - \left( \frac{EP_T^* - EP_P^*}{2\sigma_S} \right) \sqrt{N} \right] = \text{prob} \left[ x < - \left( \frac{\theta_0\sigma \mu}{2} \right) \sqrt{N} \right]$$

where $x \sim N(0, 1)$ and the last equality follows from the equilibrium results. Notice that the span or length of the time-series used is relevant, not the frequency, as emphasized in Merton (1980) and in AHS. For instance, moving from annual to quarterly data (multiplying...
will not affect $\varepsilon$ as the difference in quarterly Sharpe ratios \(\frac{EP_T^* - EP_P^*}{2\sigma_s}\) will simultaneously shrink by a factor \(\sqrt{4}\).

Table 3 reports $\varepsilon_N$ for the uncertainty aversion parameters $\theta$ that match asset prices in Table 2 for both samples. Interestingly, $\theta = 14$ in the century-long sample is reasonable. The difference between $EP_T^*$ and $EP_P^*$ of about 4% could give rise to sufficient statistical ‘confusion’ to warrant consideration. In the post-war sample however $\theta = 237$, and therefore $EP_T^* - EP_P^*$ close to 7.5%, can be ruled out as unreasonable as the probability of selecting the wrong model is merely 4%.

Finally, to determine how much robustness corresponds to different tolerance levels for $\varepsilon_N$, (33) can be inverted. Table 4 indicates that a 6% equity premium is easy to obtain in the long sample, with risk aversion as low as 7 as in Table 2. Even with one century of data available it is difficult to distinguish models with a 4.5% difference in expected returns, given their volatility. For the post-war sample, the robust equilibrium can generate a 5.8% equity premium. Put differently, in the post-war sample $\theta$ can be as high as 173 and still be plausible on statistical grounds. Effectively, this allows me to set $\gamma + \theta$ as high as 183, which would be highly problematic without robustness as it would require risk aversion of 183. Table 4 shows how $\theta$ is an environment-specific parameter as was argued before. Its value depends on the diffusion coefficient of the Brownian motion to be detected and the length of the time-series available in attempting this. Detection-error probabilities are a powerful tool in calibrating $\theta$: a given detection-error probability corresponds to $\theta$’s that are an order of magnitude different for the post-war versus century-long sample.
4 Conclusion

Model uncertainty is pervasive in many aspects of financial decision-making. Even a parameter as crucial as the expected equity return is the subject of major disagreement and dispute, based on both theoretical and empirical considerations (see for example Cochrane (1998) and Welch (2000)). This paper takes such uncertainty into account and shows its effects on both dynamic portfolio rules and equilibrium asset pricing. To incorporate a preference for robustness, I use the framework of Anderson, Hansen and Sargent (2002).

A first contribution of the paper is the modification of the AHS framework that imposes homotheticity and yields analytical tractability. I present simple closed-form solutions for the optimal consumption and portfolio rules of a robust investor with power utility (and an extension to SDU) facing constant investment opportunities. 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 \( \theta \), the parameter measuring the strength of the preference for robustness. In a calibration, reasonable parameter values lower the optimal portfolio share in equities by 50%. Secondly, I establish an isomorphy between robustness and recursive preferences in the sense of Duffie-Epstein-Zin. Essentially, robustness can be interpreted as increasing the investor’s risk aversion (from \( \gamma \) to \( \gamma + \theta \)) without affecting her willingness to substitute intertemporally. Nonetheless, a subtle difference between recursive preferences and robustness persists. Robustness, unlike recursive preferences, can reconcile relatively low degrees of risk aversion, reflected in introspection and experimental evidence, with the high ‘risk aversion’ inferred from asset prices.
Introspection or experiments in environments with well-specified probabilities and events elicit the degree of risk aversion $\gamma$, because the preference for robustness is not operational. At the same time, asset prices in addition reflect uncertainty aversion or a preference for robustness ($\theta$), so that ‘risk aversion’ can be perceived to be $\gamma + \theta$.

In an equilibrium setting, I show in closed-form how robustness helps to resolve both the equity premium and the riskfree rate puzzle. A calibration based on detection-error probabilities shows that the model can generate a 4 to 6% equity premium for reasonable parameter values. The calibration also shows that $\theta$ is environment-specific and is an order of magnitude higher for the post-war sample than for the century-long sample. The worst-case scenario supporting the equilibrium is 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, she 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.

Extending the framework considered here is important. A strong assumption in the work on robustness is that decision-makers do not engage in any learning. Epstein and Schneider (2002) present a model of learning based on their recursive multiple-prior model, which is an axiomatic version of Epstein and Wang (1994) and therefore a recursive version of the atemporal model of Gilboa and Schmeidler. Knox (2002) develops an alternative model of learning when facing uncertainty or ambiguity and obtains analytical solutions for the optimal portfolio problem. He contrasts this with both the robust approach of this
paper and the ambiguity-aversion approach of Chen and Epstein (2002).

Skiadas (2003) shows that the isomorphy between robustness and stochastic differential utility holds more generally in non-Markovian environments. Uppal and Wang (2003) extend the portfolio problem of a homothetically robust investor to allow for multiple risky assets. Liu, Pan and Wang (2002) consider jumps in the aggregate endowment and show how homothetic robustness generates a rare-event premium that can explain option prices. Maenhout (2000) studies optimal portfolio rules for homothetically robust investors facing stochastic investment opportunities (due to a mean-reverting risk premium or due to stochastic volatility) and obtains closed-form solutions for the optimal intertemporal hedging demands.
Appendix A: Proofs

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

Proof of Proposition 1: Immediate by plugging (15), (16) and (17) into (14). The boundary condition \( V(W,T) = 0 \) is satisfied. Finally, nonnegativity of consumption requires \( \alpha > 0 \). The proof for the limiting case of \( \gamma = 1 \) is given in Appendix B.

Proof of Proposition 2: It suffices to prove that the preferences of a homothetically robust agent with \( U(C) = \frac{C^{1-\gamma}}{1-\gamma} \) and \( \Psi = \frac{\theta}{(1-\gamma)\nu} \) 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 \( \frac{1}{\eta} \). Using the terminology of Duffie and Epstein (1992a and 1992b), the aggregator and variance multiplier of the robust agent are:

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

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 \( \varphi(v) = -|v|^{\frac{1-\gamma}{1-\gamma}} \), the normalized aggregator is then:

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

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 2 implies that \( \gamma = 1 - \rho \) and \( \theta = \rho - \alpha \). Thus, the normalized aggregator can be rewritten in terms of the parameters of Duffie and Epstein as\(^{19}\):

\[
 \tilde{f}(c,v) = \frac{\alpha}{\rho} \left[ \frac{c^\rho}{\rho} |v|^{\frac{\alpha-\rho}{\rho}} - \delta v \right].
\]
Finally, using the transformation \( \varphi(v) = \frac{(\delta/\rho)^{\alpha/\rho}}{v} \), 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 \( \varphi(v) = \frac{1 - \gamma}{1 - \gamma} v \), which is strictly increasing if \( \theta < 1 - \gamma \), and which yields:

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

Again, using the mapping from the proposition, and the transformation \( \varphi(v) = \frac{(\delta/\rho)^{\alpha/\rho}}{v} \), 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 3:** Immediate by plugging (20), (21) and (22) into (19). The boundary condition \( V(W, T) = 0 \) is satisfied. Finally, nonnegativity of consumption requires \( a > 0 \). ■

**Proof of Proposition 4:** Following the line of reasoning of the proof to proposition 2, it needs to be shown that the preferences of the homothetically robust agent with SDU preferences are ordinally equivalent to those of an agent with SDU without robustness. The aggregator and variance multiplier of the robust agent are \( f(c, v) = \frac{1 - \gamma - \theta}{1 - \gamma} \left[ \frac{c^{\gamma - 1} - (1 - \gamma \alpha v)^{\gamma - \theta}}{(1 - \gamma \alpha v)^{\gamma - \theta}} - \delta v \right] \) and \( A(v) = \frac{\theta}{(1 - \gamma)v} \) respectively. With \( \gamma > 1 \), both \( 1 - \gamma \) and \( v \) are negative. Hence, \( A(v) = \frac{\theta}{(1 - \gamma)v} = \frac{\theta}{(1 - \gamma)v} \). Applying as before \( \varphi(v) = -|v|^{\frac{1 - \gamma - \theta}{1 - \gamma}} \):

\[
\overline{f}(c, v) = \frac{1 - \gamma - \theta}{1 - \gamma} \left[ \frac{c^{\gamma - 1} - (1 - \gamma \alpha v)^{\gamma - \theta}}{(1 - \gamma \alpha v)^{\gamma - \theta}} - \delta v \right].
\]
The parameter definitions of Duffie and Epstein (see proof to proposition 2) and the mapping for the preference parameters of proposition 4 imply that $\psi = 1 - \rho$ and $\gamma + \theta = 1 - \alpha$.

Thus, in terms of the parameters of Duffie and Epstein:

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

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

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

the normalized aggregator for SDU as defined in Duffie and Epstein.

**Proof of Proposition 5**: The equilibrium can be constructed as follows. The optimality conditions are:

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

$$C^* = aW,$$

where $a \equiv \frac{1}{\psi} \left[ \delta - (1 - \psi) \rho - \frac{1 - \psi}{2[\gamma + \theta]} \left( \frac{\mu - r}{\sigma} \right)^2 \right]$. Substituting these into the stochastic differential equation for wealth (27) gives:

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

Clearing of the security market together with (34) immediately gives a CCAPM result:

$$\mu - r = (\gamma + \theta) \sigma^2.\]$$

Using this in the stochastic differential equation for wealth:

$$\frac{dW_t}{W_t} = \left[ \frac{r}{\psi} - \frac{\delta}{\psi} + \frac{1 + \psi}{2\psi} [\gamma + \theta] \sigma^2 \right] dt + \sigma dW_t.$$

Also, market clearing in the goods market implies that:

$$C_t = D_t = aW_t.$$
Moreover $S_t = \frac{1}{\alpha} D_t$, so that in equilibrium $S_t = W_t$. In addition $S_t = \frac{1}{\alpha} D_t$ implies immediately that $\frac{dS_t}{dt} = \mu_D dt + \sigma_D dB_t$. Combining this with the implied wealth dynamics in (36) yields the following equilibrium condition:

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

This results in:

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

Rearranging the last two equations immediately produces the desired results in (30) and (31). Finally, it is straightforward to verify that the optimal consumption and investment strategies, (34) and (35) respectively, are consistent with this equilibrium. Similarly, the transversality condition is satisfied as long as $\alpha > 0$, 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 $\theta$ appropriately even for log utility, despite the fact that logarithmic utility leads to homothetic preferences even for $\Psi(W, t) = \theta$. 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 $\alpha$ 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 \( \Psi(W, t) = \theta \). The source of this horizon effect can be understood from equation (9): when \( \theta \) 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 \( \theta \) 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 \( \Psi_{\log}(W, t) = \lim_{\gamma \to 1} \frac{\theta}{(1-\gamma)V(W, t)} \), where the limit is computed using the explicit solution for \( V \).

**Proof of Proposition 1 for \( \gamma = 1 \):** Based on the solution given in Proposition 1 for \( V \) when \( \gamma \neq 1 \), \( \Psi_{\log} = \lim_{\gamma \to 1} \frac{\theta}{(1-\gamma)V} = \frac{\delta \theta}{1-e^{-\delta(t-T)}} \). The appropriate limit of HJB equation (14) for \( \gamma \to 1 \) is then:

\[
0 = \sup_{\alpha, C} \left[ \log(C) - \delta V + V_w [W (r + \alpha(\mu - r)) - C] + V_t + \frac{1}{2} \left( V_{ww} - \frac{\delta \theta}{1-e^{-\delta(t-T)}} V_w \right) \alpha^2 \sigma^2 W^2 \right],
\]

subject to (2). It is straightforward to verify that this PDE is solved by \( V(W, t) = \frac{1-e^{-\delta(T-t)}}{\delta} \log W + \phi(t) \), where \( \phi(t) \) solves

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

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

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

\[
\alpha^* = \frac{\mu - r}{1 + \theta \sigma^2},
\]
i.e. the special cases of (16) and (17) for $\gamma = 1$. The portfolio rule contrasts with the solution that would have obtained when using $\Psi_{\log}(W, t) = \theta$, namely $\alpha^* = \frac{\delta}{\delta + \theta(1 - e^{-\gamma(1 - \theta)})} \frac{u - r}{\sigma^2}$ as in Schroder and Skiadas. ■
References


Notes


2Dispersion of opinions or lack of consensus need not reflect high uncertainty. However, dispersion of opinions among experts may instill uncertainty or ambiguity in the mind of non-expert decision-makers that seek guidance from them.

3Entropy plays a fundamental role in information theory, statistics and econometrics (see e.g. the special issue of the *Journal of Econometrics* (March 2002)). More references and discussion can be found in AHS.

4This type of (absolutely continuous) change of measure is of course well-known in finance and underlies risk-neutral pricing techniques introduced by Cox and Ross (1976) and Harrison and Kreps (1979). The connection between robustness and the seminal contribution of He and Pearson (1991) is especially noteworthy. They tackle the problem of consumption and portfolio choice for dynamically incomplete markets using martingale techniques. When markets are incomplete, the equivalent martingale measure is not unique. He and Pearson introduce the notion of a minimax martingale measure, i.e. the least-favorable state prices consistent with no arbitrage. The minimax martingale measure ensures the marketability of consumption plans. When investors worry about model uncertainty, the situation is quite different, albeit formally similar. Although markets are dynamically complete and the equivalent martingale unique, robust investors entertain a family of similar alternative
measures to guard against misspecification. Optimizing against the least-favorable measure ensures that the decisions are robust to model uncertainty.

5See Wang (2003) for an axiomatization of a class of preferences that includes the multiplier formulation.

6If anything, robustness would be expected to make the value function more concave, thereby reinforcing the argument.

7Although homotheticity obtains for log utility even for a constant \( \tilde{\theta} \), it is shown in Appendix B that one should still scale \( \theta \) appropriately, so that the homothetic results for power utility converge to the solution for logarithmic utility as \( \gamma \) tends to unity. This is achieved by using \( \lim_{\gamma \to 1} \Psi(W,t) \) in the HJB for log utility.

8Using the value function \( V \) itself to scale \( \theta \) may make it difficult to formulate a time-zero problem, as \( V \) is only known once the problem is solved. However, up to a factor independent of wealth, one can also scale by \( \kappa W^{1-\gamma} \) as argued before. This suffices for homotheticity. The precise form of the function in \( W \) (initial wealth in the time-zero problem) is easily conjectured from the homotheticity of the non-robust problem. The fact that homotheticity requires scaling by a function of beginning-of-period wealth is also apparent from the work of Knox (2002), who shows that the preferences proposed here obtain in discrete time when scaling the entropy penalty by \( V(W_t, t+1) \), the value function evaluated at time \( t+1 \), but at wealth \( W_t \).

9I divide by \( (1-\gamma)V \), rather than by \( V \), for two reasons. First, this ensures that \( \Psi \) is always positive for positive \( \theta \), and avoids the use of the absolute value operator. Second,
this scaling is economically meaningful in the sense that its limit as $\gamma \to 1$ is nontrivial once one knows the explicit form of $V$ for $\gamma \neq 1$, and yields decision rules for general CRRA utility that nest logarithmic utility. See Appendix B for details.

10 Alternatively, re-scaling all variables in the problem by wealth would help equally in avoiding the wealth dependence and imposing stationarity. In the one-dimensional case as considered here, one would expect the same results. The alternative scaling might lead to different results in multi-dimensional environments however. I am grateful to the referee for pointing this out.

11 This detection-error probability approach is due to AHS and Hansen and Sargent (2002) and is a powerful calibration tool. For the values of Table 1, $EP_P = 3\%$ corresponds to $\varepsilon_{N=100} = 0.174$, while $EP_P = 2\%$ is associated with $\varepsilon_{N=100} = 0.105$. Both are therefore reasonable as will be argued below.

12 By construction, the simple model considered here does not address the excess volatility puzzle (see for instance the survey in Campbell (1999)). Richer dynamic models are required for this. Incorporating robustness in those settings is an interesting direction for future research. Cagetti et al. (2002) is a first example of this promising line of work.

13 It is noteworthy that the results for $\gamma \neq \psi$ and $\theta = 0$ differ from the pricing relations obtained by Epstein and Zin (1991) and Duffie and Epstein (1992b). They prove that the risk premium on an asset is a linear combination of the CCAPM and the standard CAPM. This result has been criticized because the return on wealth is itself an endogenous variable (Campbell (1999)). I obtain a different result because I use the explicit portfolio weights
and consider a single risky asset.

14Campbell annualizes the quarterly raw data (assuming no serial correlation) to facilitate interpretation. All magnitudes here are also expressed in annual terms.

15The so-called correlation puzzle (Cochrane and Hansen (1992)) is therefore not ignored in this calibration. By construction, the theoretical model predicts $\sigma_c = \sigma_S = \sigma_D$ and imposes a counterfactual unit correlation between shocks to consumption growth and shocks to stock returns. This feature is prevalent in most simple exchange economies and often built into stochastic discount factor analysis. As noted by Cochrane and Hansen, some calibration exercises implicitly ignore the low empirical correlation, making it somewhat easier to generate high equity premia.

16Ignoring the correlation puzzle and setting $\rho = 1$ would obviously help as $EP^*_\rho$ becomes 1.65%. Less robustness would then be needed to match the historical premium. Explaining the riskfree rate becomes more difficult given the reduced precautionary savings motive associated with the lower $\theta$.

17This 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.

18AHS introduce Chernoff entropy which plays a role in deriving bounds on $\varepsilon_N$. I do not further explore this here since the actual $\varepsilon_N$ can easily be calculated given the Gaussian i.i.d. nature of the model. AHS demonstrate that Chernoff entropy and the associated $\varepsilon_N$ bounds are powerful model selection tools in more complicated environments.
This is also the parametrization used in Schroder and Skiadas (1999).

Schroder and Skiadas impose the same restriction in their proofs of existence and optimality.
Table 1: Portfolio Share Invested in Equity and Implied Pessimistic Scenario

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\alpha^*$</th>
<th>$EP_\gamma$</th>
<th>$\alpha^*$</th>
<th>$EP_\gamma$</th>
<th>$\alpha^*$</th>
<th>$EP_\gamma$</th>
<th>$\alpha^*$</th>
<th>$EP_\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.3438</td>
<td>0.0600</td>
<td>0.4688</td>
<td>0.0600</td>
<td>0.3348</td>
<td>0.0600</td>
<td>0.2344</td>
<td>0.0600</td>
</tr>
<tr>
<td>0.1</td>
<td>2.1307</td>
<td>0.0545</td>
<td>0.4596</td>
<td>0.0588</td>
<td>0.3301</td>
<td>0.0592</td>
<td>0.2321</td>
<td>0.0594</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5625</td>
<td>0.0400</td>
<td>0.4261</td>
<td>0.0545</td>
<td>0.3125</td>
<td>0.0560</td>
<td>0.2232</td>
<td>0.0571</td>
</tr>
<tr>
<td>1</td>
<td>1.1719</td>
<td>0.0300</td>
<td>0.3906</td>
<td>0.0500</td>
<td>0.2930</td>
<td>0.0525</td>
<td>0.2131</td>
<td>0.0545</td>
</tr>
<tr>
<td>2</td>
<td>0.7813</td>
<td>0.0200</td>
<td>0.3348</td>
<td>0.0429</td>
<td>0.2604</td>
<td>0.0467</td>
<td>0.1953</td>
<td>0.0500</td>
</tr>
<tr>
<td>5</td>
<td>0.3906</td>
<td>0.0100</td>
<td>0.2344</td>
<td>0.0300</td>
<td>0.1953</td>
<td>0.0350</td>
<td>0.1563</td>
<td>0.0400</td>
</tr>
<tr>
<td>10</td>
<td>0.2131</td>
<td>0.0055</td>
<td>0.1563</td>
<td>0.0200</td>
<td>0.1379</td>
<td>0.0247</td>
<td>0.1172</td>
<td>0.0300</td>
</tr>
</tbody>
</table>

This table reports the optimal portfolio weight $\alpha^*$ allocated to the risky asset according to $\alpha^* = \frac{1}{\gamma + \theta} \frac{\mu - \tau}{\sigma^2}$, along with the associated pessimistic scenario $EP_\gamma$ supporting the portfolio, $EP_\gamma \equiv E_t^u \left[ \frac{dS_t}{S_t} - rdt \right] = \frac{\gamma}{\gamma + \theta} (\mu - \tau) dt$. The parameters $\mu - \tau$ and $\sigma$ are 6% and 0.16, respectively.
Table 2: Preference Parameters Required to Match Riskfree Rate and Equity

<table>
<thead>
<tr>
<th>Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
</tr>
<tr>
<td>Estimated consumption and return parameters</td>
</tr>
<tr>
<td>$\mu_C$</td>
</tr>
<tr>
<td>$\sigma_C$</td>
</tr>
<tr>
<td>$\sigma_S$</td>
</tr>
<tr>
<td>$\rho$</td>
</tr>
<tr>
<td>$r$</td>
</tr>
<tr>
<td>$\mu_S - r$</td>
</tr>
<tr>
<td>Required preference parameters</td>
</tr>
<tr>
<td>$\delta$</td>
</tr>
<tr>
<td>$\psi^{-1}$</td>
</tr>
<tr>
<td>$\gamma$</td>
</tr>
<tr>
<td>$\theta$</td>
</tr>
<tr>
<td>$EP_p^*$</td>
</tr>
</tbody>
</table>

The top panel reports the annualized consumption and return parameters estimated by Campbell (1999) for a long annual sample (1891 to 1994) and for a quarterly post-war sample (1947.2 to 1996.3). For both sets of parameters, the bottom panel gives the preference parameters required by the equilibrium asset-pricing model, both for standard expected utility (no robustness or $\theta = 0$) and for SDU with robustness ($\theta \neq 0$). The last row reports $EP_p^*$, the pessimistic scenario for the expected equity premium supporting the equilibrium.
Table 3: Detection-Error Probabilities for $\theta$ as Required in Table 2

| Sample       | 1891 - 1994 | 1947.2 - 1996.3 |
|--------------|-------------|----------------|---|---|---|
| $EP^*_p$     | 2.1%        | 0.32%          |   |   |   |
| $\theta$     | 14          | 237            |   |   |   |
| $\varepsilon_N$ | 12.39%      | 3.98%          |   |   |   |

This table gives the detection-error probabilities $\varepsilon_N$ for $\theta$ and associated $EP^*_p$, as required by the equilibrium model in Table 2, both for the long and short sample. For $x \sim N(0,1)$,

$$\varepsilon_N = \text{prob} \left[ x < - \left( \frac{\theta \sigma \epsilon \rho}{\sqrt{2}} \right) \sqrt{N} \right].$$
Table 4: Calibration of $\theta$ as a function of $\varepsilon_N$

<table>
<thead>
<tr>
<th>Sample</th>
<th>1891 - 1994</th>
<th>1947.2 - 1996.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_N$</td>
<td>10%  20%  30%</td>
<td>10%  20%  30%</td>
</tr>
<tr>
<td>$\theta$</td>
<td>15.5  10.2  6.4</td>
<td>173  114  71</td>
</tr>
<tr>
<td>$EP_T^* - EP_P^*$</td>
<td>4.65%  3.06%  1.92%</td>
<td>5.51%  3.63%  2.26%</td>
</tr>
<tr>
<td>$EP_T^*$</td>
<td>6.75%  5.16%  4.02%</td>
<td>5.83%  3.95%  2.58%</td>
</tr>
</tbody>
</table>

For both samples, $\theta$ (and the corresponding $EP_T^* - EP_P^*$ and $EP_T^*$) are computed for different detection-error probabilities $\varepsilon_N$. 