Dynamic Asset Allocation
with Ambiguous Return Predictability[*]

Hui Chen[†]   Nengjiu Ju[‡]   Jianjun Miao[§]

March 19, 2013

Abstract

We study an investor's optimal consumption and portfolio choice problem when he is confronted with two possibly misspecified submodels of stock returns: one with IID returns and the other with predictability. We adopt a generalized recursive ambiguity model to accommodate the investor's aversion to model uncertainty. The investor deals with specification doubts by slanting his beliefs about submodels of returns pessimistically, causing his investment strategy to be more conservative than the Bayesian strategy. This effect is especially strong when the submodel with a low Bayesian probability delivers a much smaller continuation value. Unlike in the Bayesian framework, the hedging demand against model uncertainty may cause the investor's stock allocation to decrease sharply given a small doubt of return predictability, even though the expected return according to the VAR model is large. Adopting the Bayesian strategy can lead to sizable welfare costs for an ambiguity-averse investor, especially when he has a strong prior of return predictability.

Keywords: ambiguity aversion, model uncertainty, learning, portfolio choice, robustness, return predictability, model misspecification

JEL Classification: D81, D83, G11, E21

[†]MIT Sloan School of Management, 77 Massachusetts Ave, Cambridge, MA 02139. Email: huichen@mit.edu. Tel.: 617-324-3896.

[‡]Department of Finance, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: nengjiu@ust.hk. Tel: (+852) 2358 8318.

[§]Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215. Email: miaoj@bu.edu. Tel.: 617-353-6675.

1. Introduction

One of the most debated questions in recent financial research is whether asset returns or equity premia are predictable.[1] This question is of significant importance for portfolio choice. If asset returns are independently and identically distributed (IID) over time, then the optimal asset allocation is constant over time (Merton (1969) and Samuelson (1969)). However, if asset returns are predictable, then the optimal asset allocation depends on the investment horizon and the predictive variables (Brennan, Schwartz, and Lagnado (1997), Campbell and Viceira (1999) and Kim and Omberg (1996)). Economists have different views on whether asset returns are predictable. Welch and Goyal (2008) argue that the existing empirical models of predicting asset returns do not outperform the simple IID model both in sample and out of sample, and thus are not useful for investment advice. Campbell and Thompson (2008) argue that the empirical models of predictability can yield useful out-of-sample forecasts if one restricts parameters in economically justified ways. Cochrane (2008) points out that poor out-of-sample performance is not a test against the predictability of asset returns.

While many estimation models deliver significant variations in expected returns, the predictive relation is statistically weak and unstable. The estimated predictability coefficient is typically not quite significant and R^2 is generally low. In addition, the sample period and predictive variables are important for regression performance. This suggests that the estimation models may be misspecified. The contrast between the economic significance of various return predictability models and their marginal statistical significance presents a dilemma for investors. The significant variation in expected returns predicted by these models implies aggressive market timing strategies, which can be very costly if they turn out to be wrong.

How should a long-term investor make consumption and portfolio choice decisions when facing alternative possibly misspecified models of asset returns? To address this question, we build a dynamic model in which an investor is concerned about model misspecification and averse to model uncertainty.[2] Following most papers in the portfolio choice literature, we consider a simple environment in which the investor allocates his wealth between a risky stock and a risk-free bond with a constant real interest rate. We depart from this literature and the rational expectations hypothesis by assuming that there are two submodels of the stock return process: an IID model and a vector autoregressive (VAR) model. For simplicity, we adopt the (demeaned) dividend yield as the single predictive variable in the VAR estimation and abstract away from parameter uncertainty. The investor is unsure which one is the true model of the stock return, and thus faces a model selection problem. The investor can learn about the asset return model by observing past data.

The standard Bayesian approach to this learning problem is to impose a prior over the possible

---

[1] For an example, see the July 2008 issue of the *Review of Financial Studies*.

[2] Our notion of model uncertainty is in the sense of Knightian uncertainty or ambiguity in that no known probabilities are available to guide choices. A classical example to illustrate ambiguity and ambiguity aversion is the Ellsberg Paradox (Ellsberg (1961)).

stock return submodels. The posteriors and likelihoods are derived by Bayesian updating. They can be reduced to a single predictive distribution by Bayesian averaging. One can then solve the investor's decision problem using this predictive distribution in the standard expected utility framework (see Barberis (2000), Pastor and Stambaugh (2012), Wachter and Warusawitharana (2009), and Xia (2001)). We depart from this Bayesian approach in that we assume that posteriors and likelihoods cannot be reduced to a predictive distribution in the investor's utility function. This irreducibility of compound distributions captures attitudes towards model uncertainty or ambiguity, as discussed by Segal (1987), Klibanoff, Marinacci, and Mukerji (2005, 2009), Hansen (2007), Seo (2009), Hayashi and Miao (2011) and Ju and Miao (2012). The standard Bayesian approach implies ambiguity neutrality.

To accommodate model ambiguity and ambiguity aversion, we adopt a recursive ambiguity utility model recently proposed by Hayashi and Miao (2011) and Ju and Miao (2012), who generalize the model of Klibanoff, Marinacci, and Mukerji (2005, 2009). This generalized recursive ambiguity model is tractable in that it is smooth and allows for flexible parametric specifications, e.g., a homothetic functional form, as in Epstein and Zin (1989). We may alternatively interpret this utility model as a model of robustness in that the investor is averse to model misspecification and seeks robust decision making. We find that an ambiguity-averse investor slants his beliefs towards the submodel of stock returns that delivers the lowest continuation value. The endogenous evolution of these pessimistic beliefs has important consequences in the consumption and portfolio choice decision and welfare implications.

We calibrate the ambiguity aversion parameter using thought experiments related to the Ellsberg Paradox (see Halevy (2007) and the references cited therein). Our calibrated value is consistent with the experimental finding reported by Camerer (1999), which suggests that the ambiguity premium is typically about 10 to 20 percent of the expected value of bets. We use our calibrated value of the ambiguity aversion parameter, the standard value of risk aversion parameter, and econometric estimates of the stock return process to solve an ambiguity-averse investor's decision problem numerically. We refer to the optimal stock allocation rule for an ambiguity-averse investor as the robust strategy. We compare this robust strategy with three other investment strategies widely studied in the literature: the IID strategy, the VAR strategy, and the Bayesian strategy. The IID and VAR strategies refer to the optimal investment strategies when the investor knows for sure that the stock return follows an IID model and a VAR model, respectively. The Bayesian strategy refers to the optimal investment strategy under Epstein-Zin utility in the Bayesian framework.[3]

We show that the robust stock allocation depends on the investment horizon, the beliefs about the model of stock returns, and the predictive variable. Compared to the Bayesian strategy with identical values of the intertemporal elasticity of substitution and the risk aversion parameter, the robust strategy is more conservative in that it recommends an ambiguity-averse investor to hold less

---

[3]Assuming that the investor maximizes expected utility from next-period wealth, Kandel and Stambaugh (1996) study myopic strategy in a Bayesian framework.

stocks than a Bayesian investor, inducing more nonparticipation in the stock market. To understand the differences between the Bayesian and the robust strategies, we first review the portfolio rule under the Bayesian strategy studied by Xia (2001) for the case of parameter uncertainty. The Bayesian stock demand can be decomposed into a myopic demand and an intertemporal hedging demand. The myopic demand depends on the expected return, which is the weighted average of the expected returns from the two submodels of stock returns. The hedging demand can be further decomposed into two components. The first component is the hedging demand associated with the predictive variable. This component is analyzed by Campbell and Viceira (1999) and Kim and Omberg (1996) in settings without model uncertainty. The second component is the hedging demand against model uncertainty. High realized returns lead the Bayesian investor to shift his posterior beliefs towards (away from) the VAR model when the predictive variable is sufficiently large (small), which may make this hedging demand negative (positive).

What makes the robust strategy different from the Bayesian strategy is that an ambiguity-averse investor effectively makes investment decisions using endogenously distorted beliefs, instead of the actual predictive distribution. For a given nondegenerate prior, the distortion in beliefs is large when the difference in continuation values under the two submodels of stock returns is large. In this case, an ambiguity-averse investor is concerned about the potential large utility loss due to model misspecification and hence shifts his beliefs towards the submodel that delivers a lower continuation value. Consequently, both the myopic and the hedging demands implied by the robust strategy can be quite different from those implied by the Bayesian strategy.

Given a nondegenerate prior, large differences in continuation values under the VAR and IID submodels of stock returns occur when the predictive variable takes relatively high or low values, causing large differences between the expected returns under the two submodels. If the submodel that delivers a significantly worse outcome has a small Bayesian probability, then the distorted belief is very sensitive to small changes in the Bayesian posterior, inducing a large negative hedging demand against model uncertainty. This negative hedging demand lowers stock allocation significantly. For example, when the predictive variable takes a large value and the Bayesian probability of the VAR model is high, a small shift of the Bayesian belief towards the VAR model following a high realized stock return causes a much larger shift of the distorted belief. Thus, the negative hedging demand against model uncertainty under the robust strategy is much larger than that under the Bayesian strategy. Consequently, an ambiguity-averse investor's stock demand may be only half as much as the Bayesian investor's or less, and is even lower than what is delivered under the IID strategy.

An important finding of our paper is that the robust and the Bayesian strategies may deliver different market timing behavior and different stock allocations over time, both quantitatively and qualitatively. First, take a sufficiently small prior probability of the IID model as given. The stock allocation rises with the predictive variable under the Bayesian strategy but declines with it under the robust strategy for a wide range of values of the predictive variable. Second, take a sufficiently

large value of the predictive variable as given. If the investor believes that the stock return follows the VAR model for sure, then he would invest all his wealth in the stock. According to the Bayesian approach, the investor's stock allocation should decrease monotonically as his prior probabilities of the IID model rises. In contrast, we show that a very small prior probability of the IID model can lead an ambiguity-averse investor to decrease his stock allocation sharply and then to increase it gradually as the prior probability of the IID model rises. The large negative hedging demand against model uncertainty under the robust strategy plays a key role in these two results.

To evaluate the welfare cost of adopting the Bayesian strategy for an ambiguity-averse investor, we compute the wealth compensation that leaves him indifferent between adopting the Bayesian and the robust strategies. We find that welfare costs depend crucially on the values of the predictive variable and the prior probabilities. They are large when the predictive variable takes large values and the prior probability of the VAR model is large. In this case, the welfare costs are more than 15 percent of initial wealth.

We emphasize that our findings of large welfare costs and large differences between the Bayesian and robust strategies are empirically relevant and apply to ambiguity-averse investors with strong priors about the VAR model of stock returns. This model seems to be favored in the data, but there is still a small Bayesian probability that the IID model is on the table in a finite sample of data. In addition, the dividend-price ratio – the predictive variable used in our study – rose in recent years, especially during the recent recession.

Our paper is related to a large literature on the portfolio choice problem (see Campbell and Viceira (2002) and Wachter (2010) for a survey). In addition to the papers cited above, other papers using the Bayesian framework include Brennan (1998), Brandt, Goyal, Santa-Clara, and Stroud (2005), Detemple (1986), Dothan and Feldman (1986), Gennotte (1986), Gollier (2004), and Veronesi (1999), among others. These papers often study parameter uncertainty and do not consider investors' aversion to model uncertainty.

Our paper is more closely related to the literature on applications of ambiguity aversion preferences to the study of the portfolio choice problem (e.g., Cao, Wang, and Zhang (2005), Garlappi, Uppal, and Wang (2007), Maenhout (2004), and Uppal and Wang (2003)). This literature typically applies either the multiple-priors approach or the robust control approach. Some papers use one of these approaches to study equilibrium asset prices (e.g., Anderson, Hansen, and Sargent (2003), Chen and Epstein (2002), Epstein and Miao (2003), and Epstein and Wang (1994), Liu, Pan, and Wang (2005)). Boyle, Garlappi, Uppal, and Wang (2010) and Cao, Han, Hirshleifer, and Zhang (2011) use other models of ambiguity. All these papers do not allow for learning.

Epstein and Schneider (2007) and Miao (2009) introduce learning to the recursive multiple-priors model. Unlike the present paper, they study a portfolio choice problem in which investors are ambiguous about the mean stock return. Campanale (2011) also applies the multiple-priors approach to quantitatively explain the stock market participation rates. Hansen (2007) and Hansen

and Sargent (2007a,b) develop models of learning in the robust control framework. Hansen and Sargent (2010) apply this framework to the study of the equilibrium price of model uncertainty. They emphasize that fragile beliefs cause time-varying uncertainty premium. They refer to fragile beliefs as responsiveness of pessimistic probabilities to the arrival of news, as determined by the state dependent value functions that define what the consumer is pessimistic about. In our partial equilibrium model, these fragile beliefs have important impact on portfolio choice decisions. In a general equilibrium setup, Ju and Miao (2012) use the generalized recursive ambiguity utility model to study the implications of fragile beliefs for asset pricing.

To the best of our knowledge, the present paper provides a first dynamic portfolio choice model in which investors face a model selection problem using the generalized recursive ambiguity utility model. As discussed in Hayashi and Miao (2011) and Ju and Miao (2012), this utility model includes some other models of ambiguity as special cases, e.g., the recursive expected utility model of Epstein and Zin (1989), the recursive smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2009), the recursive multiple-priors model of Epstein and Wang (1994) and Epstein and Schneider (2003), and the robust control model of Hansen and Sargent (2001, 2007b).

In particular, when the ambiguity aversion parameter approaches infinity, our generalized ambiguity model approaches the limit of a version of the multiple-priors utility model. In this case, the investor follows the worst-case scenario by adopting either the IID strategy when the predictive variable is sufficiently large, or the VAR strategy when the predictive variable is sufficiently small. This portfolio rule is the extreme case of our model.

The rest of the paper proceeds as follows. Section 2 presents the recursive ambiguity model. Section 3 presents an ambiguity-averse investor's decision problem. Section 4 conducts calibration. Section 5 analyzes dynamic asset allocations. Section 6 conducts welfare costs analysis. Section 7 concludes. Appendices collect proofs and numerical methods.

2. Recursive Ambiguity Preferences

In this section, we introduce the recursive ambiguity utility model adopted in our paper. In a static setting, this utility model delivers essentially the same functional form that has appeared in some other papers, e.g., Chew and Sagi (2008), Ergin and Gul (2009), Klibanoff, Marinacci, and Mukerji (2005), Nau (2006), and Seo (2009).[4] These papers provide various axiomatic foundations and interpretations. Our adopted dynamic model is based on Ju and Miao (2012) and is axiomatized by Hayashi and Miao (2011). It is closely related to Klibanoff, Marinacci, and Mukerji (2005, 2009). Here we focus on the utility representation and refer the reader to the preceding papers for axiomatic foundations.

---

[4]See Epstein (2010) for a recent critique of this model and Klibanoff, Marinacci and Mukerji (2012) for a reply. Also see Hayashi and Miao (2011) for a related discussion.

2.1.   Utility

We start with a static setting in which a decision maker's ambiguity preferences over consumption are represented by the following utility function:

$$v^{-1}\left(\int_{\Pi} v \circ u^{-1}\left(\int_{S} u\left(C\right) d\pi\right) d\mu\left(\pi\right)\right), \ \forall C : S \to \mathbb{R}_{+}, \tag{1}$$

where u and v are increasing functions and μ is a subjective prior over the set Π of probability measures on S that the decision maker thinks possible. When we define $\phi = v \circ u^{-1}$, the utility function in (1) is ordinally equivalent to the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005):

$$\mathbb{E}_{\mu}\phi\left(\mathbb{E}_{\pi}u\left(C\right)\right). \tag{2}$$

A key feature of this model is that it achieves a separation between ambiguity, identified as a characteristic of the decision maker's subjective beliefs, and ambiguity attitude, identified as a characteristic of the decision maker's tastes. Specifically, ambiguity is characterized by properties of the subjective set of measures Π. Attitudes towards ambiguity are characterized by the shape of ϕ, while attitudes towards pure risk are characterized by the shape of u. In particular, the decision maker displays risk aversion if and only if u is concave, while he displays ambiguity aversion if and only if ϕ is concave or, equivalently, if and only if v is a concave transformation of u. Note that there is no reduction between μ and π in general. It is this irreducibility of compound distribution that captures ambiguity (Segal (1987)). When ϕ is linear, the decision maker is ambiguity neutral and the smooth ambiguity model reduces to the standard expected utility model.

We now embed the static model (1) in a dynamic setting. Time is denoted by $t = 0, 1, 2, ..., T$, where T is finite. The state space in each period is denoted by S. At time t, the decision maker's information consists of history $s^{t} = \{s_{0}, s_{1}, s_{2}, ..., s_{t}\}$ with $s_{0} \in S$ given and $s_{t} \in S$. The decision maker ranks adapted consumption plans $C = (C_{t})_{t\geq 0}$, where C_{t} is a measurable function of s^{t}. The decision maker is ambiguous about the probability distribution on the full state space S^{T}. This uncertainty is described by an unobservable parameter z in the space Z. The parameter z can be interpreted in several different ways. It could be an unknown model parameter, a discrete indicator of alternative models, or a hidden state that evolves over time in a regime-switching process.

The decision maker has a prior μ_{0} over the parameter z. Each parameter z gives a probability distribution π_{z} over the full state space. The posterior μ_{t} and the conditional likelihood can be obtained by Bayes' Rule. Following Jun and Miao (2012), we adopt the specification:

$$u\left(c\right) = \frac{c^{1-\gamma}}{1-\gamma}, \ \gamma > 0, \neq 1, \tag{3}$$

$$v\left(x\right) = \frac{x^{1-\eta}}{1-\eta}, \ \eta > 0, \neq 1, \tag{4}$$

and consider the following homothetic recursive ambiguity utility function:

$$V_t\left(C\right) = \left[C_t^{1-\rho} + \beta\left\{v^{-1}\mathbb{E}_{\mu_t}v \circ u^{-1}\mathbb{E}_{\pi_{z,t}}\left[u\left(V_{t+1}\left(C\right)\right)\right]\right\}^{1-\rho}\right]^{\frac{1}{1-\rho}}, \quad V_{T+1} = 0 \tag{5}$$

where $\beta \in (0,1)$ is the subjective discount factor, $1/\rho$ is the elasticity of intertemporal substitution (EIS), and γ and η are the coefficients of constant relative risk aversion and ambiguity aversion, respectively. If $\eta = \gamma$, the decision maker is ambiguity neutral and (5) reduces to the standard time-additive expected utility model. In this case, the posterior μ_t and the likelihood distribution $\pi_{z,t}$ can be reduced to a *predictive distribution*, which is the key idea underlying the Bayesian analysis. The decision maker displays ambiguity aversion if and only if $\eta > \gamma$. The coefficient of relative ambiguity aversion may be measured by $|(\eta - \gamma)/(1-\gamma)|$, which is the coefficient of relative risk aversion of $\phi\left(x\right) = v \circ u^{-1}\left(x\right) = \left[(1-\gamma)x\right]^{\frac{1-\eta}{1-\gamma}}/\left(1-\eta\right)$, $x \in \mathbb{R}$.

When the decision maker displays infinite ambiguity aversion ($\eta \to \infty$), we deduce from Klibanoff, Marinacci, and Mukerji (2005) that (5) converges to a version of the recursive multiple-priors model of Epstein and Schneider (2007):

$$V_t\left(C\right) = \min_z\left\{C_t^{1-\rho} + \beta\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}\left(C_{t+1}\right)\right]^{\frac{1}{1-\gamma}}\right\}^{\frac{1}{1-\rho}}. \tag{6}$$

In this case, the decision maker has multiple priors with Dirac measures and a single likelihood.

We may alternatively interpret the utility model defined in (5) as a model of robustness in which the decision maker is concerned about model misspecification, and thus seeks robust decision making. Specifically, each distribution π_z describes an economic model. The decision maker is ambiguous about which is the right model specification. He has a subjective prior μ_0 over alternative models. He is averse to model uncertainty, and thus evaluates different models using a concave function v. We may also interpret u and v in (5) as describing source-dependent risk attitudes (Chew and Sagi (2008)). That is, u captures risk attitudes for a given model distribution π_z and v captures risk attitudes towards model uncertainty.

2.2. How Large is Ambiguity Aversion Parameter?

Any new utility model other than the standard expected utility model will inevitably introduce some new parameters. A natural question is: How does one calibrate these parameters? In general, there are two approaches. First, one may derive equilibrium implications using the new utility model, and then use market data to estimate preference parameters by matching moments or using other econometric methods (e.g., Hansen and Singleton (1982)). Second, one may use experimental or field data to estimate the new preference parameters, like the standard way to elicit the risk aversion parameter. In our recursive ambiguity utility model (5), the new parameter is the ambiguity aversion parameter η. We will follow the second approach to calibrate this parameter in the static

setting (1).[5] We acknowledge that this calibration strategy from static experimental evidence might also be to some extent inconsistent with the dynamic setting with learning and predictability (see Leippold et al. (2008) and Ju and Miao (2012)).

We elicit the ambiguity aversion parameter by introspection using thought experiments related to the Ellsberg Paradox. Consider the following experiment similar to that in Halevy (2007).[6] Suppose there are two urns. One urn contains 50 black balls and 50 white balls. The other urn contains 100 balls, either all black or all white. But the exact composition is unknown to the subjects. Subjects are asked to place a bet on the color of the ball drawn from each urn. The bet on the second urn is placed before the color composition is known. If a bet on a specific urn is correct, the subjects win a prize of d dollars. Otherwise, the subjects do not win or lose anything. The experiments reported in Halevy (2007) show that most subjects prefer to bet on the first urn over the second urn. Halevy (2007) also uses the Becker-DeGroot-Marschack mechanism to elicit the certainty equivalent of a bet. As a result, one can compute the ambiguity premium as the difference between the certainty equivalents of the bet on the first and the second urns. We can then use the ambiguity premium to calibrate the ambiguity aversion parameter η.

Formally, we define the ambiguity premium as

$$
u^{-1}\left(\int_{\Pi}\int_S u\left(c\right)d\pi d\mu\left(\pi\right)\right) - v^{-1}\left(\int_{\Pi} v\left(u^{-1}\left(\int_S u\left(c\right)d\pi\right)\right)d\mu\left(\pi\right)\right). \tag{7}
$$

We then evaluate the bet in the previous experiment using the following parametric form: Let u and v be given by (3) and (4), respectively. Let w be the decision maker's wealth level. Suppose the subjective prior $\mu = (0.5, 0.5)$ for the bet.[7] For the bet on the second urn, Π has two probability measures over the ball color: $(0, 1)$ and $(1, 0)$. We then derive the ambiguity premium as

$$
\left(0.5\left(d+w\right)^{1-\gamma} + 0.5w^{1-\gamma}\right)^{\frac{1}{1-\gamma}} - \left(0.5\left(d+w\right)^{1-\eta} + 0.5w^{1-\eta}\right)^{\frac{1}{1-\eta}}, \tag{8}
$$

for $\eta > \gamma$. We may express the ambiguity premium as a percentage of the expected value of the bet $(d/2)$. Clearly, the size of the ambiguity premium depends on the size of a bet or the prize-wealth ratio d/w. Table 1 reports the ambiguity premium for various parameter values. Panel A considers the prize-wealth ratio of 1%. Panel B considers a smaller bet, with the prize-wealth ratio of 0.5%.

[Insert Table 1 Here.]

---

[5] Anderson, Hansen, and Sargent (2003) advocate to use model detection error probabilities to calibrate the absolute ambiguity aversion parameter θ for $\phi\left(x\right) = -e^{-\frac{x}{\theta}}$ in equation (2). They interpret θ as a robustness parameter. Because our model is different from their model, we have not followed their calibration approach.

[6] See Strzalecki (2010) for a similar experiment. In an axiomatized model, he suggests the same approach as ours to calibrate the ambiguity aversion parameter.

[7] Strictly speaking the bet deals with objective lotteries and the subjective probability measure may not be the same as the objective measure. Seo's (2009) utility model can accommodate the bet discussed in the paper. His utility model gives the same expression as (8) for the ambiguity premium.

Camerer (1999) reports that the ambiguity premium is typically in the order of 10-20 percent of the expected value of a bet in the Ellsberg Paradox type experiments. Halevy (2007) finds a similar value. Table 1 Panel A shows that the implied ambiguity premium falls in this range when the ambiguity aversion parameter η is in the range of 50-90 and when the risk aversion parameter γ is between 0 and 10. Our calibration depends crucially on the size of the bet. In experimental studies, researchers typically consider small bets. For example, Halevy (2007) considers the prize money of 2 or 20 Canadian dollars. It is likely that these prizes account for a very small fraction of a subject's wealth. In Panel B, when the prize-wealth ratio drops to 0.5%, even larger values of η are needed to match the ambiguity premium from experimental studies. In our quantitative study below, we focus on $\gamma \in \{2, 5, 10\}$. Based on the results from Table 1, we take three values (60, 80, 100) for η.

3. Decision Problem

We consider an investor's consumption and portfolio choice problem in a finite-horizon discrete-time environment. Time is denoted by $t = 0, 1, ..., T$. The investor is endowed with initial wealth W_0 in period zero, and his only source of income is his financial wealth. In each period t, he decides how much to consume and how much to invest in the financial markets. We assume that there is no bequest motive, so the investor consumes all his wealth $C_T = W_T$ in period T.

3.1. Investment Opportunities

There are two tradeable assets: a risky stock and a risk-free bond. The stock has gross real stock return R_{t+1} from t to $t + 1$. The risk-free bond has a constant gross real return R_f each period. Define log returns $r_{t+1} = \log(R_{t+1})$ and $r_f = \log(R_f)$. Observing data of the risk-free rate, the stock return and a predictive variable x_t, the investor faces the following two model specifications:

- Model 1 (IID):

$$r_{t+1} - r_f = m_1 + \varepsilon^r_{1,t+1}, \tag{9}$$

$$x_{t+1} = \rho_1 x_t + \varepsilon^x_{1,t+1}, \tag{10}$$

  where the expected return (m_1) is constant, and $\varepsilon_{1,t+1} = \left[\varepsilon^r_{1,t+1}, \varepsilon^x_{1,t+1}\right]'$ is normally distributed white noise with mean zero and covariance matrix:

$$\Omega_1 = \begin{bmatrix} (\sigma^r_1)^2 & \sigma^{rx}_1 \\ \sigma^{rx}_1 & (\sigma^x_1)^2 \end{bmatrix}. \tag{11}$$

- Model 2 (VAR):

$$r_{t+1} - r_f = m_2 + bx_t + \varepsilon^r_{2,t+1}, \tag{12}$$

9

$$x_{t+1} = \rho_2 x_t + \varepsilon_{2,t+1}^x, \tag{13}$$

where the conditional expected return $(m_2 + bx_t)$ varies with the predictive variable, and $\varepsilon_{2,t+1} = \left[\varepsilon_{2,t+1}^r, \varepsilon_{2,t+1}^x \right]'$ is normally distributed white noise with mean zero and covariance matrix

$$\Omega_2 = \begin{bmatrix} (\sigma_2^r)^2 & \sigma_2^{rx} \\ \sigma_2^{rx} & (\sigma_2^x)^2 \end{bmatrix}. \tag{14}$$

Assume that $\varepsilon_{1,t+1}$ is independent of $\varepsilon_{2,t+1}$. We estimate the parameters of both models using the same historical data, which implies $m_1 = m_2$, $\rho_1 = \rho_2$ and $\sigma_1^x = \sigma_2^x$. Hence, we will drop the subscripts for m, ρ and σ^x in the remainder of the paper. But generally $\sigma_1^{rx} \neq \sigma_2^{rx}$ and $\sigma_1^r \neq \sigma_2^r$ so that the above two model specifications are not nested.

More generally, x_t may be a vector of predictive variables. In our empirical application in this paper, we will focus on the cases with a single predictive variable. The investor faces model uncertainty because he is concerned that both of the above two models of stock returns may be misspecified. He does not know which of these two models generates the data. He can learn about the true model by observing past data. During the process of learning, he is averse to model uncertainty. To capture his aversion to model uncertainty, we adopt the recursive ambiguity model presented in Section 2 and assume that the investor's utility function is given by (5).

Alternatively, one may assume a regime switching structure in which the degree of return predictability is time-varying and agents learn about the regime over time. For example, Veronesi (1999, 2000) and Ju and Miao (2012) assume that aggregate consumption follows a regime switching process in general equilibrium models. By contrast, this paper assumes that the agent faces two submodels of stock returns in a partial equilibrium framework and has ambiguous beliefs about which one is the true model.

One may also assume that both of the two submodels of stock returns contain predictability, but with different properties in the error distribution of stock returns and/or dividend yields.[8] To the extent that the two submodels would deliver large difference in continuation value, the results of the paper will likely hold overall.

3.2. Bayesian Posterior Dynamics

Let $\mu_t = \Pr\left(z = 1 | s^t\right)$ denote the posterior probability that Model 1 is the true model for the return process, given the history of data $s^t = \{(r_0, x_0), (r_1, x_1) ..., (r_t, x_t)\}$. By Bayes' Rule, we can derive the evolution of μ_t :

$$\mu_{t+1} = \frac{\mu_t L_{1,t+1}}{\mu_t L_{1,t+1} + (1 - \mu_t) L_{2,t+1}}, \tag{15}$$

---

[8]See Camponovo et al. (2012) for a justification of this approach.

where for $s_{t+1} = [r_{t+1}, x_{t+1}]'$,

$$L_{z,t+1} = \frac{1}{2\pi |\Omega_z|^{1/2}} \exp\left[-\frac{1}{2}(s_{t+1} - m_{z,t})' \Omega_z^{-1} (s_{t+1} - m_{z,t})\right], \quad z = 1, 2 \tag{16}$$

and

$$m_{1,t} = [r_f + m, \rho x_t]', \tag{17}$$

$$m_{2,t} = [r_f + m + bx_t, \rho x_t]'. \tag{18}$$

The intuition for how the investor updates his Bayesian beliefs after observing the data of the predictive variable and the stock return is as follows. The expected return is constant according to the IID model, but it depends on the predictive variable in the VAR model. Assume that the volatilities of returns are similar in the two models (which is true in our estimation below). If the predictive variable is above average (i.e., $x_t > 0$), the VAR model will predict above average returns. A high realized return will be more likely in the VAR model than in the IID model. Thus, the observation of a high stock return makes the investor revise downward his belief about the IID model (μ_{t+1}). However, if the predictive variable is below average (i.e., $x_t < 0$), then the observation of high stock return is more consistent with the IID model, causing the investor to revise μ_{t+1} upward. This updating process is important for understanding the hedging demand analyzed in Section 5.1.

3.3. Optimal Consumption and Portfolio Choice

Let W_t and ψ_t denote respectively the wealth level and the portfolio share of the stock in period t. We can then write the investor's budget constraint as

$$W_{t+1} = R_{p,t+1} (W_t - C_t), \tag{19}$$

where

$$R_{p,t+1} = R_{t+1}\psi_t + R_f (1 - \psi_t)$$

is the portfolio return. We suppose that there are short-sale and margin restrictions such that $\psi_t \in [0, 1]$. Otherwise, wealth and consumption may be negative when ψ_t is negative or larger than 1 because R_{t+1} can go to positive infinity or zero. The investor's problem is to choose a consumption plan $\{C_t\}_{t=0}^T$ and a portfolio plan $\{\psi_t\}_{t=0}^T$ so as to maximize his utility given by (5). We derive the investor's decision problem using dynamic programming. In each period t, the investor's information may be summarized by three state variables: wealth level W_t, the predictive variable x_t, and the Bayesian belief μ_t. Let $J_t(W_t, x_t, \mu_t)$ denote the value function. Then it satisfies

the Bellman equation:

$$J_t\left(W_t, x_t, \mu_t\right) = \max_{C_t>0, \psi_t \in [0,1]} \left[C_t^{1-\rho} + \beta \left\{ \mu_t \left(\mathbb{E}_{1,t} \left[J_{t+1}^{1-\gamma}\left(W_{t+1}, x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right. \right. \tag{20}$$

$$\left. \left. + (1-\mu_t) \left(\mathbb{E}_{2,t} \left[J_{t+1}^{1-\gamma}\left(W_{t+1}, x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right\}^{\frac{1-\rho}{1-\eta}} \right]^{\frac{1}{1-\rho}},$$

subject to the budget constraint (19), the dynamics of x_t (10) or (13), and the Bayesian belief dynamics (15). Here, $\mathbb{E}_{1,t}$ is the conditional expectation operator conditioned on information available in period t, when the IID model (Model 1) is the true model for the stock return r_{t+1}. In this case, we substitute equations (9)-(10) for (r_{t+1}, x_{t+1}) into (15), and then substitute the resulting expression for μ_{t+1} into $\mathbb{E}_{1,t}\left[J_{t+1}\left(W_{t+1}, x_{t+1}, \mu_{t+1}\right)\right]$. Similarly, $\mathbb{E}_{2,t}$ is the conditional expectation operator conditioned on information available in period t, when the VAR model (Model 2) is the true model for the stock return r_{t+1}. In this case, we substitute equations (12)-(13) for (r_{t+1}, x_{t+1}) into (15), and then substitute the resulting expression for μ_{t+1} into $\mathbb{E}_{2,t}\left[J_{t+1}\left(W_{t+1}, x_{t+1}, \mu_{t+1}\right)\right]$.

In Appendix A (also see Ju and Miao (2012)), we derive the following Euler equation when the optimal portfolio weight ψ_t^* is an interior solution in $(0,1)$:

$$\mathbb{E}_t\left[M_{z,t+1}\left(R_{t+1} - R_f\right)\right] = 0, \quad t = 0, 1, \ldots, T-1, \tag{21}$$

where $M_{z,t+1}$ denotes the pricing kernel for the recursive smooth ambiguity utility model, which is given by:

$$M_{z,t+1} = \left(\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1} \left(\mathbb{E}_{z,t} \left[\left(\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}. \tag{22}$$

In period T, the investor consumes all his wealth $C_T = W_T$ and the portfolio choice has no consequence. When $\gamma = \eta$, the investor is indifferent to ambiguity and the model reduces to the standard expected utility model. We then obtain the familiar Euler equation for the power utility function. When the investor is averse to model ambiguity, the standard pricing kernel is distorted by a multiplicative factor in (22). To interpret this distortion, we normalize the multiplicative factor and show in Appendix A that the Euler equation can be written as:

$$0 = \hat{\mu}_t \mathbb{E}_{1,t} \left[\left(\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1} \left(R_{t+1} - R_f \right) \right] \tag{23}$$

$$+ (1-\hat{\mu}_t) \mathbb{E}_{2,t} \left[\left(\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1} \left(R_{t+1} - R_f \right) \right],$$

where $\hat{\mu}_t$ is given by:

$$\hat{\mu}_t = \frac{\mu_t \left(\mathcal{R}_t^1(J_{t+1})\right)^{-(\eta-\gamma)}}{\mu_t \left(\mathcal{R}_t^1(J_{t+1})\right)^{-(\eta-\gamma)} + (1-\mu_t) \left(\mathcal{R}_t^2(J_{t+1})\right)^{-(\eta-\gamma)}}, \tag{24}$$

and the term

$$\mathcal{R}_t^i(J_{t+1}) \equiv \mathbb{E}_{i,t} \left[J_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}$$

gives the certainty equivalent continuation value associated with submodel i. We interpret $\hat{\mu}_t$ as the distorted belief about the IID model.

Equation (23) implies that an ambiguity-averse investor makes decisions as if he has distorted beliefs $\hat{\mu}_t$ in a Bayesian framework. We shall emphasize that $\hat{\mu}_t$ is endogenous (preference dependent) in our model and cannot be generated from a Bayesian posterior according to (15) using any prior μ_0 given the history of data s^t. In addition, the pricing kernel (22) cannot be generated from any Bayesian model. Thus, our model cannot be reduced to a Bayesian framework and is not equivalent to any recursive expected utility model.

Equation (24) is key to understanding how an ambiguity-averse investor's belief is distorted. We rewrite it as:

$$\hat{\mu}_t = \frac{\mu_t}{\mu_t + (1-\mu_t) \left(\frac{\mathcal{R}_t^2(J_{t+1})}{\mathcal{R}_t^1(J_{t+1})}\right)^{-(\eta-\gamma)}}. \tag{25}$$

Suppose that the investor obtains higher certainty equivalent continuation value when data are generated by the VAR model than by the IID model, i.e., $\mathcal{R}_t^2(J_{t+1}) > \mathcal{R}_t^1(J_{t+1})$. If $\eta > \gamma$, then equation (25) implies that $\hat{\mu}_t > \mu_t$. That is, an ambiguity-averse investor attaches more weight on the IID model than does a Bayesian investor. The opposite is true when the IID model generates a higher certainty equivalent continuation value. Thus, the ambiguity-averse investor expresses his concerns about model misspecification by slanting his beliefs towards the "worse model," the one that implies a lower certainty equivalent continuation value.

Equation (25) also shows that the amount of distortion in beliefs is large when the difference between the certainty equivalent continuation values under the two submodels of stock returns is large (holding μ_t fixed). This case happens when the conditional expected return under the VAR model is far from the unconditional mean; that is, when x_t takes large positive or negative values.

The amount of distortion is especially large when the Bayesian belief also favors the submodel with a higher certainty equivalent continuation value. For example, when x_t takes a large positive value and μ_t is small, the Bayesian belief attaches a high probability to the VAR model of the stock return which gives a much higher expected stock return than the IID model. In this case, an ambiguity-averse investor is concerned that the IID model might be the true model of the stock return, which may generate a large utility loss. He then pessimistically slants his belief heavily toward the IID model, generating a high $\hat{\mu}_t$. When x_t takes a large negative value and μ_t is large,

the ambiguity-averse investor fears that the VAR model is the true model of the stock return and hence adjusts his belief about the IID model downward aggressively. In both cases, the large distortions in beliefs lead to large differences in investment strategies between an ambiguity-averse investor and a Bayesian investor.

Does a more ambiguity-averse investor invest less in the stock? Not necessarily, as shown by Gollier (2011) in a static portfolio choice model. The intuition is simple. The effect of ambiguity aversion is reflected by a pessimistic distortion of beliefs about the model of the stock return process. A change of the subjective distribution of asset payoffs may not induce the investor to demand the asset in a monotonic way. For example, Rothchild and Stiglitz (1971) show that an increase in the riskiness of an asset's payoffs does not necessarily reduce the demand for this asset by all risk-averse investors. In our dynamic portfolio choice problem, we cannot derive analytical results of an ambiguity-averse investor's portfolio choice, we thus use numerical solutions to conduct comparative static analyses.

4. Calibration

In order to provide quantitative predictions, we need to calibrate parameters and solve the calibrated model numerically. In Section 4.1, we discuss how to estimate models of stock returns specified in Section 3.1. In Section 4.2, we then calibrate preference parameters. In Appendix B, we present the numerical method.

4.1. Data and Model Estimation

There is a large literature documenting that stock returns are forecastable (see references cited in Campbell and Thompson (2008) and Welch and Goyal (2008)). The predictive variables include valuation ratios, payout ratios, short rates, slope of the yield curve, consumption-wealth-income ratio, and other financial variables.

Researchers typically use a VAR system as in (12)-(13) to capture predictability. We estimate this system and the IID model (9)-(10) using annual data for the U.S. stock market over the period 1927-2010. For stock returns, we use the log returns (cum-dividend) of the CRSP value-weighted market portfolio (including stocks from the NYSE, AMEX and NASDAQ). We roll over the 90 Day T-Bill return series from the CRSP Fama Risk-Free Rate file to compute the annual risk-free rates. All nominal quantities are deflated using the Consumer Price Index (CPI). We find the mean real risk-free rate $r_f = 0.0078$. Panel A of Figure 1 plots the realized excess log returns ($r_t - r_f$) over the sample.

[**Insert Figure 1 Here.**]

Following the portfolio choice literature (e.g. Barberis (2000), Campbell and Viceira (2002),

14

Xia (2001)), we choose the dividend yield as the predictive variable. We take the demeaned log dividend yield (ldy) as x_t in the regression. We compute it as the log difference between cum- and ex-dividend returns of the CRSP value-weighted market portfolio. The demeaned series is plotted in Panel A of Figure 1. This panel reveals that the log dividend yield dropped significantly during the 1990s. Lettau and Van Nieuwerburgh (2008) argue that this change is a structural break in the mean of dividend yields.

We estimate both the IID and the VAR models using the maximum likelihood method, with the restriction that the unconditional means of the excess log stock returns and payout yields equal their sample means. Table 2 reports the estimation results. We take the point estimates as our parameter values in the IID and the VAR models. We use these parameter values to conduct numerical analyses below.

[**Insert Table 2 Here.**]

Table 2 shows that the estimates of the persistence parameter ρ of the predictive variables are identical to the OLS estimates and hence are identical in both the IID model and the VAR model. In addition, the estimates of the volatility parameter σ^x are also identical in these two models. In the IID model, even though the expected excess stock return is constant over time, the innovation of the excess stock return is negatively correlated with that of the predictive variable.

In the VAR model, when using the log dividend yield (ldy) as the return predictor, we obtain results similar to those reported in the literature (e.g., Cochrane (2008)). The coefficient b is 0.122, with standard error 0.048. The R-squared is 6.26%. Moreover, the estimate of the coefficient b is sensitive to the sample period. When estimated using 30-year moving windows (see Figure 1 of Lettau and Van Nieuwerburgh (2008)), the coefficient fluctuates between 0 and 0.5, and drops substantially towards the late 1990s. These features highlight the statistical uncertainty confronting investors who try to use dividend yields to predict stock returns.

The expected excess returns generated by the predictor have three properties. First, the volatility of the expected excess return is high (20.25 percent). Second, the persistence of the expected returns is quite high. Since the predicted excess returns are assumed to be linear functions of the predictor, they inherit the persistence of the predictor. Third, the correlation between unexpected returns and innovations in expected returns is negative (-0.6223). The negative correlation means that stock returns are mean-reverting: An unexpected high return today reduces expected returns in the future, and thus high short-run returns tend to be offset by lower returns over the long run. This negative correlation is what generates intertemporal hedging demand for the stock by long-term investors. The predictive variable summarizes investment opportunities. The correlation between the stock return and the predictive variable measures the ability of the stock to hedge time variation in investment opportunities.

In Panel B of Figure 1, we plot the Bayesian posterior probabilities of the IID model using the

historical data of stock returns and log dividend yields from 1927 to 2010. The prior in 1927 is set at $0.25, 0.5$ and 0.75, respectively. The three series of posterior probabilities all trend downward over time, suggesting that the data is overall more consistent with time-varying expected returns. For the log dividend yield, the posterior probability of the IID model is about 0.014 towards the end of the sample when the prior is 0.5, and the rise in posterior probabilities in the 1990s is consistent with the poor performance of the dividend yield as a predictor during that time.

Although Panel B of Figure 1 shows that historical data favor the VAR model over a long sample period from 1927-2010, there is a small probability that the IID model is on the table for a finitely-lived investor. In particular, the posterior about the IID model wanders in $(0, 1)$ and is still positive, when the prior of the IID model is 0.5. Since different model specifications imply drastically different dynamics of stock returns, concerns about model misspecification, sample biases, and out-of-sample performances will expose a finitely-lived investor to considerable model uncertainty. We will show in Section 6 that the welfare costs of ignoring model uncertainty is sizable, even though there is a small prior probability that the IID model is on the table.

Trojani et al. (2013) point out that a robust estimation approach is potentially important to estimate a model explicitly seen by ambiguity-averse investors as potentially misspecified. In particular, the difference in estimated utilities implied by the maximum likelihood and robust estimators can be as large as the difference in utility between an ambiguity-averse agent and an expected utility maximizer knowing the model parameters. Our paper's results below are robust to these features because they depend on the difference in continuation value under the IID and VAR submodels, which is likely large for the given sample when predictive regressions are estimated by robust methods.[9]

4.2. Preference Parameters

We need to assign values to preference parameters β, γ, ρ, and η. We set $\beta = 0.99$ so that it is approximately equal to $1/(1+r_f)$. We set $\gamma = 5$, $\rho = 0.5$, and $\eta = 80$ as the benchmark parameter values. For comparison, we also consider $\gamma \in \{2, 5, 10\}$. These values are commonly used in the macroeconomics and finance literature. Following Bansal and Yaron (2004), we set $\rho = 0.5$, implying EIS is equal to 2.[10] There is no independent study of the ambiguity aversion parameter η in the literature. We use the hypotetical experiment described in Section 2.2 to calibrate this parameter. As discussed there, we take $\eta \in \{60, 80, 100\}$. When $\eta = \gamma$, our model reduces to the standard Epstein-Zin model in the Bayesian framework. Finally, we consider a $T = 40$ years investment horizon.

---

[9]We thank an anonymous referee for making this point to us.

[10]We have also considered various other values of ρ and found the optimal stock allocation is not sensitive to changes in ρ. This result is available upon request. Campbell and Viceira (1999) find a similar result.

5. Dynamic Asset Allocation

In this section, we analyze how learning under ambiguity affects dynamic asset allocation. We first examine its effects on the hedging demand, and compares that with the hedging demand for a Bayesian investor with Epstein-Zin preferences. We then study how learning under ambiguity alters the market timing, uncertainty, and horizon effects often analyzed in the Bayesian framework. Following most papers in the portfolio choice literature, we focus on the case in which the dividend yield is the single predictive variable in this section.

[**Insert Figure 2 Here.**]

Before studying the portfolio implications, we first plot the distorted belief $\hat{\mu}_0$ as a function of the prior belief μ_0 and the predictive variable x_0 for an ambiguity-averse investor with a 40 year investment horizon. Consistent with the intuition discussed in Section 3, $\hat{\mu}_0$ is slanted upward in favor of the IID model when the VAR model predicts high expected returns (x_0 is large), and downward in favor of the VAR model when the VAR model predicts low expected returns. In contrast, there is relatively little distortion in beliefs when the predictive variable is close to its mean.

The most significant distortion in beliefs occurs in two regions: (i) when the prior probability of the VAR model is high (μ_0 close to 0) and the expected return according to the VAR model is also high (x_0 is large); (ii) when the prior probability of the VAR model is low (μ_0 is close to 1) and the expected return according to the VAR model is low (x_0 is small). In these regions, the distorted belief $\hat{\mu}_0$ is very sensitive to small changes in μ_0. As discussed earlier, these are the cases where the submodel that is deemed unlikely delivers particularly unfavorable outcomes relative to the other submodel. These results are crucial for understanding the investment strategy of the ambiguity-averse investor.

5.1. Portfolio Weights and Hedging Demand

As explained in Section 2, we can interpret our ambiguity model as a model of robustness. To distinguish from other popular investment strategies studied in the literature and in the analysis below, we refer to an ambiguity-averse investor's optimal investment strategy as the robust strategy. Let ψ_t^* be his optimal stock allocation in period t. We define ψ_t^M as his myopic demand for the stock, which is the optimal portfolio weight when the investor behaves myopically by choosing a stock allocation to maximize the utility derived from his wealth in the next period. We then define the ambiguity-averse investor's hedging demand as $\psi_t^H \equiv \psi_t^* - \psi_t^M$.

The top panel of Table 3 reports the total stock demand ψ_0^* of an investor with $T = 40$ years investment horizon and with various values of risk aversion and ambiguity aversion parameters (γ, η). Because ψ_0^* is a function of the state vector (μ_0, x_0), we also report the values of ψ_0^* at

various values of (μ_0, x_0). When $\gamma = \eta$, the investor is ambiguity neutral and our model reduces to the Epstein-Zin model in the Bayesian framework. In this case, we denote the total stock demand as ψ_0^B and refer to this investment strategy as the Bayesian strategy.

[**Insert Table 3 Here.**]

The bottom panel of Table 3 reports the stock demands when the stock return is described by the IID model (9)-(10) and the VAR model (12)-(13), respectively. The former investment strategy corresponds to the case with $\mu_0 = 1$ and is studied by Merton (1969, 1971) and Samuelson (1969). The latter corresponds to the case with $\mu_0 = 0$ and is similar to that derived in Campbell and Viceira (1999) with the difference that we have imposed short-sale and margin constraints to ensure nonnegative wealth. We refer to these two investment strategies as the IID strategy and the VAR strategy, respectively. As is well known from these studies, the stock demand under the IID strategy is constant over time. By contrast, the stock demand under the VAR strategy depends on the investment horizon and the predictive state variable. In particular, it increases with the predictive variable reflecting the market timing effect.

The top panel of Table 3 reveals that the risk aversion parameter γ has larger effects on the optimal stock allocation than the ambiguity aversion parameter η. The optimal stock allocation is very sensitive to the risk aversion parameter γ and decreases significantly when γ increases from 2 to 10. By contrast, the optimal stock allocation is less sensitive to the ambiguity aversion parameter η, especially when the investor's uncertainty about submodels is low (i.e. μ_0 is close to 0 or 1), and decreases with η for various values of the state variables (μ_0, x_0) considered in Table 3.

The difference between the robust and Bayesian strategies can be large. For example, an ambiguity-averse investor with the ambiguity aversion parameter $\eta = 80$ and the risk aversion parameter $\gamma = 5$ invests 42.6% of his wealth in the stock, which is about 20 percentage points less than that for a Bayesian investor with Epstein-Zin utility with $\gamma = \eta = 5$, when they both assign identical prior probabilities, $\mu_0 = 0.1$, to the IID submodel of stock returns and when the demeaned predictive variable $x_0 = x_0^3 \equiv 0.4504$.[11]

Next, we turn to the analysis of the hedging demand. We decompose it into two components. The first hedge component is associated with changes in the predictive variable x_t. This component has been analyzed by Campbell and Viceira (1999) and Kim and Omberg (1996). Recall that the shock to the predictive variable is negatively correlated with the shock to future stock returns (i.e., $\sigma^{rx} < 0$). This negative correlation implies that stocks tend to have high returns when their expected future returns fall. Since the investor is normally long in stocks for large values of the predictive variable, a decline in the expected future returns represents a deterioration of the investment opportunity set. Since an investor with high risk aversion ($\gamma > 1$) wants to hedge the risk of deteriorating investment opportunities by holding assets that deliver increased wealth when

---

[11]One standard deviation of x_t is $\sigma^x / \sqrt{1 - \rho^2} = 0.4504$.

investment opportunities are poor, he has a positive hedging demand.[12] If the expected excess return (or x_t) becomes sufficiently negative, the investor has an incentive to short stocks so that the no short-sale constraint binds.

The second hedge component reflects the agent's incentive to hedge against model uncertainty (or changes in μ_t). This hedge component is negative and large when x_t takes large positive or negative values. The intuition is as follows. To hedge against the change in the investment opportunity set, the investor wants to sell (buy) assets with payoffs positively (negatively) correlated with it. When the investor observes a large positive value of x_t, an unexpectedly high stock return induces the investor to attach more weight on the VAR model as discussed in Section 3.2. The persistence in x_t then implies that returns are more likely to be high for a while. Thus, the future investment opportunity set is positively correlated with the stock return, which induces a negative hedging demand for the stock associated with model uncertainty. Conversely, if x_t takes a large negative value, an unexpectedly high stock return induces the investor to attach more weight on the IID model. This also represents an improvement in investment opportunities because the expected return under the IID model is higher than the VAR model when x_t takes a large negative value. As a result, the investor also has a negative hedging demand if x_t takes a large negative value.

Although a similar decomposition appears in the standard Bayesian analysis (e.g., Xia (2001)), ambiguity aversion affects both hedge components quantitatively, causing significant differences between the robust and the Bayesian strategies. The component of the hedge demand against model uncertainty is especially important in our model. As we have seen in Figure 2, concerns about model misspecification lead to a pessimistic distortion in beliefs. When x_t is large, an unexpectedly high stock return not only shifts the Bayesian belief μ_t towards 0, but also lowers the distorted belief $\hat{\mu}_t$. When μ_t is sufficiently small, the investor is highly concerned about model misspecification, which is reflected in the large and sensitive change of $\hat{\mu}_t$ with respect to small changes in μ_t (see Figure 2). Thus, a small change in μ_t induced by a change in stock returns can lead to a large change in the distorted belief $\hat{\mu}_t$, which amplifies the negative hedging demand against model uncertainty. Similarly, the hedging demand against model uncertainty is also amplified when x_t takes a large negative value and μ_t is close to 1. Thus, we expect the difference between the robust strategy and the Bayesian strategy to be most significant in these two cases.

[**Insert Table 4 Here.**]

To examine the hedging component quantitatively, we present the total hedging demand as a percentage of the total stock demand for various values of (γ, η) and for various values of (μ_0, x_0) in Table 4.[13] The rows with $\gamma = \eta$ correspond to the Bayesian strategy.

Table 4 reveals two interesting results. First, compared to the myopic demand, the hedging

---

[12]As discussed by Campbell and Viceira (1999) and Kim and Omberg (1996), an investor with low risk aversion $\gamma < 1$ has a different hedging behavior. Here, we do not study the case with $\gamma < 1$.

[13]We are unable to present the two hedge components separately, because they do not admit analytical expressions.

demand typically accounts for a small fraction of the total stock demand. In particular, for small values of the risk aversion parameter and large values of predictive variables (e.g., $\gamma = 2$ and $x_0 > 0$), the myopic demand under both the Bayesian and the robust strategies accounts for almost all the stock demand. The total hedging demands under both the Bayesian and the robust strategies are relatively large when μ_0 is small. In addition, they are negative when x_0 takes a large positive value and μ_0 is small, reflecting the fact that the hedge component associated with model uncertainty dominates the hedge component associated with changes in the predictive variable. The opposite result holds when x_0 takes a large negative value and μ_0 is large.

Second, when the investor attaches a high prior on the VAR model (e.g., $\mu_0 = 0.1$) and when the predictive variable takes a high value (e.g., $x_0 = 0.4504$), an ambiguity-averse investor may take a much larger short hedging position than a Bayesian investor. For example, the total hedging demand is -16.7% of the total stock demand for an ambiguity-averse investor with $\gamma = 5$ and $\eta = 80$, while it is 0.7%, 0.3% and -0.9 for a Bayesian investor with $\gamma = 5$, 5.764, and 10, respectively. For the Bayesian investor, the negative hedging demand against model uncertainty is dominated by the positive hedging demand against changes in the predictive variable. In addition, the ambiguity-averse investor reduces his short hedging position quickly as the prior on the IID model increases. For example, when μ_0 increases from 0.1 to 0.9 given $x_0 = x_0^3 = 0.4504$, the ambiguity-averse investor with $\gamma = 5$ and $\eta = 80$ reduces his short hedging position sharply from 16.7 to 0.5 percent of the total stock demand. By contrast, the Bayesian investor with $\gamma = \eta = 5$ changes his hedging position from 0.7 to -2.7 percent of the total stock demand.

5.2. Comparison with Epstein-Zin Utility

The fact that an ambiguity-averse investor chooses a more conservative investment strategy than a Bayesian investor under the same values of relative risk aversion and intertemporal elasticity of substitution is not surprising. However, it does beg the question of whether a high degree of risk aversion for a Bayesian investor will have the same effect on the stock allocation as does ambiguity aversion. We answer this question by computing the stock allocations for a Bayesian investor with Epstein-Zin utility whose risk aversion parameter is selected to match the average portfolio weight (about 33.3%) for an ambiguity-averse investor with $\gamma = 5$ and $\eta = 80$, conditional on $\mu_0 = 0.5$ (the average is taken with respect to the stationary distribution of x). The calibrated risk aversion parameter is 5.764.

The results are shown in the last row of the top panel of Table 3 and 4. Conditional on $\mu_0 = 0.5$, even though the initial average portfolio weight is identical, the robust strategy (with $\gamma = 5, \eta = 80$) recommends more conservative stock allocations than the Bayesian strategy when the dividend yield takes extreme values, e.g., x_0^1 or x_0^3, but more aggressive stock allocations for middle values of dividend yields. The same feature also applies when the probability of the IID model is low (e.g., when $\mu_0 = 0.1$). Results in Table 4 show that the cause of these differences is

the difference in the hedging demand.

To further demonstrate the different portfolio implications of the Epstein-Zin utility and the recursive smooth ambiguity utility under uncertainty about return predictability, we plot in Figures 3 and 4 the robust strategy with $\gamma = 5$ and $\eta = 80$ and the Bayesian strategy with Epstein-Zin utility under various values of the risk aversion parameter γ. We fix the EIS at $1/\rho = 2$ for all strategies (the portfolio allocation is not sensitive to different values of EIS).

[Insert Figures 3-4 Here.]

Figure 3 shows that, when compared to the Bayesian strategy that has the same average portfolio weight conditional on $\mu_0 = 0.5$ (achieved with the Epstein-Zin utility under $\gamma = 5.764$), the robust strategy recommends more conservative stock allocations for extreme values of dividend yields and more aggressive stock allocations for middle values of dividend yields than the Bayesian strategy when μ_0 is close to 0.5. As μ_0 moves away from 0.5, the differences in the two strategies become larger, with the robust strategy becoming more conservative than the Bayesian strategy whenever the submodel of returns that is considered less likely predicts significantly lower expected returns (i.e., when μ_0 is small and x_0 is large, or when μ_0 is large and x_0 is small).

Although a sufficiently high risk aversion parameter for the Epstein-Zin utility can reduce the portfolio weight to the same level as the robust strategy for a particular initial state (μ_0, x_0) or to the same initial average level, there are significant quantitative differences in the portfolio weights in some states. We get similar results in Figure 4 when we fix the dividend yield and examine the changes in portfolio weight against the beliefs. These numerical results show that one cannot quantitatively replicate the portfolio strategy of an ambiguity-averse investor by using a model with Epstein-Zin utility and higher risk aversion.

In addition, when μ_0 is small (see Panels A and B of Figure 3), there are cases where the stock allocation under the robust strategy decreases with the dividend yield, but the opposite is true under the Bayesian strategy in the Epstein-Zin model for various values of the risk aversion parameter. Also, when the dividend yield is high (see Panel D of Figure 4), the portfolio weight decreases monotonically with μ_0 for the Epstein-Zin utility with various values of γ, yet under the robust strategy it falls sharply for small values of μ_0 and then increases as μ_0 increases. These results are all due to the large negative hedging demand against model uncertainty analyzed before. Although we do not have a theoretical proof, this non-monotonic pattern in portfolio weight never happens in the Epstein-Zin model for a wide range of preference parameters we have examined.

5.3. The Market Timing Effect

An important implication of return predictability for the portfolio choice is market timing. That is, the optimal stock allocation may depend on the current value of the predictive variable. Figure 5 shows the market timing effect by plotting the optimal stock allocation against the predictive

variable for various values of prior probabilities of the IID model of the stock return process. We consider five investment strategies: the IID strategy, the VAR strategy, the Bayesian strategy with $\gamma = \eta = 5$, the robust strategy with $\gamma = 5$ and $\eta = 80$, and the multiple-priors strategy with $\gamma = 5$ and $\eta = \infty$.

As is well known, the IID strategy does not have any market timing effect. The VAR strategy advises the investor to invest more in the stock when the value of the predictive variable is higher. In particular, for intermediate values of the predictive variable, the stock demand is approximately linear, confirming the approximate analytical solution derived by Campbell and Viceira (1999). When the predictive variable takes sufficiently large values, the investor invests all his wealth in the stock because expected excess returns are too high. When the predictive variable takes sufficiently small values, the investor does not hold the stock because expected excess returns are too low.[14]

By contrast, the Bayesian strategy implies that the optimal stock allocation first increases with the predictive variable and then decreases with it. Xia (2001) obtains a similar result in the Bayesian framework with parameter uncertainty. The intuition is that the negative hedge component associated with model uncertainty dominates the positive hedge component associated with the predictive variable, when the predictive variable takes sufficiently high values.

Compared to the Bayesian strategy, our robust strategy is more conservative in the sense that it recommends the investor to invest less in the stock. In particular, the robust strategy curve is obtained by bending the Bayesian strategy curve downward. When the demeaned predictive variable takes low and negative values, an ambiguity-averse investor is more likely to not participate in the stock market than a Bayesian investor. A similar result appears in the multiple-priors model (e.g., Epstein and Schneider (2007)). When the demeaned predictive variable is close to zero, the IID and VAR models of the stock returns are similar, and hence deliver similar continuation values to the investor. As a result, the Bayesian and the robust strategies offer very similar portfolio advice to the investor.

When the demeaned predictive variable takes large positive values, the robust strategy recommends a smaller stock allocation than both the VAR strategy and the Bayesian strategy. This difference is particularly significant when μ_0 is small, because, as we discussed before, the distortion in beliefs is large when the predictive variable takes a large value and μ_0 is small, which leads to large (negative) hedging demand against model uncertainty. Panel A of Figure 5 shows that when the predictive variable x_0 is two standard deviations above the mean and when the prior probability attached to the IID model is 0.05, the stock allocation advised by the Bayesian strategy is more than twice as large as that advised by the robust strategy. Similarly, the stock allocation under the robust strategy is again significantly lower than that under the Bayesian strategy when the predictive variable takes a small value and μ_0 is large.

[Insert Figure 5 Here.]

---

[14]Recall that we have imposed the short-sale and margin (borrowing) constraints such that $\psi_t \in [0, 1]$.

Importantly, the robust strategy also implies qualitatively different market timing behavior than the Bayesian strategy in some states. Figure 5 shows that there is a region of states (μ_0 is not too high and x_0 is positive and not too large) such that the stock allocation under the Bayesian strategy increases with the predictive variable, while the stock allocation under the robust strategy decreases with it. This difference is particularly large for small μ_0. The intuition is that in those states the large negative hedging demand against model uncertainty under the robust strategy dominates the positive hedging demand against changes in the predictive variable, but the opposite is true under the Bayesian strategy.

Figure 5 also presents the investor's investment strategy when he displays infinite ambiguity aversion with his utility function given by the recursive multiple-priors model (6). This figure illustrates that a more ambiguity-averse investor does not necessarily invest less in the stock. The extremely ambiguity-averse investor invests according to the worst-case scenario. In particular, he does not invest in the stock for sufficiently low values of the predictive variable because at these values the VAR model of stock returns gives a lower continuation value. The investor invests according to the IID strategy for sufficiently high values of the predictive variable because at these values the IID model gives a lower continuation value. As a result, the portfolio rule of an investor with infinite ambiguity aversion is constant whenever the value of the predictive variable is sufficiently above average, which is qualitatively different from the robust strategy in our model. Only for intermediate values of the predictive variable, the investor times the market by increasing his stock allocations when the current value of the predictive variable increases.

5.4. The Uncertainty Effect

How does the investor's stock allocation change when he has different initial prior over the IID model of the stock return process? Figure 6 plots this uncertainty effect for the Bayesian strategy with $\gamma = \eta = 5$ and the robust strategy with $\gamma = 5$ and $\eta = 80$ at various values of the predictive variable. Panel A of this figure shows that under both strategies the investor does not invest in the stock when he believes that the stock return is more likely to follow the VAR model (i.e., μ_0 is small) and when the demeaned predictive variable takes a large negative value. In this case, the stock is likely to have a negative expected excess return, and hence the short-sale constraint binds. As the prior probability of the IID model rises, the investor starts to invest more in the stock, while the robust strategy advises less stock allocation than the Bayesian strategy. Panel B has a similar feature except that the short-sale constraint does not bind because the predictive variable takes a larger value.

Panel C plots the case where the demeaned predictive variable takes one standard deviation. It shows that under both the robust and the Bayesian strategies the investor starts with 100 percent allocation of initial wealth in the stock when the prior probability of the VAR model is 1. As the prior probability of the IID model rises, the investor starts to invest less in the stock. Again, the

robust strategy advises less stock allocation than the Bayesian strategy.

Panel D of Figure 6 shows that the investor invests all his wealth in the stock when he believes the stock return follows the VAR model ($\mu_0 = 0$) and when the predictive variable is two-standard-deviation above the mean. In this case, the expected excess return under the VAR model is much higher than under the IID model. Surprisingly, even if there is a very small prior probability that the stock return follows the IID model, an ambiguity-averse investor will decrease his stock allocation sharply from 100 percent to about 35 percent. This dramatic drop in the stock allocation is due to the fact that the investor becomes extremely concerned about model misspecification when the predictive variable takes a large value, and when the VAR model has a high Bayesian probability. In this case, the hedging demand against model uncertainty can be quite large and negative (see the discussion in Section 5.1 and Table 4). As the ambiguity-averse investor raises his prior beliefs about the IID model, he gradually reduces his short hedging positions and starts to invest more in the stock. This result is in sharp contrast to that obtained in the Bayesian framework: Conditional on a high value of the predictive variable, a Bayesian investor decreases his stock allocation monotonically when his prior probability of the IID model rises, which is mainly driven by the fact that the expected stock return falls as more probabilities are assigned to the lower IID return.

[**Insert Figure 6 Here.**]

5.5. The Horizon Effect

When stock returns are predictable, the optimal stock allocation depends on the investment horizon. Figure 7 presents the horizon effect for the VAR strategy, the Bayesian strategy with $\gamma = \eta = 5$, and the robust strategy with $\gamma = 5$ and $\eta = 80$, when we fix the belief at the value $\mu_t = 0.1$ and x_t at a value in $\{-0.4504, 0, 0.4504\}$ over time. Under the assumption of the VAR strategy, the investor has complete confidence that the stock return follows the VAR model. In this case, because the shocks to the expected returns and to the future returns are negatively correlated, the stock appears to be less risky for a longer investment horizon. Thus, the VAR strategy recommends the investor to invest more in the stock when he faces a longer investment horizon.

However, if the investor faces model uncertainty, the stock allocation may not be monotonically increasing in the investment horizon. Consistent with Xia's (2001) finding, the stock allocation under the Bayesian strategy may decrease with the investment horizon. This case happens when the predictive variable takes a high value as shown in panel C of Figure 7. The intuition is the following: The horizon effect depends crucially on the intertemporal hedging demand. As we discuss earlier, this hedging demand consists of two components having effects on opposite directions. When the investment horizon is longer, the hedging component associated with model uncertainty is more important. Because this hedge component is negative when the predictive variable takes a large positive value, the investor invests less in the stock when he has a longer investment horizon.

24

Figure 7 reveals that the robust strategy implies a similar horizon effect to that under the Bayesian strategy. The difference is that the robust strategy recommends a smaller stock allocation over time than the Bayesian strategy. The intuition is that an ambiguity-averse investor is more concerned about model uncertainty and hence the hedging component associated with model uncertainty is more negative. When the predictive variable x_0 is equal to zero, the VAR model and the IID model are very similar and hence both the robust strategy and the Bayesian strategy advise similar stock allocations, as shown in Panel B of Figure 7. By contrast, when $x_0 = 0.4504$, the difference in continuation values between the IID and the VAR models is large, causing a large difference in stock allocations between the Bayesian and the robust strategies, as shown in Panel C of Figure 7.

6. Welfare Costs of Suboptimal Investment Strategies

In the previous section, we have shown that the robust strategy may give very different advice to an ambiguity-averse investor than other popular investment strategies such as the IID strategy, the VAR strategy, and the Bayesian strategy. These strategies maximize expected utility and are not optimal to an ambiguity-averse investor. An important question is the following: How costly is it to an ambiguity-averse investor if he does not follow the robust strategy when facing model uncertainty? To study this question, we suppose that an ambiguity-averse investor follows one of the preceding suboptimal investment strategies given that he has an ambiguity utility function defined in (5). We then compute this investor's value function under a suboptimal investment strategy and compare it with the value function under the optimal robust investment strategy.

Let the initial value function implied by the suboptimal strategy k be $J_0^k(W_0, x_0, \mu_0)$, where $k \in \{1, 2, 3\}$ indexes one of the preceding three investment strategies. As is standard in the literature (e.g., Campbell and Viceira (1999) and Xia (2001)), we define the welfare cost as the percentage wealth compensation Δ^k needed to leave the ambiguity-averse investor indifferent between the suboptimal investment strategy and the optimal robust investment strategy, i.e.,

$$J_0^k \left(W_0 \left(1 + \Delta^k \right), x_0, \mu_0 \right) = J_0 \left(W_0, x_0, \mu_0 \right), \tag{26}$$

where J_0 is the initial value function of an ambiguity-averse investor. Note that the welfare cost Δ^k depends on the initial state (x_0, μ_0). In Appendix B, we describe the method to compute the welfare cost. For space limitation, here we only report the welfare cost of following the Bayesian strategy rather than the robust strategy in Figure 8. Xia (2001) has already studied the welfare implications for a Bayesian investor by comparing the IID, VAR, and Bayesian strategies. We do not repeat this analysis here.[15]

---

[15] In results not reported here, we have computed welfare costs of adopting each of these investment strategies for

Figure 8 shows that welfare costs are sizable for small prior probabilities μ_0 of the IID model and large values of the predictive variable x_0. In particular, Panel A of Figure 8 shows that welfare costs can exceed 15 percent when an ambiguity-averse investor with $(\rho, \gamma, \eta) = (0.5, 5, 60)$ adopts a Bayesian strategy with $(\rho, \gamma, \eta) = (0.5, 5, 5)$ by ignoring model ambiguity. Panel B shows that when the ambiguity aversion parameter is increased to $\eta = 100$, welfare costs rise even higher. These high welfare costs happen precisely when the investor believes that it has a high probability that the stock return follows the VAR model and when the predictive variable takes high values. This result is consistent with Figures 2–6. As we discussed earlier, for the state vector (μ_0, x_0) in that region, the robust and the Bayesian strategies have both large qualitative and quantitative differences. The intuition is that in that region the IID model of stock returns delivers a lower continuation value to the investor, which leads him to slant his beliefs towards the IID model because of his pessimistic behavior. Even though the Bayesian posterior favors the VAR model in the data as shown in Figure 1, there is always a small probability that the IID model is on the table in a finite sample of data. Concerning about model misspecification that may lead to a large utility loss, the slanted subjective beliefs about the IID model can be far apart from the Bayesian beliefs. For example, when $\mu_0 = 0.04$, the distorted belief $\hat{\mu}_0$ ranges from about 0.2 to 0.6 when x_0 takes values between one and two standard deviations as shown in Figure 2. These large differences in beliefs cause the large differences in investment strategies.

Figure 8 also shows that the welfare costs are also impressive (about 3 to 7 percent) for large values of μ_0 and small values of x_0, although they are not as large as those in the previous case. The intuition is similar. For this region of states, the VAR model leads to a smaller continuation value because it implies small expected stock returns. The ambiguity-averse investor slants his beliefs toward the VAR model, causing $\hat{\mu}_0$ to be smaller than μ_0 as shown in Figure 2.

For the remaining region of states, welfare costs are very small because both the IID and VAR models of stock returns give similar continuation values to the investor. Model misspecification does not lead to too large utility losses so that the investor is less concerned about it.

7. Conclusion

Whether the stock return is predictable is highly debated. In this paper, we study an investor's optimal consumption and portfolio choice problem when he is confronted with two possibly misspecified models of stock returns: the IID model and the VAR model. The investor does not know which one is the true model and fears that both models may be misspecified. He learns about the stock return model under ambiguity and his learning problem departs from the standard Bayesian

---

an ambiguity-averse investor using different measures of dividend yields and different parameter values. We find that weflare costs of adopting the VAR strategy are extremely large. These results are available upon request.

approach. He copes with the specification doubts by slanting his beliefs pessimistically. We find that an ambiguity-averse investor's robust investment strategy is qualitatively and quantitatively different from some other investment strategies studied in the literature. In particular, the robust strategy is more conservative than the Bayesian strategy. This effect is especially large for sufficiently high or low values of the predictive variable. For low values of the predictive variable, an ambiguity-averse investor is more likely to not participate in the stock market. For high values, the robust strategy recommends much smaller stock holdings than the Bayesian strategy or the VAR strategy.

In addition, when the prior probability of the IID model is small, there is a region of states such that the robust strategy advises the investor to decrease his stock allocation with the predictive variable while the Bayesian strategy advises the opposite. Another interesting finding is for the case in which the predictive variable takes large values. In this case, even a small prior probability of the IID model can lead an ambiguity-averse investor to decrease his stock allocation sharply and then increase it gradually as the prior probability of the IID model rises. This is in contrast to the Bayesian strategy which implies that the stock allocation decreases monotonically with the prior probability of the IID model. The key intuition is that the ambiguity-averse investor slants his belief heavily towards the IID model and this distorted belief is sensitive to small changes in Bayesian probability. As a result, the hedging demand against model ambiguity is negative and large when his prior probability of the IID model is small and when the predictive variable takes a large value.

We also find that the welfare cost of adopting the Bayesian strategy for the ambiguity-averse investor can be sizable. It is more than 15 percent of his initial wealth when the predictive variable takes large values and when the prior probability of the IID submodel of stock returns is small. We emphasize that our results of large welfare costs and large differences between the Bayesian and the robust strategies are meaningful because they happen in the empirically relevant region of states and apply to ambiguity-averse investors with strong priors about the VAR submodel.

In our model, we have assumed that the investor knows the parameters in the submodels of stock returns. It would be interesting to extend our model to incorporate uncertainty about these parameters. Such an extension would significantly complicate our analysis. We leave this extension for future research.

Appendices

A Proofs of Results in Section 3.3

We conjecture the value function takes the form:

$$J_t\left(W_t, x_t, \mu_t\right) = W_t G_t\left(x_t, \mu_t\right), \ \ G_T = 1. \tag{A.1}$$

where G_t is a function to be determined. We substitute this conjecture into the Bellman equation (20) to derive:

$$W_t G_t\left(x_t, \mu_t\right) = \max_{C_t>0, \psi_t \in [0,1]} \left[C_t^{1-\rho} + \beta \left\{\mu_t \left(\mathbb{E}_{1,t}\left[W_{t+1}^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right)\right]\right)^{\frac{1-\eta}{1-\gamma}} \right.\right. \tag{A.2}$$

$$\left.\left. + \left(1 - \mu_t\right)\left(\mathbb{E}_{2,t}\left[W_{t+1}^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right)\right]\right)^{\frac{1-\eta}{1-\gamma}}\right\}^{\frac{1-\rho}{1-\eta}}\right]^{\frac{1}{1-\rho}}.$$

We substitute the budget constraint (19) into the above Bellman equation to obtain:

$$W_t G_t\left(x_t, \mu_t\right) = \max_{C_t, \psi_t}\left[C_t^{1-\rho} + \beta\left(W_t - C_t\right)^{1-\rho} H_t\left(\psi_t, x_t, \mu_t; G_{t+1}\right)\right]^{\frac{1}{1-\rho}}, \tag{A.3}$$

where we define

$$H_t\left(\psi_t, x_t, \mu_t; G_{t+1}\right) = \left\{\mu_t\left(\mathbb{E}_{1,t}\left[\left(R_{t+1}\psi_t + R_f\left(1 - \psi_t\right)\right)^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right)\right]\right)^{\frac{1-\eta}{1-\gamma}} \right. \tag{A.4}$$

$$\left. + \left(1 - \mu_t\right)\left(\mathbb{E}_{2,t}\left[\left(R_{t+1}\psi_t + R_f\left(1 - \psi_t\right)\right)^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right)\right]\right)^{\frac{1-\eta}{1-\gamma}}\right\}^{\frac{1-\rho}{1-\eta}}.$$

We use the first-order condition for C_t to derive

$$\left(\frac{C_t}{W_t - C_t}\right)^{-\rho} = \beta H_t\left(\psi_t, x_t, \mu_t; G_{t+1}\right). \tag{A.5}$$

Solving yields a linear consumption rule:

$$C_t = a_t W_t, \tag{A.6}$$

where we define

$$a_t = \frac{1}{1 + \left(\beta H_t\left(\psi_t, x_t, \mu_t; G_{t+1}\right)\right)^{1/\rho}}. \tag{A.7}$$

We may equivalently write the portfolio choice problem as

$$\max_{\psi_t \in [0,1]} \frac{1}{1-\rho} H_t \left(\psi_t, x_t, \mu_t; G_{t+1} \right). \tag{A.8}$$

In an interior solution, the optimal portfolio weight ψ_t satisfies the following first-order condition:

$$0 = \mu_t \left(\mathbb{E}_{1,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{1-\gamma} G_{t+1}^{1-\gamma} \left(x_{t+1}, \mu_{t+1} \right) \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}} \tag{A.9}$$

$$\times \left\{ \mathbb{E}_{1,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{-\gamma} G_{t+1}^{1-\gamma} \left(x_{t+1}, \mu_{t+1} \right) (R_{t+1} - R_f) \right] \right\}$$

$$+ (1 - \mu_t) \left(\mathbb{E}_{2,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{1-\gamma} G_{t+1}^{1-\gamma} \left(x_{t+1}, \mu_{t+1} \right) \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}$$

$$\times \left\{ \mathbb{E}_{2,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{-\gamma} G_{t+1}^{1-\gamma} \left(x_{t+1}, \mu_{t+1} \right) (R_{t+1} - R_f) \right] \right\}.$$

Substituting the consumption rule (A.6) into the Bellman equation (A.3), we obtain:

$$G_t \left(x_t, \mu_t \right) = \left[a_t^{1-\rho} + \beta (1 - a_t)^{1-\rho} H_t \left(\psi_t, x_t, \mu_t; G_{t+1} \right) \right]^{\frac{1}{1-\rho}}$$

$$= \left[1 + (\beta H_t \left(\psi_t, x_t, \mu_t; G_{t+1} \right))^{1/\rho} \right]^{\frac{\rho}{1-\rho}} = a_t^{\frac{-\rho}{1-\rho}} = \left(\frac{C_t}{W_t} \right)^{\frac{-\rho}{1-\rho}}. \tag{A.10}$$

Substituting this equation into (A.9), we obtain:

$$0 = \mu_t \left(\mathbb{E}_{1,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{1-\gamma} \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\frac{\rho(1-\gamma)}{1-\rho}} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}} \tag{A.11}$$

$$\times \left\{ \mathbb{E}_{1,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{-\gamma} \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\frac{\rho(1-\gamma)}{1-\rho}} (R_{t+1} - R_f) \right] \right\}$$

$$+ (1 - \mu_t) \left(\mathbb{E}_{2,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t)) \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\frac{\rho(1-\gamma)}{1-\rho}} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}$$

$$\times \left\{ \mathbb{E}_{2,t} \left[(R_{t+1}\psi_t + R_f (1-\psi_t))^{-\gamma} \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\frac{\rho(1-\gamma)}{1-\rho}} (R_{t+1} - R_f) \right] \right\}.$$

Using the budget constraint and (A.6), we obtain:

$$W_{t+1} = (W_t - C_t) R_{p,t+1} = R_{p,t+1} C_t \frac{1 - a_t}{a_t}. \tag{A.12}$$

So we have

$$\mathbb{E}_{\mu_t}\left\{\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\mathbb{E}_{z,t}\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}R_{t+1}\right]\right\}\qquad\text{(A.13)}$$

$$=\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_{z,t}\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}R_f\right]\right\}.$$

Multiply (A.13) by ψ_t and use $\psi_t R_{t+1}=R_{p,t+1}-(1-\psi_t)R_f$ to obtain

$$\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{1-\eta}{1-\gamma}}\qquad\text{(A.14)}$$

$$=\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_{z,t}\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{(1-\rho)}}R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}R_f\right]\right\}.$$

Using (A.6), we can rewrite (A.5) as

$$\left(\frac{1-a_t}{a_t}\right)^{\rho}=\beta H_t\left(\psi_t,x_t,\mu_t;G_{t+1}\right).\qquad\text{(A.15)}$$

Using (A.12) , we can rewrite (A.10) as:

$$G_{t+1}(x_{t+1},\mu_{t+1})=\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho}{1-\rho}}R_{p,t+1}^{\frac{\rho}{1-\rho}}\left(\frac{1-a_t}{a_t}\right)^{\frac{\rho}{1-\rho}}.\qquad\text{(A.16)}$$

Substitute equation (A.16) into the definition of H_t on the right hand side of (A.15) and simplify to obtain:

$$\beta^{\frac{1-\eta}{1-\rho}}\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\rho(1-\gamma)/(1-\rho)}(R_{p,t+1})^{\frac{1-\gamma}{1-\rho}}\right]\right)^{\frac{1-\eta}{1-\gamma}}=1.\qquad\text{(A.17)}$$

Then, we can write (A.13) as

$$\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_{z,t}\left[\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}}R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}R_{t+1}\right]\right\}$$

$$=\mathbb{E}_{\mu_t}\left(\mathbb{E}_{z,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_{z,t}\left[\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{(1-\gamma)}{(1-\rho)}}R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}R_f\right]\right\}=1.$$

This equation can be written as the Euler equation

$$\mathbb{E}_t[M_{z,t+1}R_{t+1}] = \mathbb{E}_t[M_{z,t+1}R_f] = 1,$$

where the pricing kernel is given by equation is given by

$$M_{z,t+1} = \left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}-1}\left(\mathbb{E}_{z,t}\left[\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}} R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}. \tag{A.18}$$

We then obtain:

$$\mathbb{E}_t[M_{z,t+1}(R_{t+1} - R_f)] = 0. \tag{A.19}$$

By defining the distorted beliefs about the IID model $\hat{\mu}_t$ as:

$$\hat{\mu}_t = \frac{\mu_t\left(\mathbb{E}_{1,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\mu_t\left(\mathbb{E}_{1,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}} + (1-\mu_t)\left(\mathbb{E}_{2,t}\left[R_{p,t+1}^{\frac{1-\gamma}{1-\rho}}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right)^{\frac{(1-\gamma)}{(1-\rho)}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}, \tag{A.20}$$

we can rewrite (A.19) as equation (23).

By (A.1) and (A.10), we derive

$$J_t = W_t G_t = W_t\left(\frac{C_t}{W_t}\right)^{\frac{-\rho}{1-\rho}},$$

and

$$\frac{W_t}{C_t} = \left(\frac{J_t}{C_t}\right)^{1-\rho},$$

where, without risk of confusion, we have suppressed arguments for J_t and G_t. Using (A.12) we have

$$R_{p,t+1} = \frac{W_{t+1}}{C_t}\left(\frac{a_t}{1-a_t}\right) = \frac{C_{t+1}}{C_t}\left(\frac{J_{t+1}}{C_{t+1}}\right)^{1-\rho}\left(\frac{a_t}{1-a_t}\right). \tag{A.21}$$

This equation implies that

$$J_{t+1} = R_{p,t+1}^{\frac{1}{1-\rho}}C_{t+1}^{\frac{-\rho}{1-\rho}}\left(\frac{1-a_t}{a_t}\right)^{\frac{1}{1-\rho}}C_t^{\frac{1}{1-\rho}}.$$

Substituting this equation into (A.20) and cancelling terms adapted to time t information from the denominator and the numerator, we obtain (24).   Q.E.D.

B   Computation Method

We use the standard discrete state space value function iteration method, similar to that in Barberis (2000), to solve the model by backward induction. We choose the state space for the state variables (x, μ) as $\left[-4\sigma^x/\sqrt{1-\rho^2}, 4\sigma^x/\sqrt{1-\rho^2}\right] \times [0, 1]$. We discretize this space using 1001×201 equally spaced points. Increasing grid points does not change our results much. We compute the expectation in the Bellman equation using the Gaussian quadrature method. In the last period T, $C_T = W_T$, there is no portfolio choice, and $G_T = 1$. In period $T - 1$, the optimal portfolio weight $\psi_{T-1}^* \in [0, 1]$ solves the problem:

$$\max_{\psi_{T-1} \in [0,1]} \quad \frac{1}{1-\gamma} H_{T-1}\left(\psi_{T-1}, x_{T-1}, \mu_{T-1}; G_T\right). \tag{B.1}$$

We next solve for the optimal consumption-wealth ratio a_{T-1}^* using equation (A.7). Substituting $a_{T-1}^*\left(x_{T-1}, \mu_{T-1}\right)$ into (A.10) for $t = T-1$, we obtain G_{T-1}. In general, suppose at time t, we know G_{t+1}. We then use equation (A.8) to solve for the optimal portfolio weight $\psi_t^*\left(x_t, \mu_t\right)$. Substitute $\psi_t^*\left(x_t, \mu_t\right)$ into equation (A.7) to obtain a_t^*. Substituting a_t^* into (A.10), we obtain G_t. We then go to time $t - 1$ and repeat the above procedure again, until we reach $t = 0$.

To solve for the welfare costs of suboptimal investment strategies, we use the following procedure. Let $k \in \{1, 2, 3\}$ index one of the three investment strategies defined in Section 5. These strategies together with the implied consumption rules satisfy the budget constraint (19) and hence are feasible for an ambiguity-averse investor's consumption/portfolio choice problem with the recursive ambiguity utility (5). Thus they deliver lower life-time utility values than the robust strategy when the utility function (5) is used. As in Appendix A, we can show that the three suboptimal investment strategies also give linear consumption rules $C_t^k = a_t^k W_t$. Substituting these rules into (5), we obtain $J_0^k\left(W_0, x_0, \mu_0\right) = W_0 G_0^k\left(x_0, \mu_0\right)$, where G_0^k is some function. The welfare cost for suboptimal investment strategy k is given by

$$\Delta^k\left(x_0, \mu_0\right) = \frac{G_0\left(x_0, \mu_0\right)}{G_0^k\left(x_0, \mu_0\right)} - 1.$$

References

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, 2003, A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection, *Journal of the European Economic Association* 1, 68-123.

Bansal, Ravi and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481-1509.

Barberis, Nicholas, 2000, Investing for the Long Run when Returns are Predictable, *Journal of Finance* 55, 225-264.

Boyle, Phelim P., Lorenzo Garlappi, Raman Uppal, and Tan Wang, 2012, Keynes Meets Markowitz: The Trade-off Between Familiarity and Diversification, *Management Science* 58, 253-272.

Brandt, Michael W., Amit Goyal, and Pedro Santa-Clara, and Jonathan R. Stroud, 2005, A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability, *Review of Financial Studies* 18, 831-873.

Brennan, Michael, 1998, The Role of Learning in Dynamic Portfolio Decisions, *European Finance Review* 1: 295-306.

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, Strategic Asset Allocation, *Journal of Economic Dynamics and Control* 21, 1377-1403.

Camerer, Colin F., 1999, Ambiguity-Aversion and Non-Additive Probability: Experimental Evidence, Models and Applications, *Uncertain Decisions: Bridging Theory and Experiments*, ed. by L. Luini, pp. 53-80. Kluwer Academic Publishers.

Campanale, Claudio, 2011, Learning, ambiguity and life-cycle portfolio allocation, *Review of Economic Dynamics* 14, 339-367.

Campbell, John Y. and Samuel B. Thompson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies* 21, 1509-1531.

Campbell, John Y. and Luis M. Viceira, 1999, Consumption and Portfolio Decisions When Expected Returns are Time Varying, *Quarterly Journal of Economics* 114, 433-495.

Campbell, John Y. and Luis M. Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press, New York.

Camponovo, Lorenzo, O. Scaillet, and Fabio Trojani, 2012, Predictive Regression and Robust Hypothesis Testing: Predictability Hidden by Anomalous Observations, working paper, Swiss Finance Institute.

Cao, H. Henry, Bing Han, David Hirshleifer, and Harold H. Zhang, 2011, Fear of the Unknown: Familiarity and Economic Decisions, *Review of Finance* 15, 173-206.

Cao, H. Henry, Tan Wang, and Harold H. Zhang, 2005, Model Uncertainty, Limited Market Participation, and Asset Prices, *Review of Financial Studies* 18, 1219 - 1251.

Chen, Zengjin, and Larry G. Epstein, 2002, Ambiguity, Risk and Asset Returns in Continuous-Time, *Econometrica* 4, 1403-1445.

Chew, Song Hong and Jacob Sagi, 2008, Small Worlds: Modelling Attitudes towards Sources of Uncertainty, *Journal of Economic Theory* 139, 1-24.

Cochrane, John H., 2008, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1534-1575.

Detemple, Jerome B., 1986, Asset Pricing in an Economy with Incomplete Information, *Journal of Finance* 61, 383-392.

Dothan, Michael U., and David Feldman, 1986, Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy, *Journal of Finance* 41, 369–382.

Ellsberg, Daniel, 1961, Risk, Ambiguity and the Savage Axiom, *Quarterly Journal of Economics* 75, 643-669.

Epstein L. G., 2010, A Paradox for the "Smooth Ambiguity" Model of Preference, *Econometrica* 78, 2085-2099.

Epstein, Larry G. and Jianjun Miao, 2003, A Two-Person Dynamic Equilibrium under Ambiguity, *Journal of Economic Dynamics and Control* 27, 1253-1288.

Epstein, Larry G. and Martin Schneider, 2003, Recursive multiple-priors, *Journal of Economic Theory 113, 1-31.*

Epstein, Larry G. and Martin Schneider, 2007, Learning under ambiguity, *Review of Economic Studies* 74, 1275-1303.

Epstein, Larry G. and Tan Wang, 1994, Intertemporal Asset Pricing under Knightian Uncertainty, *Econometrica* 62, 283-322.

Epstein, Larry G. and Stanley Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937-969.

Ergin, Huluk I. and Faruk Gul, 2009, A Theory of Subjective Compound Lotteries, *Journal of Economic Theory* 144, 899-929.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang, 2007, Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach, *Review of Financial Studies* 20, 41-81.

Gennotte, Gerard, 1986, Optimal Portfolio Choice under Incomplete Information, *Journal of Finance* 41, 733-749.

Gilboa, Itzach and David Schmeidler, 1989, Maxmin Expected Utility with Non-unique Priors, *Journal of Mathematical Economics* 18, 141-153.

Gollier, Christian, 2004, Optimal Portfolio Risk with First-Order and Second -Order Predictability, *Contributions to Theoretical Economics* Vol 4, No 1.

Gollier, Christian, 2011, Portfolio Choices and Asset Prices: The Comparative Statics of Ambiguity Aversion, *Review of Economic Studies* 78, 1329-1344.

Halevy, Yoran, 2007, Ellsberg Revisited: An Experimental Study, *Econometrica* 75, 503-536.

Hansen, Lars Peter, 2007, Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk, *American Economic Review, Papers and Proceedings* 97, 1-30.

Hansen, Lars Peter and Thomas J. Sargent, 2001, Robust Control and Model Uncertainty, *American Economic Review* 91, 60-66.

Hansen, Lars Peter and Thomas J. Sargent, 2007a, Recursive Robust Estimation and Control without Commitment, *Journal of Economic Theory* 136, 1-27.

Hansen, Lars Peter, and Thomas J. Sargent, 2007b, *Robustness*, Princeton University Press.

Hansen, Lars Peter and Thomas J. Sargent, 2010, Fragile Beliefs and the Price of Uncertainty, *Quantitative Economics* 1, 129-162.

Hansen, Lars Peter and Kenneth J. Singleton, 1982, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, *Econometrica* 50, 1269-1286.

Hayashi, Takashi and Jianjun Miao, 2011, Intertemporal Substitution and Recursive Smooth Ambiguity Preferences, *Theoretical Economics* 6, 423-472.

Ju, Nengjiu and Jianjun Miao, 2012, Ambiguity, Learning, and Asset Returns, *Econometrica* 80, 559-591.

Kandel, Shmuel and Robert F. Stambaugh, 1996, On the Predictability of Stock Returns: An Asset-Allocation Perspective, *Journal of Finance* 51, 385-424.

Kim, Tongsuk and Edward Omberg, 1996, Dynamic Nonmyopic Portfolio Behavior, *Review of Financial Studies* 9, 141-161.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A Smooth Model of Decision Making under Ambiguity, *Econometrica* 73, 1849-1892.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2009, Recursive Smooth Ambiguity Preferences, *Journal of Economic Theory* 144, 930-976.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2012, On the Smooth Ambiguity Model: A Reply, *Econometrica* 80, 1303-1321.

Kreps, David, M. and Evan L. Porteus, 1978, Temporal Resolution of Uncertainty and Dynamic Choice, *Econometrica* 46, 185-200.

Leippold, Markus, Fabio Trojani, and Paolo Vanini, 2008, Learning and Asset Prices under Ambiguous Information, *Review of Financial Studies* 21, 2565-2597.

Lettau, Martin and Stijn Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, *Review of Financial Studies* 21, 1607-1652.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An Equilibrium Model of Rare Event Premium and Its Implication for Option Smirk, *Review of Financial Studies* 18, 131-164.

Maenhout, Pascal J., 2004, Robust Portfolio Rules and Asset Pricing, *Review of Financial Studies* 17, 951-983.

Merton, Robert, 1969, Lifetime Portfolio Selection under Uncertainty: The Continuous-Time case, *Review of Economics and Statistics* 51, 247-257.

Merton, Robert, 1971, Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory* 3, 373-413.

Miao, Jianjun, 2009, Ambiguity, Risk and Portfolio Choice under Incomplete Information, *Annals of Economics and Finance* 10, 257-279.

Nau, Robert F., 2006, Uncertainty Aversion with Second-Order Utilities and Probabilities, *Management Science* 52, 136-145.

Pastor, Lubos and Robert F. Stambaugh, 2012, Are stocks really less volatile in the long run? *Journal of Finance* 67, 431-478.

Rothschild, Michael and Joseph E. Stiglitz, 1971, Increasing Risk: II. Its Economic Consequences, *Journal of Economic Theory* 5, 66-84.

Samuelson, Paul, 1969, Lifetime Portfolio Selection by Dynamic Stochastic Programming, *Review of Economics and Statistics* 51, 239-246.

Segal, Uzi, 1987, The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach, *International Economic Review* 28, 175-202.

Seo, Kyounwon, 2009, Ambiguity and Second-Order Belief, *Econometrica* 77, 1575-1605.

Strzalecki, Tomasz, 2010, Axiomatic Foundations of Multiplier Preferences, *Econometrica* 79, 47-73.

Trojani, Fabio, Christian Wiehenkamp, and Jan Wrampelmeyer, 2013, Taking Ambiguity to Reality: Robust Agents Cannot Trust the Data Too Much, Swiss Finance Institute Research Paper No. 11-33.

Uppal, Raman and Tan Wang, 2003, Model Misspecification and Underdiversification, *Journal of Finance* 58, 2465-2486.

Veronesi, Pietro, 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, *Review of Financial Studies* 12, 975-1007.

Veronesi, Pietro, 2000, How does Information Quality Affect Stock Returns? *Journal of Finance* 55, 807-837.

Wachter, Jessica, 2010, Asset Allocation, *Annual Reviews of Financial Economics* 2, 175-206.

Wachter, Jessica and Missaka Warusawitharana, 2009, Predictable Returns and Asset Allocation: Should a Skeptical Investor Time the Market? *Journal of Econometrics* 148, 162-178

Welch, Ivo and Amit Goyal, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455-1508.

Xia, Yihong, 2001, Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation. *Journal of Finance* 56, 205-246.

Table 1: AMBIGUITY PREMIUM AS A PERCENTAGE OF THE EXPECTED VALUE OF THE BET

γ/η	40.0	50.0	60.0	70.0	80.0	90.0	100.0	110.0
			A. Prize-wealth ratio $= 1.0\%$					
0.5	9.8	12.2	14.6	17.0	19.3	21.6	23.8	26.0
2.0	9.4	11.8	14.2	16.6	18.9	21.2	23.4	25.6
5.0	8.6	11.1	13.5	15.8	18.2	20.4	22.7	24.9
10.0	7.4	9.8	12.2	14.6	16.9	19.2	21.4	23.6
15.0	6.2	8.6	11.0	13.4	15.7	18.0	20.2	22.4
			B. Prize-wealth ratio $= 0.5\%$					
0.5	4.9	6.2	7.4	8.6	9.8	11.1	12.3	13.5
2.0	4.7	6.0	7.2	8.4	9.7	10.9	12.1	13.3
5.0	4.4	5.6	6.8	8.1	9.3	10.5	11.7	12.9
10.0	3.7	5.0	6.2	7.4	8.7	9.9	11.1	12.3
15.0	3.1	4.3	5.6	6.8	8.0	9.3	10.5	11.7

Notes: This table reports ambiguity premium as a percentage of the expected value of the bet for various different values of γ and η. The expression for ambiguity premium is given by equation (8).

Table 2: Estimations of the IID and VAR Models

	m	b	ρ	$(\sigma^r)^2 \times 1000$	$\sigma^{rx} \times 1000$	$(\sigma^x)^2 \times 1000$
IID	0.057	0	0.916	43.752	-22.763	32.656
	(0.023)	(-)	(0.059)	(6.751)	(4.814)	(5.039)
VAR	0.057	0.122	0.916	41.000	-22.772	32.656
	(0.023)	(0.048)	(0.053)	(6.326)	(4.702)	(5.039)

Notes: This table reports the results from estimating the IID and the VAR models of stock returns. The numbers in brackets are standard errors. The predictive variable is ldy, the log dividend yield. The sample period is 1926-2010. All variables are annualized when applicable.

Table 3: OPTIMAL PORTFOLIO WEIGHTS IN PERCENTAGE

(ρ,γ,η)	$\mu_0 = 0.1$			$\mu_0 = 0.5$			$\mu_0 = 0.9$		
	x_0^1	x_0^2	x_0^3	x_0^1	x_0^2	x_0^3	x_0^1	x_0^2	x_0^3
0.5, 2, 2	52.9	100.0	100.0	70.5	100.0	100.0	86.0	92.5	94.9
0.5, 2, 60	45.1	100.0	94.9	47.7	100.0	91.5	61.9	93.8	91.5
0.5, 2, 80	43.6	100.0	90.5	44.4	100.0	91.0	54.9	94.2	91.7
0.5, 2, 100	42.4	100.0	89.0	42.1	100.0	91.0	49.6	94.6	91.9
0.5, 5, 5	27.6	52.5	62.7	30.8	42.1	45.0	34.7	37.0	37.4
0.5, 5, 60	24.0	47.8	46.0	25.1	41.3	38.2	31.6	37.3	36.4
0.5, 5, 80	23.1	47.0	42.6	23.8	41.2	37.2	30.2	37.4	36.2
0.5, 5, 100	22.4	46.4	40.1	22.7	41.1	36.5	28.8	37.5	36.2
0.5, 10, 10	14.3	26.4	30.6	15.6	21.0	22.1	17.3	18.4	18.5
0.5, 10, 60	13.3	25.0	26.3	14.1	20.7	20.1	16.7	18.4	18.2
0.5, 10, 80	13.0	24.6	24.9	13.6	20.7	19.6	16.4	18.5	18.1
0.5, 10, 100	12.7	24.3	23.7	13.2	20.6	19.2	16.1	18.5	18.1
0.5, 5.764, 5.764	24.2	45.6	54.1	26.8	36.5	38.8	30.1	32.0	32.4

(ρ,γ)	IID Model	VAR Model		
		x_0^1	x_0^2	x_0^3
0.5, 2	90.1	47.4	100.0	100.0
0.5, 5	35.8	28.1	66.5	100.0
0.5, 10	17.8	15.1	35.1	56.0

Notes: This table presents the optimal portfolio weights in percentage allocated to the stock. Column 1 denotes various values of ρ, γ and η. Columns 2-10 report the optimal portfolio weights in percentage for different combinations of μ_0 and x_0 where $x_0^1 = -\sigma^x/\sqrt{1-\rho^2} = -0.4504$, $x_0^2 = 0.0$, $x_0^3 = \sigma^x/\sqrt{1-\rho^2} = 0.4504$. The predictive variable x is the price-dividend ratio.

Table 4: PERCENTAGE HEDGING DEMAND OVER TOTAL STOCK DEMAND

	$\mu_0 = 0.1$			$\mu_0 = 0.5$			$\mu_0 = 0.9$		
(ρ, γ, η)	x_0^1	x_0^2	x_0^3	x_0^1	x_0^2	x_0^3	x_0^1	x_0^2	x_0^3
0.5, 2, 2	35.7	4.6	0.0	15.5	7.1	0.0	2.3	2.0	-1.6
0.5, 2, 60	33.5	4.6	-5.4	13.6	7.2	-1.6	-1.6	3.4	1.1
0.5, 2, 80	32.7	4.7	-7.2	14.4	7.2	-0.2	-2.6	3.9	1.6
0.5, 2, 100	32.1	4.7	-4.7	15.4	7.3	0.6	-2.9	4.3	1.9
0.5, 5, 5	51.1	27.6	0.7	23.4	12.0	-9.9	3.8	2.5	-2.7
0.5, 5, 60	47.0	20.5	-15.4	20.6	10.4	-8.4	3.2	3.2	-0.9
0.5, 5, 80	46.0	19.2	-16.7	20.4	10.2	-7.1	3.0	3.5	-0.5
0.5, 5, 100	45.2	18.1	-16.7	20.4	10.1	-5.8	2.7	3.8	-0.1
0.5, 10, 10	53.0	28.3	-0.9	24.7	12.3	-11.0	4.2	2.5	-2.9
0.5, 10, 60	51.0	24.3	-10.8	23.0	11.3	-11.0	3.9	2.8	-2.0
0.5, 10, 80	50.3	23.2	-13.3	22.6	11.0	-10.4	3.8	3.0	-1.7
0.5, 10, 100	49.7	22.2	-15.2	22.4	10.9	-9.8	3.8	3.1	-1.4
0.5, 5.764, 5.764	51.6	27.8	0.3	23.7	12.1	-10.2	3.9	2.5	-2.7

Notes: This table presents the ratio of the hedging demand to the total stock demand in percentage. Column 1 denotes various values of ρ, γ and η. Columns 2-10 report the percentage hedging demand over the total stock demand for different combinations of μ_0 and x_0 where $x_0^1 = -\sigma^x/\sqrt{1-\rho^2} = -0.4504$, $x_0^2 = 0.0$, $x_0^3 = \sigma^x/\sqrt{1-\rho^2} = 0.4504$. The predictive variable x is the price-dividend ratio.
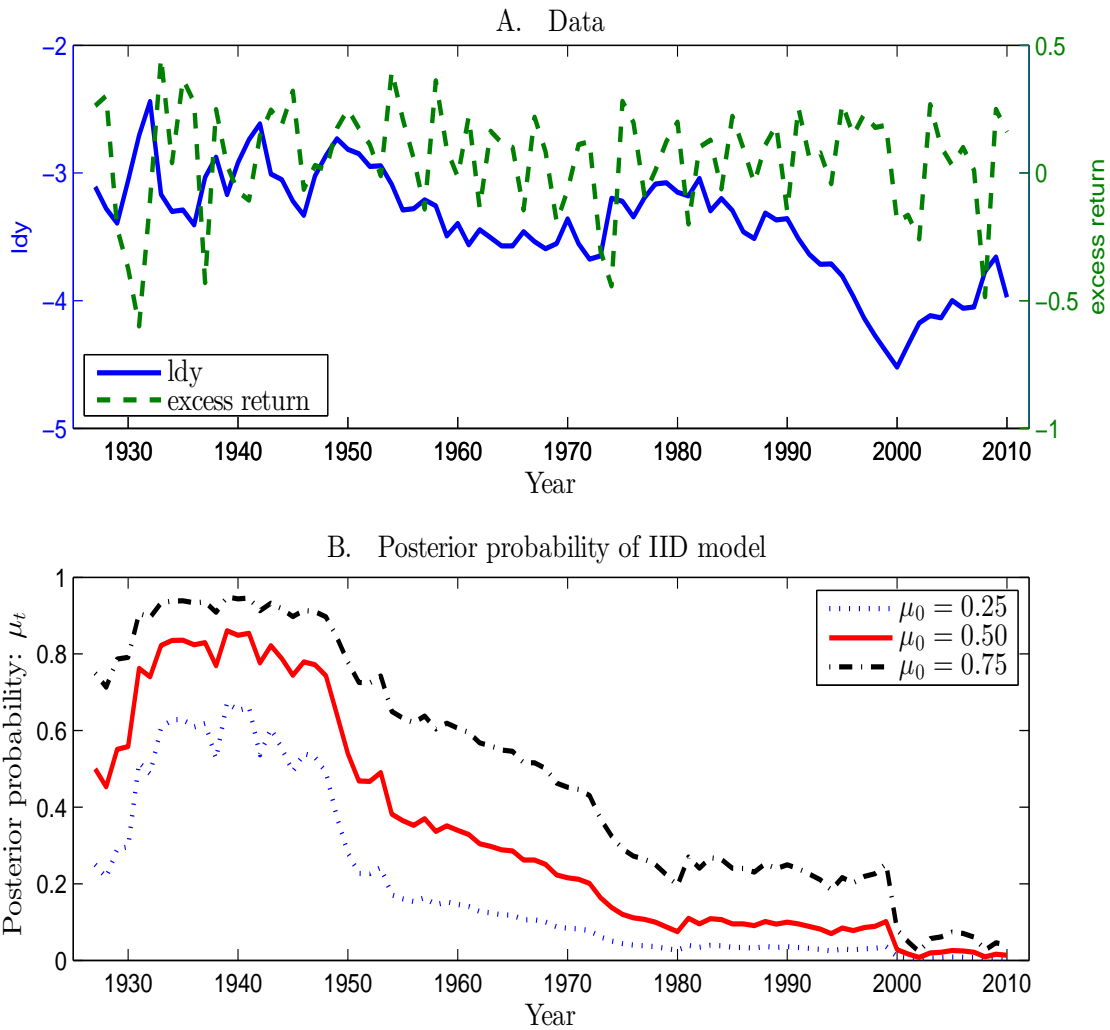
Figure 1: **Returns, predictors, and posterior probabilities of the IID model.** Panel A plots the historical data for log dividend yields and market excess returns. Panel B plots the posterior probabilities of the IID model using the historical annual data of stock returns and the log dividend yields as predictors from 1927-2010. The prior is set at $\mu_0 = 0.25, 0.5, 0.75$ in 1927.
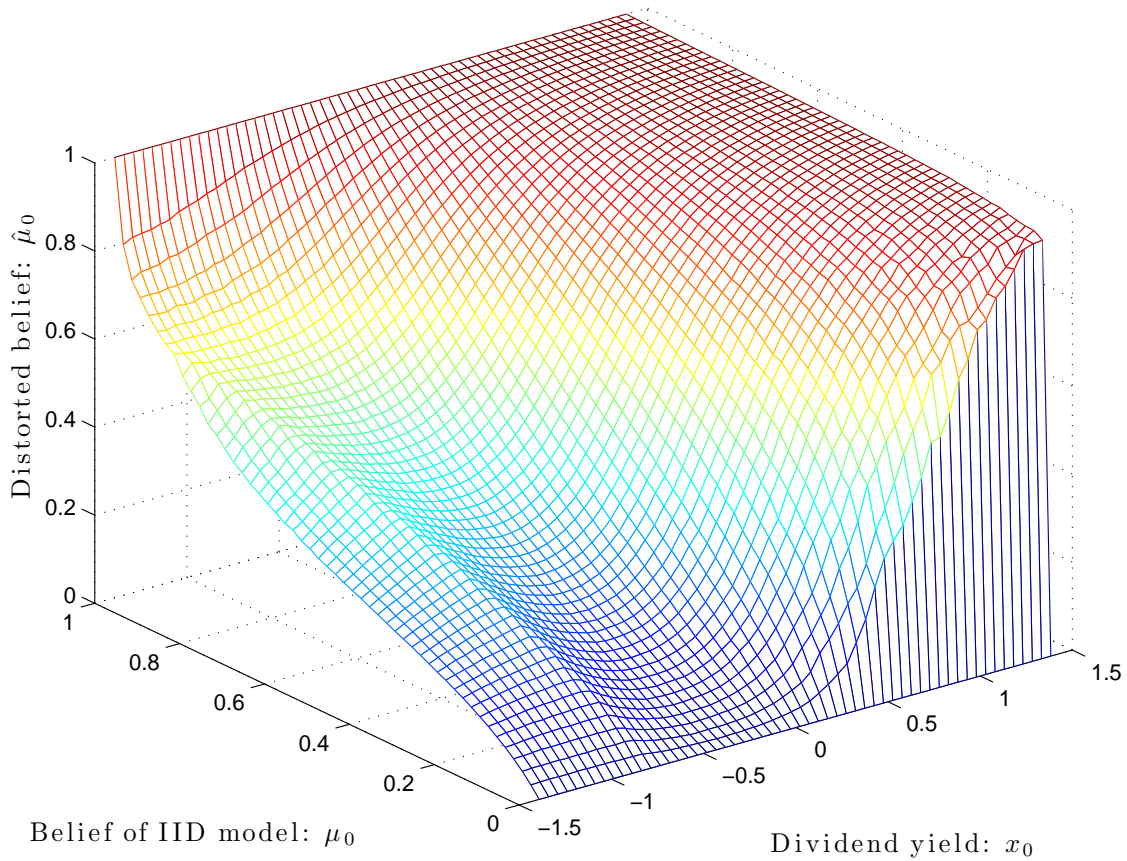
41

Figure 2: **Distorted beliefs of the IID model.** This figure plots the distorted belief of the IID model as a function of the Bayesian belief of the IID model μ_0 and the demeaned dividend yield x_0 for an ambiguity averse investor with horizon $T = 40$ years. We set $\rho = 0.5, \gamma = 5$, and $\eta = 80$. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1 - \rho^2} = 0.4504$.
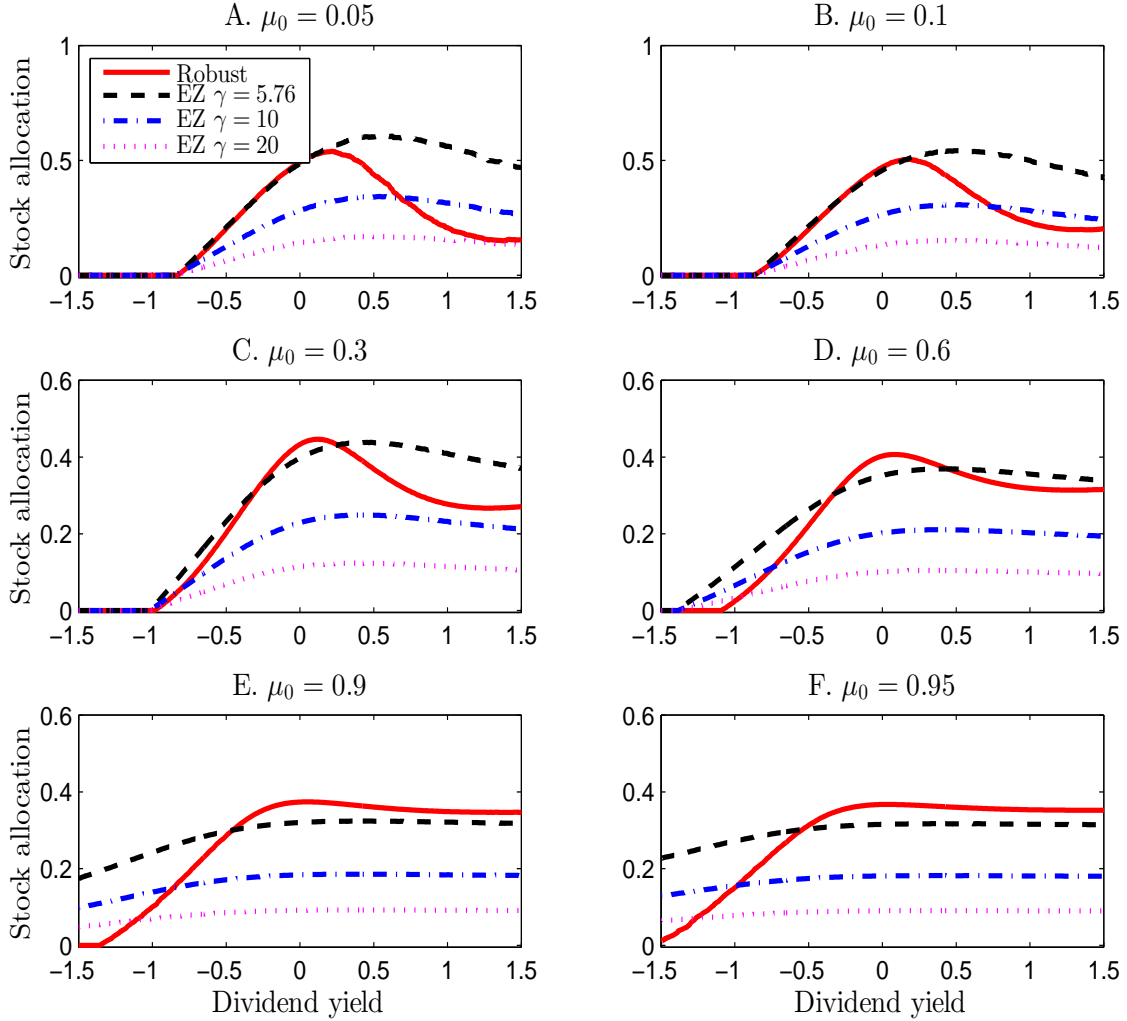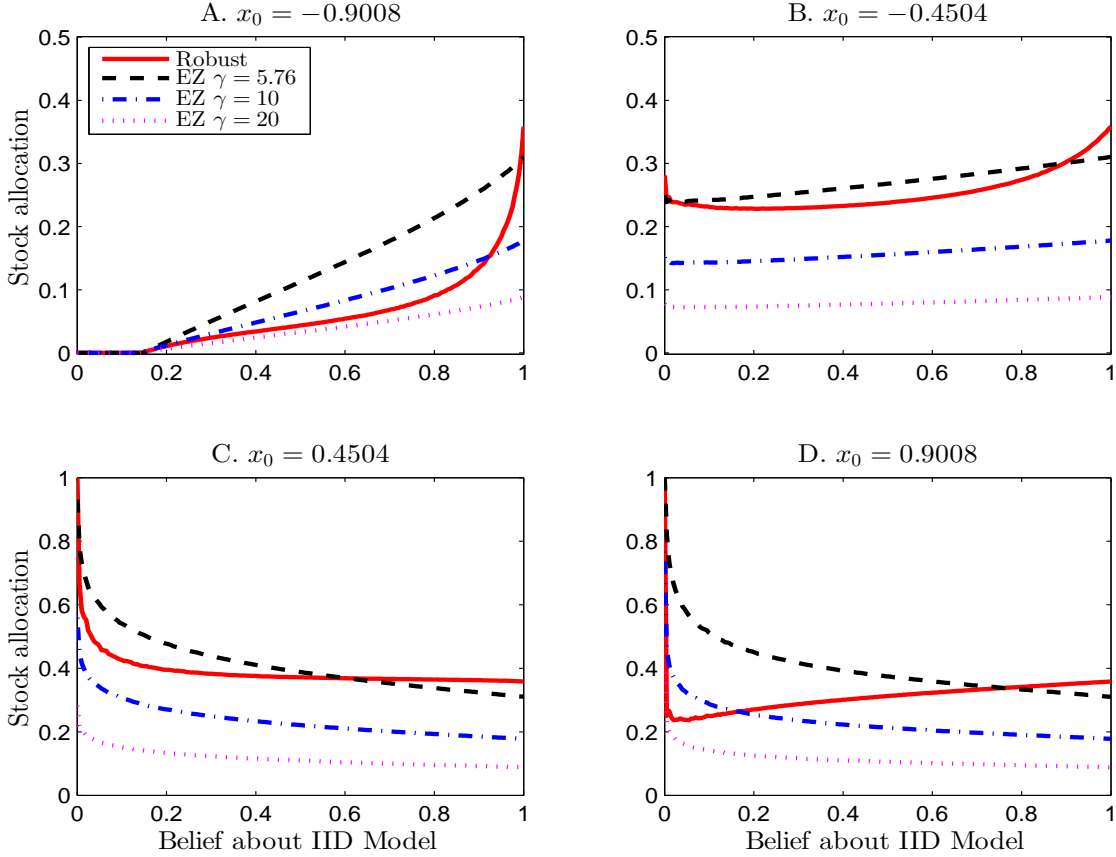
Figure 3: **Comparison with the Epstein-Zin model.** Each panel of this figure plots the stock allocations ψ_0 as functions of the demeaned log dividend yield x_0 for the ambiguity model with $\rho = 0.5, \gamma = 5$, and $\eta = 80$ and for the Epstein-Zin model with $\rho = 0.5$ and $\gamma = 5.76, 10, 20$. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1 - \rho^2} = 0.4504$.
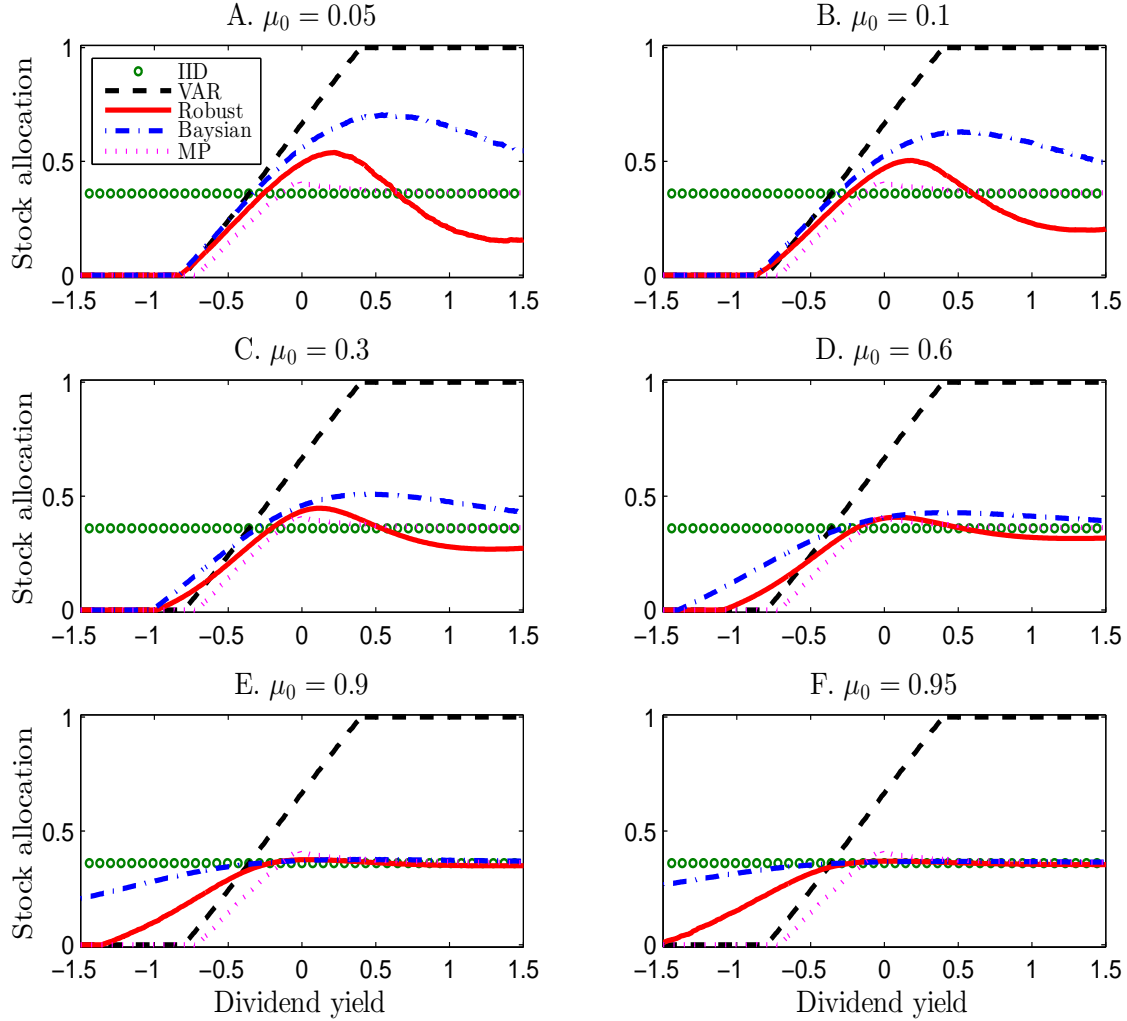
Figure 4: **Comparison with the Epstein-Zin model.** Each panel of this figure plots the stock allocations ψ_0 as functions of the prior μ_0 for the ambiguity model with $\rho = 0.5, \gamma = 5$, and $\eta = 80$ and for the Epstein-Zin model with $\rho = 0.5$ and $\gamma = 5.76, 10, 20$. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1-\rho^2} = 0.4504$.

Figure 5: **Alternative investment strategies: market-timing effect.** This figure plots the initial portfolio weights ψ_0 on the stock for five alternative investment strategies as functions of the initial observation of the demeaned log dividend yield x_0 for different initial beliefs about the IID model μ_0. For the robust strategy ('Robust'), we set $\rho = 0.5, \gamma = 5, \eta = 80$. For the Bayesian strategy ('Bayesian'), we set $\rho = 0.5, \gamma = \eta = 5$. For the multiple-priors model ('MP'), we set $\rho = 0.5, \gamma = 5$, and $\eta = \infty$. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1 - \rho^2} = 0.4504$.
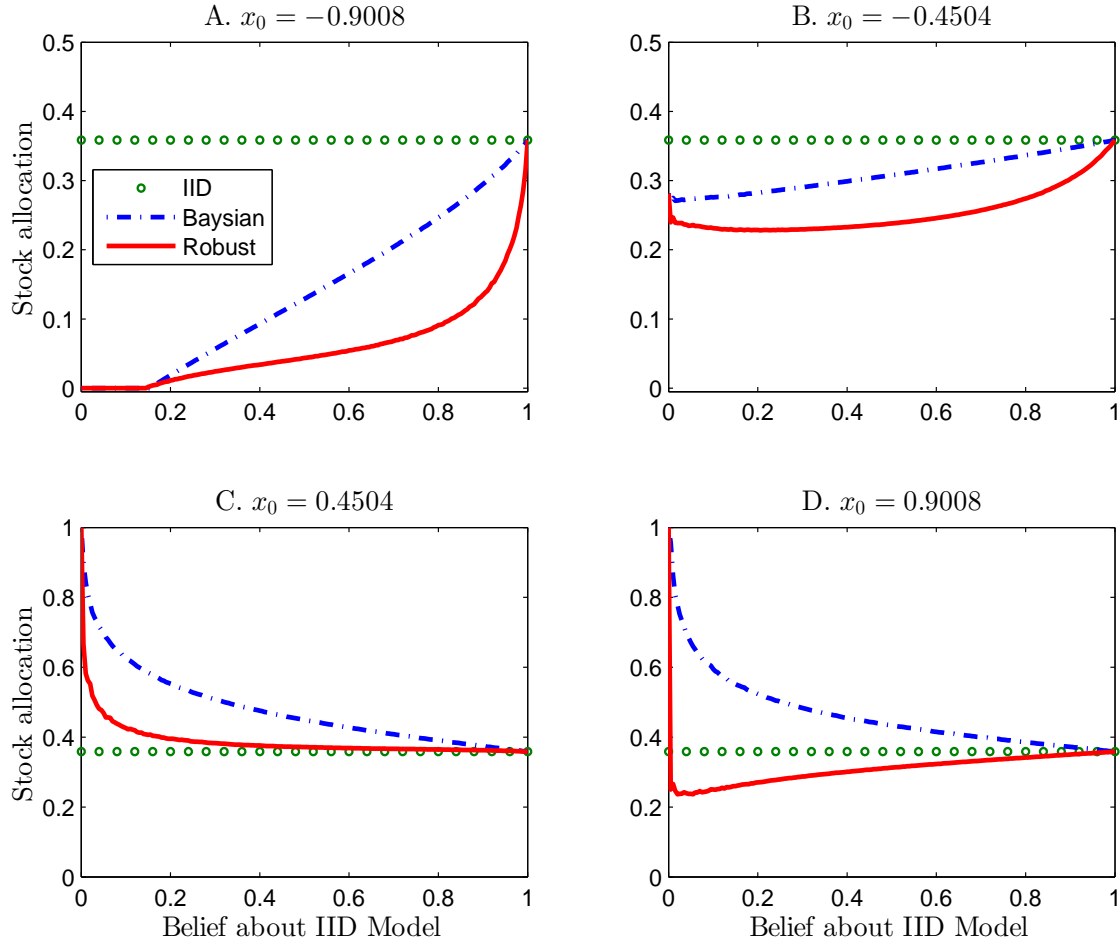
Figure 6: **Alternative investment strategies: uncertainty effect.** This figure plots the portfolio weights on the stock for three alternative investment strategies as functions of the initial beliefs about the IID model μ_0 for four different initial values of the demeaned log dividend yield x_0. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1 - \rho^2} = 0.4504$. For the robust strategy ('Robust'), we set $\rho = 0.5, \gamma = 5, \eta = 80$. For the Bayesian strategy ('Baysian'), we set $\rho = 0.5, \gamma = \eta = 5$.
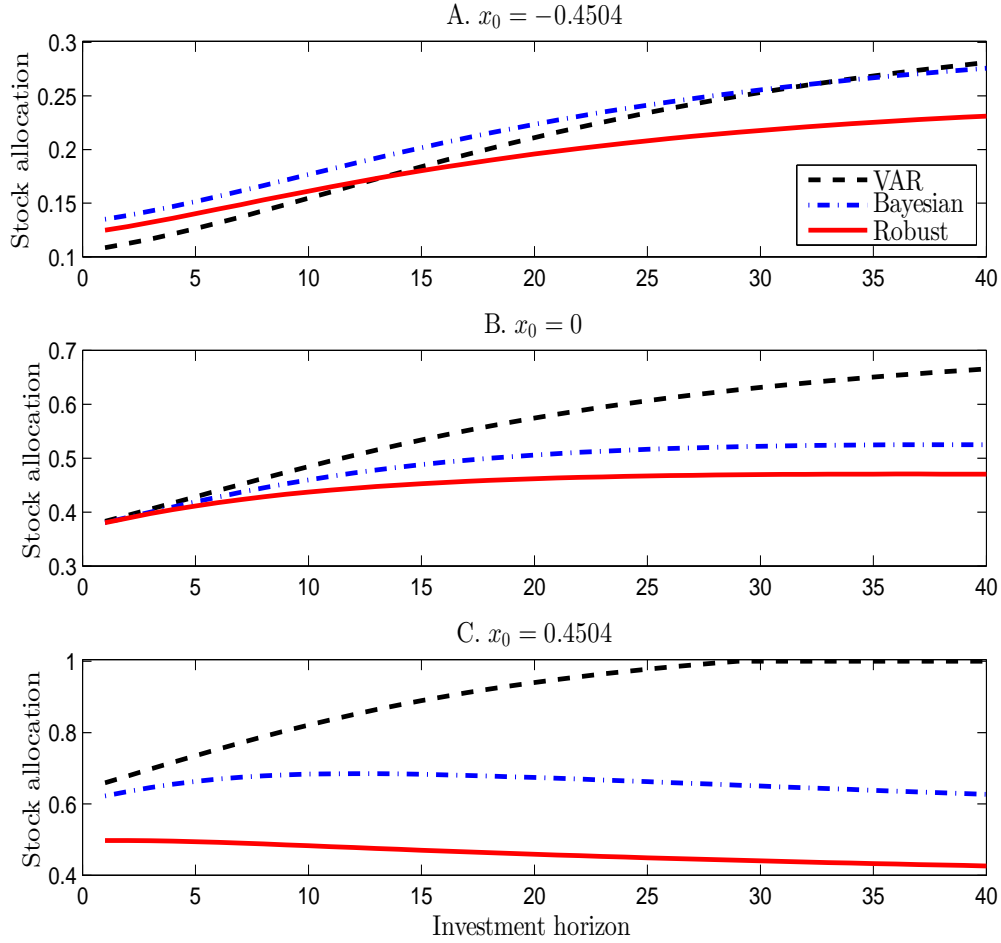
Figure 7: **Alternative investment strategies: horizon effect.** This figure plots the portfolio weights on the stock for three alternative investment strategies as functions of the investment horizon. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1-\rho^2} = 0.4504$. For the robust strategy ('Robust'), we set $\rho = 0.5, \gamma = 5, \eta = 80$. For the Bayesian strategy ('Bayesian'), we set $\rho = 0.5, \gamma = \eta = 5$.
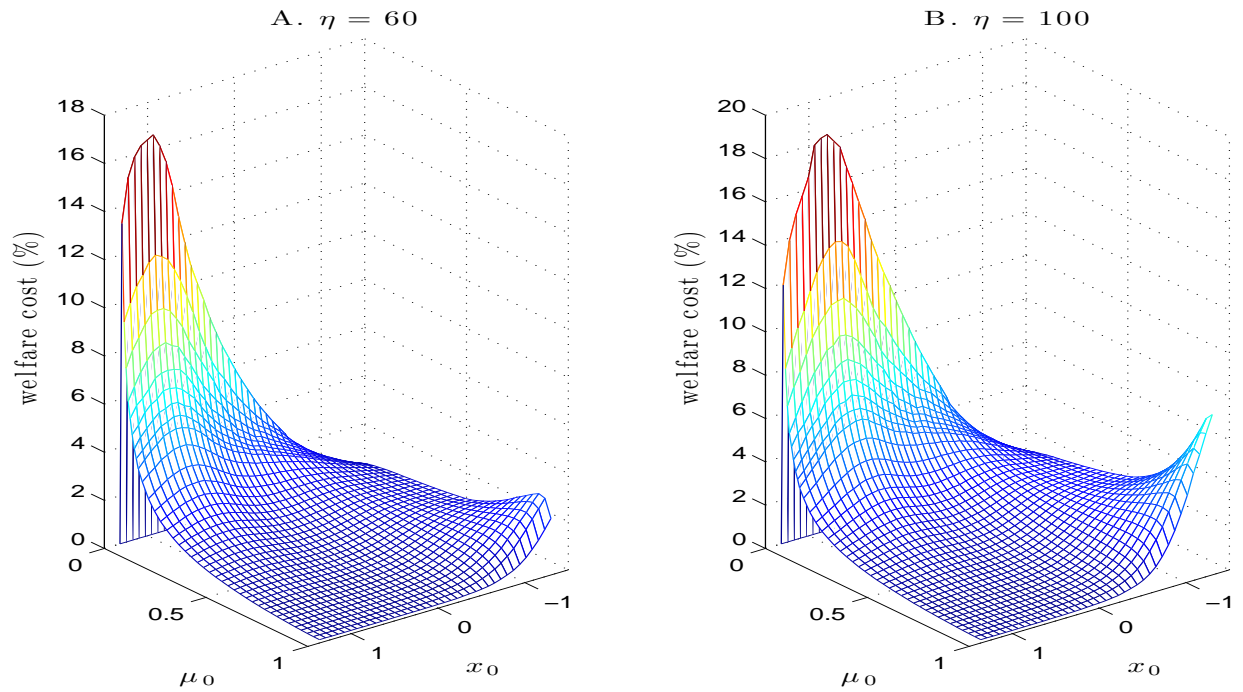
Figure 8: **Welfare costs.** This figure plots the conditional welfare costs of an ambiguity averse investor with horizon $T = 40$ years if he follows the Bayesian strategy by ignoring model ambiguity. The conditional welfare cost is a function of the belief of the IID model μ_0 and the demeaned dividend yield x_0. For the robust strategy, $\rho = 0.5$, $\gamma = 5$, and $\eta = 60, 100$. For the Bayesian strategy, we set $\rho = 0.5, \gamma = \eta = 5$. The unconditional standard deviation of the dividend yield is $\sigma^x/\sqrt{1 - \rho^2} = 0.4504$.