Dynamic Asset Allocation with Ambiguous Return Predictability*

Hui Chen† Nengjiu Ju‡ Jianjun Miao§

First draft: September 2008. This draft: February 2009

Abstract

We study an investor's optimal consumption and portfolio choice problem when he confronts 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 large for extreme values of the predictive variable. Unlike in the Bayesian framework, model uncertainty induces a hedging demand, which may cause the investor to decrease his stock allocations sharply and then increase with his prior probability of IID returns. Adopting suboptimal investment strategies by ignoring model uncertainty can lead to sizable welfare costs.

Keywords: generalized recursive ambiguity utility, ambiguity aversion, model uncertainty, learning, portfolio choice, robustness, return predictability

JEL Classification: D81, D83, G11, E21

*We thank Larry Epstein for helpful conversations, and Monika Piazzesi and Luis Viceira for useful comments. The latest version can be downloaded from Miao's homepage: http://people.bu.edu/miaoj/research.htm.

†MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142. 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 et al. (1997), Campbell and Viceira (1999) and Kim and Omberg (1996)). Economists disagree with the predictability of asset returns. 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.

An important issue of the return predictability is that the predictive relation is quite weak and unstable, and hence the estimation models may be misspecified. The 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. For example, Boudoukh et al. (2007) show that if one replaces the widely used dividend yield by other payout measures such as the total payout yield or the net payout yield, the evidence of predictability is much stronger.

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. 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 riskfree 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. Following most papers in the literature, we adopt the (demeaned) dividend yield as the single predictive variable in the VAR estimation. In addition, we also consider alternative predictive variables such as the (demeaned) total payout yield or the (demeaned) net payout yield suggested by Boudoukh et al. (2007). 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 stock return models. The posteriors and likelihoods are obtained by Bayesian updating. They can be reduced to a single predictive distribution by Bayes' Rule. One can then solve the investor's decision problem using this predictive distribution in the standard expected utility framework (see

[1] See the recent July issue of *Review of Economic Studies* in 2008.

Barberis (2000) and Xia (2001)). We depart from this Bayesian approach in that the posteriors and likelihoods cannot be reduced to a predictive distribution in the investor's decision problem. This irreducibility of compound distributions captures model uncertainty or ambiguity, as discussed by Hansen (2007).

To accommodate model ambiguity and ambiguity aversion, we adopt a recursive ambiguity utility model recently proposed by Hayashi and Miao (2008) and Ju and Miao (2007), who generalize the model of Klibanoff et al. (2008). 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, 1991). We can 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 has the lowest continuation utility 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 finding reported by Camerer (1999) 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 investment strategy. We compare this robust strategy with four other investment strategies widely studied in the literature: the IID strategy, the VAR strategy, the Bayesian strategy, and the Bayesian myopic strategy. The IID and VAR strategies refer to the optimal strategies when the investor knows for sure that the stock return follows an IID model and a VAR model, respectively. The Bayesian and Bayesian myopic strategies refer to the optimal strategies in the Bayesian framework when the investor behaves dynamically and myopically, respectively. The Bayesian myopic strategy is studied by Kandel and Stambaugh (1996).

We find the following results. First, the robust stock allocation depends on the investment horizon, the beliefs about the model of stock returns, and the predictive variable. Although the Bayesian strategy implies a similar stock allocation rule, the robust strategy is more conservative in that it recommends an ambiguity-averse investor to hold less stocks than a Bayesian investor, inducing more nonparticipation in the stock market. The intuition is that the intertemporal hedging demand plays an important role. We decompose the hedging demand into two components: a component associated with the predictive variable, and the other associated with model uncertainty. The former component is analyzed by Campbell and Viceira (1999) and Kim and Omberg (1996). The latter is positive when the predictive variable is small and negative, and is negative when the predictive is large and positive. This negative hedging component may induce an ambiguity-averse investor to hold less stocks when his investment horizon is longer, which is different from the advice of the VAR strategy.

Second, even though an ambiguity-averse investor attaches a small prior probability to the VAR model, he uses the predictive variable to time the market. Unlike the VAR strategy, but like the Bayesian strategy, the robust stock allocation may decrease with the predictive variable when it takes large values because of the negative hedging demand associated with model uncertainty. Compared to the Bayesian strategy, this decrease under the robust strategy is large for large extreme values of the predictive variable.

Third, when the predictive variable takes a large value, the VAR strategy advises the investor to allocate all his wealth to the stock. By contrast, the robust strategy advises the investor to sharply decreases his stock allocation to a level below that implied by the IID strategy, even when there is a small prior probability that the IID model is the true model of the stock return. As the investor raises his beliefs about the IID model, he gradually increases his stock allocation. Unlike this prediction, the Bayesian strategy advises the investor to gradually decreases his stock allocation as he attaches more beliefs about the IID model.

Finally, although the optimal stock demand is far more sensitive to the risk aversion parameter than to the ambiguity aversion parameter, the ambiguity aversion parameter is important for welfare implications. We compute the welfare costs of an ambiguity-averse investor if he follows the suboptimal IID, VAR, Bayesian, and Bayesian myopic investment strategies. We find that welfare costs depend crucially on the predictive variable used in the VAR estimation. In general, the VAR strategy is the most costly and the Bayesian strategy is the least costly. In terms of certainty equivalent wealth, the VAR strategy, the IID strategy, and the Bayesian strategy can cost as large as about two times, 50 percent, and 14 percent of the investor's initial wealth, respectively, when the net payout yield is used as the predictive variable in the VAR estimation.

Our paper is related to a large literature on the portfolio choice problem (see Campbell and Viceira (2002) for a survey). In addition to the papers cited above, other papers using the Bayesian framework include Brennan (1998), Brandt et al. (2005), Detemple (1986), Gennotte (1986), Gollier (2004), and Wachter and Warusawitharana (2007). All these papers do not consider model misspecification and ambiguity aversion. To study this issue in the portfolio choice problem, some papers use the multiple-priors approach or the robust control approach (Garlappi et al. (2007), Maenhout (2004), and Uppal and Wang (2003)). Other papers use one of these approaches to study equilibrium asset prices (Anderson et al. (2003), Chen and Epstein (2002), Epstein and Miao (2003), and Epstein and Wang (1994)). These papers do not allow for learning. Epstein and Schneider (2007, 2008) and Miao (2001) introduce learning to the recursive multiple-priors model. Hansen (2007) and Hansen and Sargent (2007a,b) develop models of learning in the robust control framework. Hansen and Sargent (2008) apply this framework to the study of the equilibrium price of model uncertainty.

To the best of our knowledge, the present paper provides the first portfolio choice model with learning under ambiguity using the generalized recursive ambiguity utility model. Compared to other models of ambiguity including the recursive smooth ambiguity model of Klibanoff et al. (2008), the recursive multiple-priors model of Epstein and Schneider (2003, 2007), and the robust control model of Hansen and Sargent (2001, 2007b), our generalized recursive ambiguity utility model has the unique advantage of combining smoothness and homotheticity. Both features are important to make analytical and quantitative studies tractable.

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 (2008), Klibanoff et al. (2005), Nau (2006), Neilson (2008), and Seo (2008). These papers provide various different axiomatic foundations and interpretations. Our adopted dynamic model is proposed by Hayashi and Miao (2008) and closely related to Klibanoff et al. (2008). Here we focus on the utility representation and refer the reader to the preceding papers for axiomatic foundations.

2.1. Utility

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

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

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 et al. (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.[2] 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

[2] The behavioral foundation of ambiguity and ambiguity attitude is based on the theory developed by Ghirardato and Marinacci (2002) and Klibanoff et al. (2005). Epstein (1999) provides a different foundation. The main difference is that the benchmark ambiguity neutral preference is the expected utility preference according to Ghirardato and Marinacci (2002), while Epstein's (1999) benchmark is the probabilistic sophisticated preferences.

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.

An important advantage of the smooth ambiguity model over other models of ambiguity, such as the widely adopted multiple-priors utility model of Gilboa and Schmeidler (1989), is that it is tractable and admits a clear-cut comparative statics analysis.[3] Tractability is revealed by the fact that the well-developed machinery for dealing with risk attitudes can be applied to ambiguity attitudes. In addition, the indifference curve implied by (2) is smooth under regularity conditions, rather than kinked as in the case of the multiple-priors utility model. More importantly, comparative statics of ambiguity attitudes can be easily analyzed using the function ϕ or v only, holding ambiguity fixed. Such a comparative static analysis is not evident for the multiple-priors utility model since the set of priors in that model may characterize ambiguity as well as ambiguity attitudes.

We now embed the static model (1) in a dynamic setting. Time is denoted by $t = 0, 1, 2, ..., T$, where T could be finite or infinity. 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. Inspired by Epstein and Zin (1989) and Kreps and Porteus (1978), we adopt the specification:

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

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

and consider the following homothetic recursive ambiguity utility function:[4]

$$V_t(C) = \left[C_t^{1-\gamma} + \beta \left\{ \mathbb{E}_{\mu_t} \left(\mathbb{E}_{\pi_{z,t}} \left[V_{t+1}^{1-\gamma}(C_{t+1}) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right\}^{\frac{1-\gamma}{1-\eta}} \right]^{\frac{1}{1-\gamma}}, \tag{5}$$

where $\beta \in (0, 1)$ is the subjective discount factor, 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

[3] Klibanoff et al. (2005) show that the multiple-priors model is a limiting case of the smooth ambiguity model with infinity ambiguity aversion.

[4] Hayashi and Miao (2008) propose a more general recursive ambiguity utility model that permits a three-way separation among risk aversion, ambiguity aversion, and intertemporal substitution. Hayashi and Miao (2008) also provide an axiomatic foundation for this utility model.

and (5) reduces to the standard time-additive expected utility model. In this case, the posterior μ_t and the likelihood $\pi_{z,t}$ 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 $\eta > \gamma$.

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

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

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

Klibanoff et al. (2008) propose the following closely related recursive smooth ambiguity model:

$$V_t(C) = u(C_t) + \beta \phi^{-1} \left(\mathbb{E}_{\mu_t} \phi \left(\mathbb{E}_{\pi_{z,t}} \left[V_{t+1}(C) \right] \right) \right), \tag{7}$$

where u and ϕ admit the same interpretation as in the static model. Ju and Miao (2007) find that when u is defined in (3) and $\phi(x) = x^{1-\alpha}/(1-\alpha)$ for $x > 0$ and $1 \neq \alpha > 0$, the model in (7) is not well defined for $\gamma > 1$. Thus, they consider (5), which is ordinally equivalent to (7) when $\gamma \in (0,1)$. The utility function in (7) is always well defined for the specification $\phi(x) = -e^{-\frac{x}{\theta}}$ for $\theta > 0$. An interesting feature of this specification is that it has a connection with risk-sensitive control and robustness, as studied by Hansen (2007) and Hansen and Sargent (2007b). A disadvantage of this specification is that the utility function does not have a homogeneity property. Thus, the curse of dimensionality makes the numerical analysis of the decision maker's dynamic programming problem complicated, except for the special case with $u(c) = \log(c)$ (see Ju and Miao (2007)). As a result, we will focus on the homothetic specification (5) in our analysis below.

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) and Skiadas (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 to calibrate these parameters? In general, there are two approaches. First, one may derive equilibrium implications using the new utility model, and then use asset markets data to estimate preference parameters by matching moments or using other econometric methods. This approach is pioneered by Hansen and Singleton (1982) and Epstein and Zin (1991). 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 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). 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 lose anything. The experiments reported in Halevy (2007) show that most subjects prefer to bet on the first urn to 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

$$u^{-1}\left(\int_{\Pi}\int_{S}u\left(c\right)d\pi d\mu\left(\pi\right)\right)-v^{-1}\left(\int_{\Pi}v\left(u^{-1}\left(\int_{S}u\left(c\right)d\pi\right)\right)d\mu\left(\pi\right)\right). \tag{8}$$

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.[6] 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

$$\left(0.5\left(d+w\right)^{1-\gamma}+0.5w^{1-\gamma}\right)^{\frac{1}{1-\gamma}}-\left(0.5\left(d+w\right)^{1-\eta}+0.5w^{1-\eta}\right)^{\frac{1}{1-\eta}}, \tag{9}$$

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.]

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, to match this estimate, the ambiguity aversion parameter η

[5] Anderson et al. (2003) advocate to use model detection error probabilities to calibrate θ in $\phi\left(x\right)=-e^{-\frac{x}{\theta}}$. They interpret θ as a robustness parameter.

[6] Strictly speaking the bet deals with objective lotteries and the subjective probability measure may not be the same as the objective measure. Seo's (2008) utility model can accommodate the bet discussed in the paper. His utility model gives the same expression as (9) for the ambiguity premium.

must be in the range of 50-90 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. Preferences and Investment Opportunities

There are two tradeable assets: a risky stock and a riskfree bond. The stock has gross real stock return R_{t+1} from t to $t+1$. The riskfree bond has a constant gross real return R_f each period. Define log returns $r_{t+1} = \log(R_{t+1})$ and $r_f = \log(R_f)$. The investor faces the following two model specifications for the stock return process:

- Model 1 (IID):
$$r_{t+1} - r_f = m + \varepsilon_{1,t+1}, \tag{10}$$
where $\varepsilon_{1,t+1}$ is normally distributed white noise with mean zero and variance σ_1^2.

- Model 2 (VAR):
$$r_{t+1} - r_f = m + bx_t + \varepsilon_{2,t+1}, \tag{11}$$
$$x_{t+1} = \rho x_t + \varepsilon_{3,t+1}, \tag{12}$$
where x_t represents a demeaned predictive variable and $[\varepsilon_{2,t+1}, \varepsilon_{3,t+1}]'$ is normally distributed white noise with mean zero and covariance matrix
$$\Omega = \begin{bmatrix} \sigma_2^2 & \sigma_{23} \\ \sigma_{23} & \sigma_3^2 \end{bmatrix}. \tag{13}$$
Assume that $\varepsilon_{1,t+1}$ is independent of $\varepsilon_{2,t+1}$ and $\varepsilon_{3,t+1}$. More generally, x_t may be a vector of predictive variables. In our empirical application in Section 4.1, we will consider 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 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).

3.2. Belief 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 :

$$\mu_{t+1} = \frac{\mu_t f_1\left(r_{t+1}, r_f + m\right)}{\mu_t f_1\left(r_{t+1}, r_f + m\right) + (1 - \mu_t) f_2\left(r_{t+1}, r_f + m + bx_t\right)}, \tag{14}$$

where

$$f_z\left(y, a\right) = \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{\left(y - a\right)^2}{2\sigma_z^2}\right], \; z = 1, 2. \tag{15}$$

The intuition for how the investor updates his 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 depends on the predictive variable in the VAR model. If the predictive variable is above average (i.e., $x_t > 0$), the VAR model will predict above average returns. Assuming that the volatility of returns are similar in the two models (which is true in our estimation below), 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 in the stock in period t. We can then write the investor's budget constraint as

$$W_{t+1} = R_{p,t+1}\left(W_t - C_t\right), \tag{16}$$

where

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

is the portfolio return. We suppose that there are short-sale and margin restrictions such that $\psi_t \in [0, 1]$. Otherwise, wealth may be negative because R_{t+1} can go to 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 belief μ_t. Let $J_t(W_t, x_t, \mu_t)$ denote the value function. Then it satisfies the Bellman equation:

$$J_t(W_t, x_t, \mu_t) = \max_{C_t, \psi_t} \left[C_t^{1-\gamma} + \beta \left\{ \mu_t \left(\mathbb{E}_t^1 \left[J_{t+1}^{1-\gamma}(W_{t+1}, x_{t+1}, \mu_{t+1}) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right.\right. \tag{17}$$

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

subject to the budget constraint (16) and the belief dynamics (14). Here, $\mathbb{E}_t^1$ 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 equation (10) for r_{t+1} into (14), and then substitute the resulting expression for μ_{t+1} into $\mathbb{E}_t^1[J_{t+1}(W_{t+1}, x_{t+1}, \mu_{t+1})]$. Similarly, $\mathbb{E}_t^2$ 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 equation (11) for r_{t+1} into (14), and then substitute the resulting expression for μ_{t+1} into $\mathbb{E}_t^2[J_{t+1}(W_{t+1}, x_{t+1}, \mu_{t+1})]$.

In Appendix A (also see Ju and Miao (2007)), we derive the following Euler equation when the optimal portfolio weight ψ_t^* is an interior solution in $(0, 1)$:

$$\mathbb{E}_t[M_{z,t+1}(R_{t+1} - R_f)] = 0, \ t = 0, 1, ..., T-1, \tag{18}$$

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

$$M_{z,t+1} = \left(\mathbb{E}_t^z \left[R_{p,t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}} \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}, \ z = 1, 2. \tag{19}$$

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 (19). To interpret this distortion, we normalize the multiplicative factor and show in Appendix A that we can rewrite the Euler equation as:

$$0 = \hat{\mu}_t \mathbb{E}_t^1 \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{t+1} - R_f) \right] + (1 - \hat{\mu}_t) \mathbb{E}_t^2 \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{t+1} - R_f) \right], \tag{20}$$

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

$$\hat{\mu}_t = \frac{\mu_t \left(\mathbb{E}_t^1 \left[J_{t+1}^{1-\gamma} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\mu_t \left(\mathbb{E}_t^1 \left[J_{t+1}^{1-\gamma} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}} + (1-\mu_t) \left(\mathbb{E}_t^2 \left[J_{t+1}^{1-\gamma} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}}. \tag{21}$$

Note that $\hat{\mu}_t$ is the distorted belief about the IID model. Thus, an ambiguity averse investor's decision making is observationally equivalent to that of a standard Bayesian investor with distorted beliefs. A similar observational equivalence result also appears in the multiple-priors utility model and the robust control model. We shall emphasize that $\hat{\mu}_t$ is endogenous (preference dependent) in our model and cannot be generated from (14) using any prior μ_0 in a Bayesian framework. In addition, the pricing kernel (19) cannot be generated from any Bayesian model.

We can use (21) to compare the behaviors of an ambiguity-averse investor and a Bayesian investor. Suppose the investor obtains higher expected continuation utility when returns are generated by the VAR model than the IID model. Regardless of whether $\gamma \in (0, 1)$ or $\gamma > 1$, we have

$$\mathbb{E}_t^2 \left[J_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}} > \mathbb{E}_t^1 \left[J_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}.$$

Thus, when $\eta > \gamma$, equation (21) implies that $\hat{\mu}_t > \mu_t$. That is, an ambiguity-averse investor will attach more weight on the IID model. The opposite is true when the IID model generates higher expected continuation utility. We conclude that an ambiguity-averse investor always behaves pessimistically by distorting his belief towards the "worse" model for the stock return.

Does a more ambiguity-averse investor invest less in the stock? Not necessarily, as shown by Gollier (2007) 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 (11)-(12) to capture predictability. We estimate this system and the IID model (10) using annual data for the U.S. stock market over the period 1926-2005. For stock returns, we use the log returns (cum-dividend) of the CRSP value-weighted market portfolio (including 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 riskfree rates. All nominal quantities are deflated using the Consumer Price Index (CPI), taken from the Bureau of Labor Statistics. 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), Xia (2001), Campbell and Viceira (2002)), we first 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 B of Figure 1. This panel reveals that during the 1990s, the log dividend yield dropped significantly for a long period of time. Lettau and Van Nieuwerburgh (2008) argue that this change is a structural break in the mean of dividend yields.

Boudoukh et al. (2007) argue that the structural break could be due to the enactment of an SEC rule that encourages stock repurchases. They show that share repurchases and issuances become a more important part of total payouts over the last 20 years. They construct total payout yields (adjusting dividend yield for repurchases) and net payout yields (adjusting for both repurchases and issuances), and find significantly stronger evidence for return predictability using these new payout yields as predictors. Thus, we also use the total payout yield or the net payout yield to replace the dividend yield in the VAR estimation. We take the log total payout yield (ltp) and the transformed net payout yield (lnp) series from Boudoukh et al. (2007),[7] which is updated to cover the sample 1926-2005 and is available from Michael Roberts's homepage. Panels C and D of Figure 1 present these two series, respectively. These panels reveal that the two adjusted payout yields appear to avoid the structural break issue for dividend yields. The large negative value of lnp during the Great Depression is due to the negative net payout yield (about -3%) during that time.

We estimate the VAR model (11)-(12) 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 VAR models. We use these parameter values to conduct numerical analyses below.

[Insert Table 2 Here.]

[7] Boudoukh et al. (2007) define lnp as $\log(0.1 + \textit{net payout yield})$ to avoid taking the log of a negative net payout yield.

Panel A of Table 2 shows that the average annual real riskfree rate over the sample is 0.78 percent, with standard error 0.45 percent. In the model, we assume that the riskfree rate is constant, and is fixed at the sample mean. The sample mean and volatility of excess log returns are 5.77 and 19.71 percent respectively, which characterize the IID model of stock returns.

Panel B of Table 2 reveals that 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 0.1055, with standard error 0.0527 (t-statistic of 2). The R^2 is only 4.9%. 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 total payout yield (ltp) and especially the net payout yield (lnp) show stronger predictive power than the dividend yield (ldy) in this sample, with R^2 rising to 8.8 and 26.5 percent, respectively. The predictability coefficients are also statistically more significant, taking value 0.2037 for ltp, and 0.7526 for lnp. The t-statistics are 2.7 and 5.3, respectively.

The expected excess returns generated by the three predictors have very different properties. First, the volatility of the expected excess return differs across the three predictors. It is 4.9 and 5.8 percent in the cases of ldy and ltp, respectively, and almost doubles to 10.2 percent in the case of lnp. Second, the persistence of the expected returns is also very different. Since the predicted excess returns are assumed to be linear functions of the predictors, they inherit the persistence of the predictors. As a result, the expected excess returns implied by the log dividend yield are highly persistent, with a half life of 12.3 years, but less so for the total payout yield (ltp) (half life of 4.3 years) or the net payout yield (lnp) (half life of 1.7 years). Third, the correlations between unexpected returns and innovations in expected returns are different. The correlation is about -0.67 for ldy and ltp, but drops by half to -0.33 for lnp.

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.

[Insert Figure 2 Here.]

In Figure 2, we plot the posterior probabilities of the IID model using the historical data of stock returns and the three payout yields from 1926 to 2005. The prior in 1926 is set at 0.5. The three series of posterior probabilities all trend downward over time, suggesting that the data is overall more consistent with time-varying expected returns. The lower posterior probabilities in the case of ltp and lnp are consistent with the higher R^2 for these predictors in Table 2. For the log dividend yield, the posterior probability of the IID model is still above 0.1 towards the end of the sample, and the rise in posterior probabilities in the 1990s is consistent with the structural break and the resulting poor performance of the dividend yield as a predictor during that time. In the case of the log net payout yield, the posterior probability of the IID model drops to nearly 0 at the beginning of the 1930s, which is because the big drop in the net payout yield (see Panel D of Figure 1) in 1929 and 1930 "successfully predicted" the large negative returns in 1930 and 1931.

Although Figure 2 shows that historical data favor the VAR model over a long sample period from 1926-2005, there is always 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 above 0.1 toward the end of the sample, when the dividend yield is the predictor. Since different model specifications imply drastically different dynamics of stock returns, concerns about model misspecifications, 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.

4.2. Preference Parameters

We need to assign values to preference parameters β, γ, and η. We set $\beta = 0.99$ so that it is approximately equal to $1/(1 + r_f)$. We consider $\gamma \in \{2, 5, 10\}$.[8] These values are commonly used in the macroeconomics and finance literature. There is no independent study of the ambiguity aversion parameter η in the literature. We use the hypothetical 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 Bayesian framework. Finally, we consider a $T = 40$ years investment horizon.

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. We then study how it alters the market timing, uncertainty, and horizon effects often considered 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. We will consider other payout measures such as the total payout yield and the net payout yield as the predictive variable in Section 6.

5.1. Learning, Ambiguity Aversion, and Hedging Demand

As discussed 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

[8]In our quantitative study below, we find that the investor allocates all his wealth to the stock when γ is sufficiently small including the log case with $\gamma = 1$, while he does not invest in the stock when γ is sufficiently large.

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 variables (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 standard Bayesian framework. In this case, we denote the total stock demand as ψ_0^B and the myopic demand as ψ_0^{BM}. The former demand has been analyzed by Barberis (2000) and Xia (2001) and the latter has been analyzed by Kandel and Stambaugh (1996). We refer to these two investment strategies as the Bayesian strategy and Bayesian myopic strategy, respectively.

[Insert Table 3 Here.]

The bottom panel of Table 3 reports the stock demands when the stock return is described by the IID model (10) and the VAR model (11)-(12), 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). 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 is constant over time when the stock return is IID. By contrast, when the stock return is predictable as in the VAR model (11)-(12), the stock demand depends on the investment horizon and the predictive state variable. In particular, the stock demand increases with the predictive variable reflecting the market timing effect.

The top panel of Table 3 reveals that the risk aversion parameter γ has bigger 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 the models is low (i.e. μ is close to 0 or 1), and decreases with η for the various values of the state variables (μ_0, x_0) considered in Table 3.

[Insert Table 4 Here.]

Table 4 presents the 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). This table reveals several interesting results. First, for the parameters considered, the hedging demand accounts for a smaller fraction of the total demand than the myopic demand. Second, the hedging demand as a percentage of the total demand is larger when the investor attaches a higher prior on the VAR model (i.e., μ_0 is smaller).

This result is intuitive. When the investor is more confident with the predictability of stock returns, he will hedge more aggressively against changes in the investment opportunity as proxied by the predictive variable. Third, the hedging demand as a percentage of the total demand increases with the risk aversion parameter γ, but may not be monotonic with the ambiguity aversion parameter η. In addition, it is not sensitive to η. Finally, the hedging demand is generally positive for low values of the predictive variable and negative for high values of the predictive variable. However, when the investor attaches a very high prior on the IID model (i.e., μ_0 is large), the hedging demand may be negative for high and low values of the predictive variable, but positive for intermediate values. In addition, this case happens typically for low values of the ambiguity aversion parameter η.

To understand the properties of the hedging demand, we decompose it into two components. The first hedge component is associated with changes in the predictive variable x. This component has been analyzed by Campbell and Viceira (1999) and Kim and Omberg (1996). Recall that the shocks to the predictive variable is negatively correlated with the shocks to future stock returns (i.e., $\sigma_{23} < 0$). 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, 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 hold assets that deliver wealth in bad times, i.e., when investment opportunities are poor, he has a positive hedging demand.[9] If the expected excess return (or x) becomes sufficiently negative, a decline in expected future returns can represent an improvement in the investment opportunity set because it creates a profitable opportunity to short stocks. Thus, the investor has a negative hedging demand when x is sufficiently negative.

The second hedge component reflects the agent's incentive to hedge against model uncertainty (or changes in μ_t). This hedge component is positive when x is negative and small, and is negative when x is positive and large. 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 positive and large value of x, 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 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 is negative and small, an unexpectedly high stock return induces the investor to attach more weight on the IID model. The investor becomes more convinced that the high return is only temporary, and that the future investment opportunity deteriorates. As a result, the investor buys the stock if x is negative and small.

The results in Table 4 show that the second component of the hedging demand associated with model uncertainty tends to dominate the first component associated with changes in the

[9]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.

predictive variable. Moreover, this component tends to be more important when the investor is more ambiguity averse.

5.2. The Market Timing Effect

An important implication of return predictability for the portfolio choice is market timing. That is, the optimal stock allocation depends on the current value of the predictive variable. Figure 3 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 four investment strategies: the Bayesian strategy, the IID strategy, the VAR strategy, and the robust strategy. 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 high. When the predictive variable takes sufficiently small values, the investor does not hold the stock because expected excess returns are low.

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. 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 predictive variable is close to zero, the IID and VAR models of the stock returns are similar, and hence there is little model uncertainty. As a result, both the Bayesian strategy and the robust strategy offer very similar portfolio advice to the investor. When the demeaned predictive variable takes large and positive values, the robust strategy recommends a much smaller stock allocation than the VAR strategy and the Bayesian strategy.

[Insert Figure 3 Here.]

Figure 3 also presents the investor's trading strategy when he displays infinite ambiguity aversion with his utility function given by a special case of the Epstein and Schneider model (6). This figure illustrates that a more ambiguity averse investor does not necessarily invest less in the stock, as we point out in Section 3.3. 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 the lowest continuation utilities. The investor invests according to the IID strategy for sufficiently high values of the predictive variable because at these values the IID model gives the lowest continuation utilities. For intermediate values of the predictive variable, the investor times the market by increasing his stock allocations when the predictive variable increases.

5.3. 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 4 plots this uncertainty effect for the Bayesian strategy and the robust strategy at various values of the predictive variable. Panel A of this figure shows that 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 predictive variable takes a small negative value. In this case, the stock is likely to have a negative expected excess return, and hence the investor has no incentive to invest in the stock. As the investor increases his prior probability of the IID model, he starts to invest more in the stock. Panel B has a similar feature. Panel C shows that the investor invests a large fraction of his wealth in the stock when he believes that the stock return is more likely to follow the VAR model and when the predictive variable takes a high positive value. As he increases his prior probability of the IID model, he starts to invest less in the stock.

Panel D of Figure 4 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 takes high values (i.e., the expected excess return is high). However, 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% to below 30%. As he raises his prior beliefs about the IID model, he starts to invest more in the stock. This result is in sharp contrast to that obtained in the Bayesian framework: A Bayesian investor decreases his stock allocation monotonically when his prior probability of the IID model rises. The intuition is the following: When the predictive variable takes a very high value, the investor becomes more cautious and the hedging demand associated with model uncertainty becomes more negative. This negative hedging demand can be much larger for an ambiguity-averse investor than that for a Bayesian investor, which causes the sharp decline in the stock demand at low values of initial prior probability of the IID model.

[Insert Figure 4 Here.]

5.4. The Horizon Effect

When stock returns are predictable, the optimal stock allocation depends on the investment horizon. Figure 5 presents the horizon effect for the VAR strategy, Bayesian strategy, and the robust strategy, when we fix the beliefs at the value $\mu = 0.5$ and x at a value in $\{-0.4677, 0, 0.4677\}$ over time.[10] 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 5. 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.

[Insert Figure 5 Here.]

Figure 5 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 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 5.

5.5. Simulated Stock Allocations

Figure 6 presents the simulated series of the stock allocations for an investor with $T = 40$ years investment horizon under the IID strategy, the VAR strategy, the Bayesian strategy, and the robust strategy, using the historical stock returns and the dividend yields from 1968 to 2007. We also suppose the investor starts with an initial prior $\mu_0 = 0.5$ in year 1968. Figure 6 reveals that the stock allocation under the VAR strategy exhibits the highest variation over time. The VAR strategy advises the investor to invest most of his wealth in stocks during 1970s and 1980s because of the high dividend yields during that period as shown in panel B of Figure 1. This strategy also recommends the investor not to invest in the stock during the late 1990s because the dividend yield dropped substantially during that period, as shown in panel B of Figure 1. By contrast, the other three strategies all recommend the investor to invest a part of his wealth in the stock. In particular,

[10] One standard deviation of x is $\sigma_3/\sqrt{1-\rho^2} = 0.4677$.

the IID strategy advises the investor to invest about a constant 40 percent of his wealth in the stock over time.

[Insert Figures 6-7 Here.]

The Bayesian strategy advises the investor to time the market, but less aggressively than the VAR strategy. The robust strategy recommends an ambiguity-averse investor to invest more conservatively over time than the Bayesian strategy. We may understand the intuition behind this result using the dynamics of the distorted beliefs plotted in Figure 7. Since the ambiguity-averse investor slants his probabilities towards the model with low continuation value, which could be the IID model or the VAR model, the distorted belief $\hat{\mu}_t$ can be either below or above the Bayesian belief μ_t. This feature highlights the difference between an ambiguity-averse investor and a Bayesian investor with a different prior, which amounts to parallel shifting the blue line. During 1970s and 1980s, an ambiguity-averse investor slants his beliefs toward the IID model, causing him to invest less in the stock than a Bayesian investor. After 1990, an ambiguity-averse investor slants his beliefs toward the VAR model, also causing him to invest less in the stock than a Bayesian investor.

Figure 6 reveals that the stock market nonparticipation phenomenon was more likely to occur for the VAR strategy in the late 1990s and the early 2000s when the predictive variable (dividend yields) took low values. When we reduce the initial prior belief μ_0 about the IID model, the simulated paths of stockholdings for the Bayesian strategy and the robust strategy will shift down. Consequently, both the robust strategy and the Bayesian strategy can generate nonparticipation. Nonparticipation is more likely to happen for the conservative robust strategy than for the Bayesian strategy.

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 investor than other popular investment strategies such as the IID strategy, the VAR strategy, the Bayesian myopic strategy, and the Bayesian strategy. 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 the investor follows one of the preceding suboptimal investment strategies, but adjusts his consumption optimally. We then compute the investor's value function under the suboptimal investment strategy and compare it with the value function under the robust investment strategy.

Let the value function for the suboptimal strategy k be $J_t^k(W_t, x_t, \mu_t)$. As is standard in the literature, we define the welfare cost as the percentage wealth compensation m needed to leave the investor indifferent between the suboptimal investment strategy and the optimal robust investment strategy, i.e.,

$$J_0^k(W_0(1+m), x_0, \mu_0) = J_0(W_0, x_0, \mu_0). \tag{22}$$

Note that the welfare cost m depends on the state variable (x_0, μ_0). For each initial beliefs about the IID model μ_0, we can compute the average welfare cost using the long-run stationary distribution of x. We report the average welfare costs for four different values of μ_0 in Tables 5-7. In particular, Tables 5-7 are for the cases where the predictive variable is the dividend yield, the total payout yield, or the net payout yield, respectively, in the VAR estimation.

[Insert Table 5 Here.]

Tables 5-7 show that the welfare costs depend on the level of model uncertainty and the attitudes toward uncertainty (the risk aversion parameter γ and the ambiguity aversion parameter η). When the predictive variable is the dividend yield, the predictability coefficient and R^2 are very small, as shown in Table 2. This feature implies that the VAR model and the IID model are similar, and hence the investor faces small model uncertainty. This small model uncertainty leads to a small ambiguity premium, and hence small welfare costs when the investor follows the Bayesian strategy, as shown in Table 5. Table 5 also shows that this welfare cost becomes larger when the investor is more ambiguity averse. The largest welfare cost under the Bayesian strategy appears in the case where $\gamma = 2$, $\eta = 100$ and $\mu_0 = 0.8$. This cost accounts for 2.9 percent of the investor's initial wealth.

Turn to the welfare costs of following other suboptimal investment strategies. Table 5 shows that the VAR strategy is the most costly strategy. The second most costly one is the IID strategy. The Bayesian myopic strategy is the least costly strategy among these three strategies. In particular, the VAR strategy can cost 44.6 percent of the initial wealth for an investor with $\gamma = 10, \eta = 60$, and $\mu_0 = 0.2$.

[Insert Table 6 Here.]

We now consider Table 6 where the predictive variable is the total payout yield. In this case, the predictability coefficient and R^2 are larger as shown in Table 2, and it seems that the IID model and the VAR model are more farther apart. Thus, the investor faces a larger level of model uncertainty. This naturally leads to a larger welfare cost of following the Bayesian strategy because the investor bears a larger ambiguity premium. More specifically, the welfare cost is as large as 4.5 percent of the initial wealth for an investor with $\gamma = 2, \eta = 100$, and $\mu_0 = 0.8$. Other features of Table 6 are similar to those of Table 5. A notable feature of Table 6 is that the VAR strategy is more costly than that in Table 5. In particular, the welfare cost can be more than an investor's total initial wealth for $\gamma = 10$.

[Insert Table 7 Here.]

We finally consider Table 7 where the predictive variable is the net payout yield. In this case, the predictability coefficient and R^2 are the largest as shown in Table 2, and hence the level of model uncertainty is the largest. This result implies that the welfare cost of following the Bayesian strategy in Table 7 is the largest among Tables 5-7. For reasonable values of the risk aversion parameter $\gamma = 2, 5$, and the ambiguity aversion parameter $\eta = 60, 100$, the welfare cost of following the Bayesian strategy is sizable and ranges from about 3 percent to 14 percent of the investor's initial wealth. In addition, the welfare cost of ignoring model uncertainty by following the VAR strategy can be close to two times of the initial wealth for an investor with $\gamma = 10$. This result is striking because we find in Section 4.1 that historical data favor the VAR model most when the net payout yield is the predictor. The intuition is the following. Expected stock returns display the largest time variation when the net payout yield is the predictor. Thus, the VAR strategy advises the investor to time the market very aggressively. This strategy is very costly, compared to the robust strategy which takes into account the possibility that the stock return may be IID.

7. Conclusion

Whether or not 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 approach. He copes with the specification doubts by slanting his beliefs pessimistically. We find that an ambiguity-averse investor's robust investment strategy is different from some other investment strategies often studied in the literature. In particular, the robust strategy is more conservative than the Bayesian strategy. This effect is especially large for extreme values of the predictive variable. For low extreme values, an ambiguity-averse investor is more likely to not participate in the stock market. For high extreme values of the predictive variable, the robust strategy recommends much less stockholdings than the Bayesian strategy or the VAR strategy. We also find that a very small prior probability of the IID model can lead an ambiguity-averse investor to decreasing his stock allocation sharply, unlike the prediction in the Bayesian approach.

We find that the welfare costs of following suboptimal investment strategies by ignoring model uncertainty can be sizable. Under reasonably calibrated parameter values, the welfare costs of following the VAR strategy and the Bayesian strategy can be about two times and 14 percent of the investor's initial wealth, respectively.

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. We leave this extension for future research.

Appendices

A Proofs of Results in Section 2.4

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 (17) to derive:

$$\begin{aligned} W_t G_t\left(x_t, \mu_t\right) = \max_{C_t, \psi_t} \Bigg[C_t^{1-\gamma} + \beta \Bigg\{ \mu_t \left(\mathbb{E}_t^1 \left[W_{t+1}^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \\ + \left(1-\mu_t\right) \left(\mathbb{E}_t^2 \left[W_{t+1}^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \Bigg\}^{\frac{1-\gamma}{1-\eta}} \Bigg]^{\frac{1}{1-\gamma}}. \end{aligned} \tag{A.2}$$

We substitute the budget constraint (16) into the above Bellman equation to obtain:

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

where we define

$$\begin{aligned} H_t\left(\psi_t, x_t, \mu_t; G_{t+1}\right) = \Bigg\{ \mu_t \left(\mathbb{E}_t^1 \left[\left(R_{t+1}\psi_t + R_f\left(1-\psi_t\right)\right)^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \\ + \left(1-\mu_t\right) \left(\mathbb{E}_t^2 \left[\left(R_{t+1}\psi_t + R_f\left(1-\psi_t\right)\right)^{1-\gamma} G_{t+1}^{1-\gamma}\left(x_{t+1}, \mu_{t+1}\right) \right] \right)^{\frac{1-\eta}{1-\gamma}} \Bigg\}^{\frac{1-\gamma}{1-\eta}}. \end{aligned} \tag{A.4}$$

We use the first-order condition for C_t to derive

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

Solving yields a linear consumption rule:

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

where we define

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

We may equivalently write the portfolio choice problem as

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

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

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

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

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

$$\begin{aligned}G_t\left(x_t,\mu_t\right) &= \left[a_t^{1-\gamma}+\beta\left(1-a_t\right)^{1-\gamma}H_t\left(\psi_t,x_t,\mu_t;G_{t+1}\right)\right]^{\frac{1}{1-\gamma}}\\ &= \left[1+\left(\beta H_t\left(\psi_t,x_t,\mu_t;G_{t+1}\right)\right)^{1/\gamma}\right]^{\frac{\gamma}{1-\gamma}}.\end{aligned}\tag{A.10}$$

Thus,

$$G_t=a_t^{\frac{-\gamma}{1-\gamma}}=\left(\frac{C_t}{W_t}\right)^{\frac{-\gamma}{1-\gamma}}.\tag{A.11}$$

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

$$0=\mu_t\left(\mathbb{E}_t^1\left[\left(R_{t+1}\psi_t+R_f\left(1-\psi_t\right)\right)^{1-\gamma}\left(\frac{C_{t+1}}{W_{t+1}}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\tag{A.12}$$
$$\times\left\{\mathbb{E}_t^1\left[\left(R_{t+1}\psi_t+R_f\left(1-\psi_t\right)\right)^{-\gamma}\left(\frac{C_{t+1}}{W_{t+1}}\right)^{-\gamma}\left(R_{t+1}-R_f\right)\right]\right\}$$

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

Note that $W_{t+1}=R_{p,t+1}\left(W_t-C_t\right)=R_{p,t+1}\left(1-a_t\right)C_t$. So we have

$$0=\mu_t\left(\mathbb{E}_t^1\left[R_{p,t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_t^2\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R_{t+1}-R_f\right)\right]\right\}\tag{A.13}$$
$$+\left(1-\mu_t\right)\left(\mathbb{E}_t^2\left[R_{p,t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}\left\{\mathbb{E}_t^2\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R_{t+1}-R_f\right)\right]\right\}.$$

From this equation, we deduce that the pricing kernel is given by equation (19). Using this equation and the preceding equation, we can write the Euler equation as equation (18). We can also rewrite it as equation (20), where $\hat{\mu}_t$ is distorted beliefs about the IID model given by:

$$\hat{\mu}_t=\frac{\mu_t\left(\mathbb{E}_t^1\left[R_{p,t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\mu_t\left(\mathbb{E}_t^1\left[R_{p,t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}+\left(1-\mu_t\right)\left(\mathbb{E}_t^2\left[R_{p,t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}.\tag{A.14}$$

Using (A.11) and the conjectured value function, we deduce that

$$J_t\left(W_t,x_t,\mu_t\right)=\left[C_t^{-\gamma}W_t\right]^{\frac{1}{1-\gamma}}.\tag{A.15}$$

Using this equation and the fact that $W_{t+1}=R_{p,t+1}\left(W_t-C_t\right)$, we can then rewrite (A.14) as (21).

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 $\left[-4\sigma_3/\sqrt{1-\rho^2}, 4\sigma_3/\sqrt{1-\rho^2}\right] \times [0,1]$. We discretize this space using 401×51 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:

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

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.11) 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.11), 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 $k \in \{1, 2, 3, 4\}$ index one of the four suboptimal investment strategies. We conjecture that the value function takes the form $J_t^k\left(W_t, x_t, \mu_t\right) = W_t G_t^k\left(x_t, \mu_t\right)$, where G_t^k is a function to be determined. We start with the last period $G_T^k = 1$ and then solve backward. Suppose we know G_{t+1}^k at date t. Given a suboptimal investment strategy ψ_t^k, we plug it in equation (A.4). For this strategy, we use equation (A.7) to obtain the optimal consumption-wealth ratio:

$$a_t = \frac{1}{1 + \left(\beta H_t\left(\psi_t^k, x_t, \mu_t; G_{t+1}^k\right)\right)^{1/\gamma}}. \tag{B.2}$$

We next use equation (A.10) to derive G_t^k :

$$G_t^k\left(x_t, \mu_t\right) = \left[1 + \left(\beta H_t\left(\psi_t^k, x_t, \mu_t; G_{t+1}\right)\right)^{1/\gamma}\right]^{\frac{\gamma}{1-\gamma}}. \tag{B.3}$$

We then solve backward for G_t^k for all $t = T-1, T-2, ..., 0$. We finally obtain $G_0^k\left(\mu_0, x_0\right)$ and confirm our conjecture.

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.

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

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, *Journal of Finance* 62, 877-916.

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.

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.

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, forthcoming in *Journal of Economic Theory.*

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.

Epstein, Larry G., 1999, A Definition of Uncertainty Aversion, *Review of Economic Studies* 66, 579-608.

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 Martin Schneider, 2008, Ambiguity, Information Quality and Asset Pricing, *Journal of Finance* 63, 197-228.

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.

Epstein, Larry G. and Stanley Zin, 1991, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis, *Journal of Political Economy* 99, 263-286.

Ergin, Huluk I. and Faruk Gul, 2008, A Theory of Subjective Compound Lotteries, *Journal of Economic Theory*, forthcoming.

Garlappi, Lerenzo, 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.

Ghirardato, Paolo and Massimo Marinacci, 2002, Ambiguity Made Precise: A Comparative Foundation, *Journal of Economic Theory* 102, 251-289.

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, 2007, Does Ambiguity Aversion Reinforce Risk Aversion? Applications to Portfolio Choices and Asset Prices, working paper, University of Toulouse.

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, Forthcoming in *American Economic Review, Papers and Proceedings.*

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, 2008, Fragile Beliefs and the Price of Model Uncertainty, working paper, University of Chicago and New York University.

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, 2008, Generalized Recursive Ambiguity Preferences, work in progress.

Ju, Nengjiu and Jianjun Miao, 2007, Ambiguity, Learning, and Asset Returns, working paper, Boston University and Hong Kong University of Science and Technology.

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, 2008, Recursive Smooth Ambiguity Preferences, working paper, Northwestern University and Collegio Carlo Alberto.

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

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

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, 2001, Ambiguity, Risk, and Portfolio Choice under Incomplete Information, working paper, Boston University.

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

Neilson, William S., 2008, Ambiguity Aversion: An Axiomatic Approach Using Second Order Probabilities, Working Paper, University of Tennessee.

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, 2008, Ambiguity and Second-Order Belief, *Econometrica*, forthcoming.

Skiadas, Costis, 2008, Smooth Ambiguity Aversion toward Small Risk and Continuous-Time Recursive Utility, working paper, Northwestern University.

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

Wachter, Jessica and Missaka Warusawitharana, 2007, Predictable Returns and Asset Allocation: Should a Skeptical Investor Time the Market? Working Paper, The Wharton School.

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

A. IID Model

m	σ_1	r_f
0.0577	0.1971	0.0078
(0.0219)	(0.0156)	(0.0045)

B. VAR Model

Variable	m	b	ρ	σ_2	σ_{23}	σ_3	R^2
ldy	0.0577	0.1055	0.9452	0.1935	-0.0199	0.1527	0.0488
	(0.0219)	(0.0527)	(0.0416)	(0.0154)	(4.01E-3)	(0.0121)	
ltp	0.0577	0.2037	0.8501	0.1894	-0.0191	0.1504	0.0881
	(0.0219)	(0.0742)	(0.0589)	(0.0151)	(3.86E-3)	(0.0120)	
lnp	0.0577	0.7526	0.6539	0.1700	-0.0057	0.1028	0.2653
	(0.0219)	(0.1418)	(0.0857)	(0.0135)	(2.07E-3)	(0.0082)	

Notes: This table reports the results from estimating the riskfree rates and the IID and VAR models of stock returns. The numbers in brackets are standard errors. The variable *ldy* denotes the log dividend yield. The variables *ltp* and *lnp* denote the total payout yield and the net payout yield series constructed by Boudoukh et al. (2007). The sample period is 1926-2005. All variables are annualized when applicable.

Table 3: Optimal Portfolio Weights in percentage

(γ, η)	$\mu_0 = 0.2$ x_0^1	x_0^2	x_0^3	$\mu_0 = 0.5$ x_0^1	x_0^2	x_0^3	$\mu_0 = 0.8$ x_0^1	x_0^2	x_0^3
2, 2	57.1	100.0	100.0	69.5	100.0	100.0	84.5	100.0	100.0
2, 60	52.0	100.0	98.0	55.3	100.0	98.5	64.5	100.0	99.1
2, 80	51.0	100.0	94.4	52.4	100.0	97.6	59.5	100.0	98.9
2, 100	50.3	100.0	93.3	50.3	100.0	97.4	55.6	100.0	98.9
5, 5	28.6	49.7	57.2	30.0	43.2	48.8	33.8	40.4	43.0
5, 60	28.1	46.4	42.6	27.7	42.2	41.3	31.1	40.3	40.3
5, 80	27.9	45.7	39.8	27.0	42.0	40.1	30.1	40.3	39.8
5, 100	27.7	45.2	37.8	26.3	41.9	39.3	29.2	40.2	39.6
10, 10	15.5	25.2	27.2	15.4	21.6	23.8	16.9	20.1	21.2
10, 60	15.4	24.1	23.3	14.9	21.3	21.7	16.3	20.1	20.4
10, 80	15.4	23.8	22.1	14.7	21.2	21.1	16.1	20.1	20.2
10, 100	15.3	23.5	21.1	14.5	21.2	20.7	15.9	20.0	20.1

γ	IID Model	VAR Model x_0^1	x_0^2	x_0^3
2	99.3	48.8	100.0	100.0
5	39.6	27.1	63.0	98.3
10	19.7	15.2	34.4	54.9
15	13.1	10.5	23.7	37.8

Notes: This table presents the optimal portfolio weights in percentage allocated to the stock. Column 1 denotes various values of γ and η. Columns 2-10 report the optimal portfolio weights in percentage for different combinations of μ_0 and x_0 where $x_0^1 = -\sigma_3/\sqrt{1-\rho^2} = -0.4677$, $x_0^2 = 0.0$, $x_0^3 = \sigma_3/\sqrt{1-\rho^2} = 0.4677$. The predictive variable x is the price-dividend ratio.

Table 4: PERCENTAGE HEDGING DEMAND OVER TOTAL STOCK DEMAND

	$\mu_0 = 0.2$			$\mu_0 = 0.5$			$\mu_0 = 0.8$		
(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
2, 2	18.1	0.6	0.0	4.3	0.6	0.0	-1.9	0.7	0.0
2, 60	26.3	0.6	-2.1	15.6	0.6	-1.5	6.9	0.7	-0.9
2, 80	27.6	0.6	-5.9	17.4	0.6	-2.5	9.0	0.7	-0.7
2, 100	28.5	0.6	-7.2	18.8	0.6	-2.4	10.6	0.7	-0.5
5, 5	35.2	20.1	-6.7	12.3	8.2	-7.6	-1.2	1.9	-3.9
5, 60	40.1	14.6	-22.3	19.4	6.1	-9.3	4.0	1.6	-2.9
5, 80	41.3	13.3	-23.9	21.2	5.7	-8.7	6.1	1.5	-2.6
5, 100	42.3	12.2	-24.5	22.7	5.3	-8.0	8.0	1.5	-2.2
10, 10	40.4	21.8	-11.5	15.2	8.9	-9.4	-0.8	2.1	-4.4
10, 60	42.9	18.2	-21.7	18.9	7.6	-11.1	1.6	2.0	-4.0
10, 80	43.7	17.1	-24.3	20.2	7.3	-11.1	2.7	1.9	-3.8
10, 100	44.5	16.2	-26.2	21.3	6.9	-11.0	3.8	1.8	-3.5

Notes: This table presents the ratio of the hedging demand to the total stock demand in percentage. Column 1 denotes various values of 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 $x_0^1 = -\sigma_3/\sqrt{1-\rho^2} = -0.4677$, $x_0^2 = 0.0$, $x_0^3 = \sigma_3/\sqrt{1-\rho^2} = 0.4677$. The predictive variable x is the price-dividend ratio.

Table 5: Welfare Costs of Following Suboptimal Investment Strategies

γ/η	γ	60	100	γ	60	100	γ	60	100	γ	60	100
	$\mu_0 = 0.2$						$\mu_0 = 0.4$					
	IID			VAR			IID			VAR		
2	11.8	13.6	13.8	0.3	0.2	0.1	9.1	13.0	13.6	1.1	0.3	0.2
5	7.1	5.7	5.5	16.6	22.7	23.6	4.4	4.1	4.2	19.6	22.9	23.3
10	3.5	3.0	2.9	43.7	44.6	44.5	2.2	2.0	2.0	39.4	40.5	40.7
	BM			Bayesian			BM			Bayesian		
2	0.4	0.5	0.6	0.0	0.2	0.3	0.2	0.7	0.9	0.0	0.5	0.7
5	1.6	2.5	2.8	0.0	0.9	1.5	0.6	1.2	1.3	0.0	0.5	0.7
10	1.1	1.3	1.4	0.0	0.2	0.4	0.4	0.5	0.6	0.0	0.1	0.2
	$\mu_0 = 0.6$						$\mu_0 = 0.8$					
	IID			VAR			IID			VAR		
2	6.6	12.5	13.4	2.4	0.8	0.5	4.1	11.5	13.0	4.5	1.5	0.8
5	3.1	3.5	3.7	20.7	22.6	22.7	2.1	2.9	3.3	21.4	22.2	22.2
10	1.5	1.6	1.6	35.8	37.2	37.8	1.1	1.2	1.4	32.9	34.2	35.0
	BM			Bayesian			BM			Bayesian		
2	0.1	1.4	1.9	0.0	1.1	1.6	0.0	2.4	3.4	0.0	1.9	2.9
5	0.3	0.6	0.8	0.0	0.3	0.5	0.1	0.4	0.7	0.0	0.2	0.4
10	0.2	0.3	0.3	0.0	0.0	0.1	0.1	0.1	0.2	0.0	0.0	0.1

Notes: This table reports the welfare costs of following four suboptimal investment strategies: the IID strategy, the VAR strategy, the Bayesian myopic (BM) strategy, and the Bayesian strategy, when the investor is averse to model uncertainty. Column 1 denotes various values of γ. Welfare costs are measured by percentage of initial wealth. Row 1 denotes various values of η. When $\eta = \gamma$, the model reduce to the standard Bayesian framework. The predictive variable is the dividend yield.

Table 6: Welfare Costs of Following Suboptimal Investment Strategies

γ/η	γ	60	100	γ	60	100	γ	60	100	γ	60	100
	$\mu_0 = 0.2$						$\mu_0 = 0.4$					
	IID			VAR			IID			VAR		
2	12.4	14.2	14.5	0.4	0.3	0.4	8.9	13.5	14.4	1.5	0.7	0.6
5	10.4	7.6	7.3	18.6	30.4	32.5	5.7	4.7	4.8	23.7	30.7	31.6
10	5.3	4.4	4.1	136.7	141.9	142.1	2.9	2.4	2.4	128.0	132.8	133.6
	BM			Bayesian			BM			Bayesian		
2	0.4	0.7	0.9	0.0	0.4	0.7	0.2	0.9	1.4	0.0	0.9	1.5
5	2.1	4.2	4.9	0.0	2.1	3.2	0.9	2.1	2.5	0.0	1.2	1.8
10	1.6	2.1	2.5	0.0	0.4	0.9	0.7	1.0	1.1	0.0	0.2	0.5
	$\mu_0 = 0.6$						$\mu_0 = 0.8$					
	IID			VAR			IID			VAR		
2	6.0	13.1	14.5	3.0	1.1	0.7	3.4	12.3	14.4	5.2	1.9	1.1
5	3.5	3.6	4.0	26.1	30.2	30.7	1.9	2.9	3.5	27.6	29.7	29.9
10	1.7	1.7	1.7	120.4	126.2	128.1	1.0	1.1	1.3	113.4	119.8	122.8
	BM			Bayesian			BM			Bayesian		
2	0.1	1.7	2.6	0.0	1.7	2.7	0.0	3.2	4.6	0.0	3.0	4.5
5	0.4	1.1	1.5	0.0	0.7	1.2	0.1	0.7	1.1	0.0	0.5	1.0
10	0.3	0.5	0.6	0.0	0.1	0.3	0.1	0.2	0.3	0.0	0.1	0.2

Notes: This table reports the welfare costs of following four suboptimal investment strategies: the IID strategy, the VAR strategy, the Bayesian myopic (BM) strategy, and the Bayesian strategy, when the investor is averse to model uncertainty. Welfare costs are measured by percentage of initial wealth. Column 1 denotes various values of γ. Row 1 denotes various values of η. When $\eta = \gamma$, the model reduce to the standard Bayesian framework. The predictive variable is the total payout yield.

Table 7: Welfare Costs of Following Suboptimal Investment Strategies

γ/η	γ	60	100	γ	60	100	γ	60	100	γ	60	100
	$\mu_0 = 0.2$						$\mu_0 = 0.4$					
	IID			VAR			IID			VAR		
2	48.6	48.9	49.3	1.7	3.3	3.6	34.2	46.4	48.4	4.2	3.7	3.2
5	29.2	21.5	21.0	18.9	36.5	37.4	19.1	16.9	17.4	25.9	35.6	35.4
10	14.3	11.5	10.9	165.7	189.7	190.4	9.5	8.4	8.3	170.9	185.4	185.3
	BM			Bayesian			BM			Bayesian		
2	0.1	5.3	8.1	0.0	5.3	8.2	0.1	6.2	9.1	0.0	6.4	9.4
5	1.1	8.6	10.5	0.0	5.2	7.2	0.9	5.4	6.9	0.0	3.4	5.0
10	1.0	3.7	4.6	0.0	1.3	2.2	0.7	2.1	2.7	0.0	0.8	1.4
	$\mu_0 = 0.6$						$\mu_0 = 0.8$					
	IID			VAR			IID			VAR		
2	22.1	45.3	48.8	7.3	4.1	3.1	12.0	44.1	49.3	11.1	5.0	3.3
5	12.4	14.4	15.8	30.0	35.1	34.5	7.1	12.2	14.5	33.2	35.1	34.2
10	6.3	6.5	6.9	173.1	183.1	182.7	3.8	4.8	5.7	174.6	181.7	181.2
	BM			Bayesian			BM			Bayesian		
2	0.1	7.3	10.5	0.0	7.6	10.9	0.0	9.4	13.2	0.0	9.9	13.9
5	0.6	3.5	5.0	0.0	2.6	4.3	0.3	2.5	4.2	0.0	2.3	4.2
10	0.4	1.2	1.7	0.0	0.5	1.1	0.2	0.7	1.1	0.0	0.4	0.9

Notes: This table reports the welfare costs of following four suboptimal investment strategies: the IID strategy, the VAR strategy, the Bayesian myopic (BM) strategy, and the Bayesian strategy. Welfare costs are measured by percentage of initial wealth, when the investor is averse to model uncertainty. Column 1 denotes various values of γ. Row 1 denotes various values of η. When $\eta = \gamma$, the model reduce to the standard Bayesian framework. The predictive variable is the net payout yield.

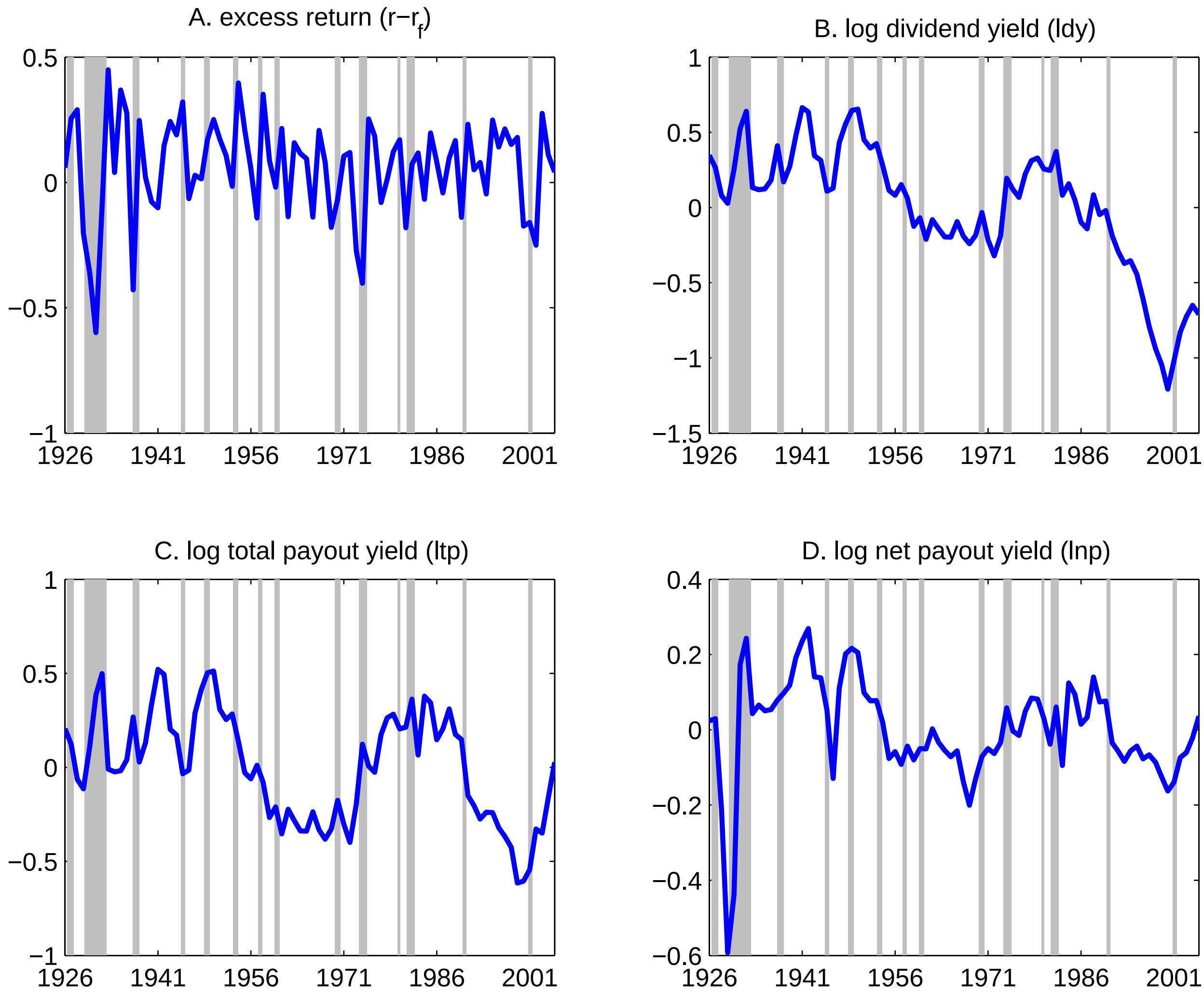

Figure 1: **Returns and predictors.** This figure plots the historical data of the real return from the CRSP value-weighted market portfolio, the demeaned log dividend yield, and the demeaned log total and net payout yields from 1926-2005. The latter two payout yields are constructed by Boudoukh et al. (2007) and downloaded from Michael Roberts' homepage.

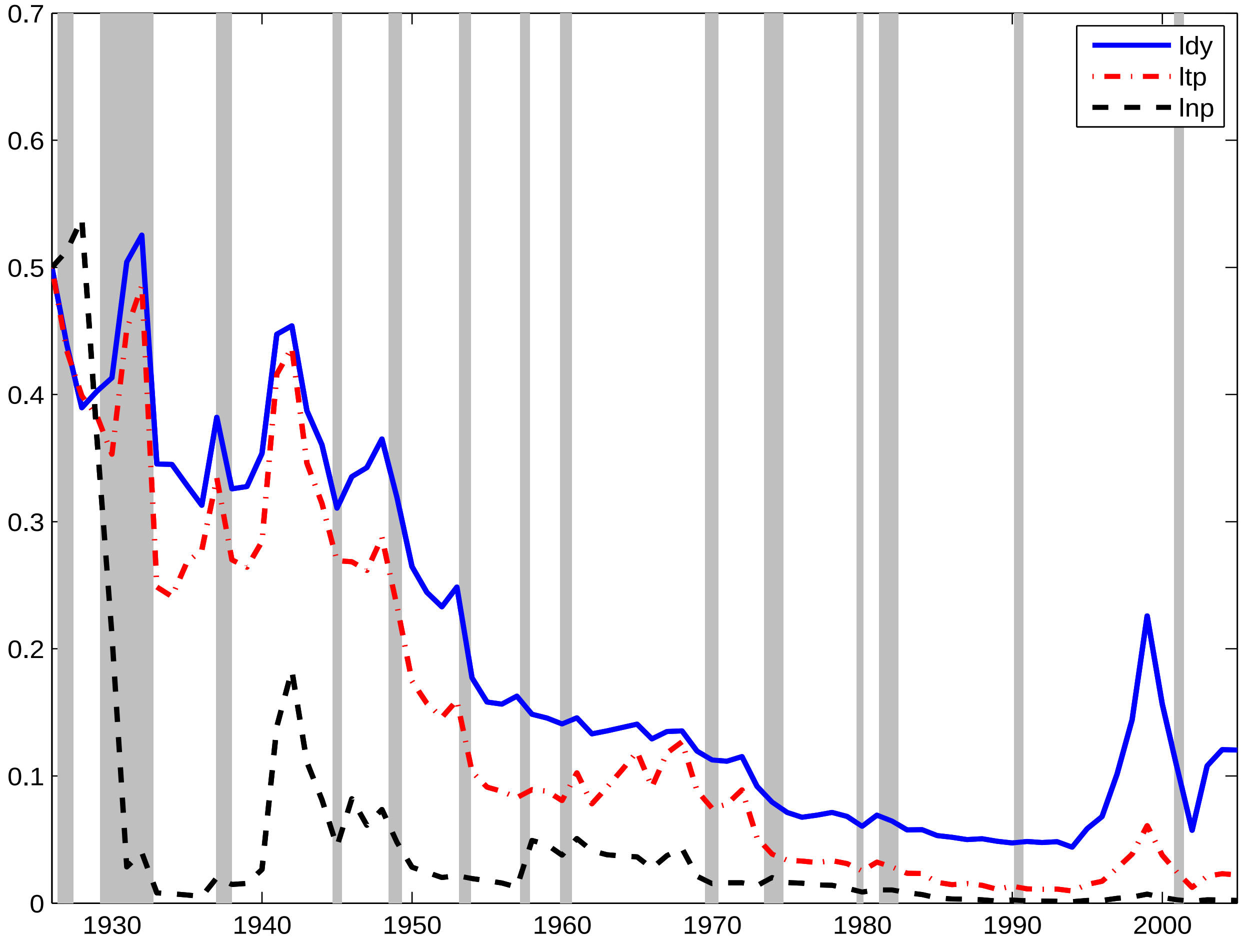

Figure 2: **Posterior probabilities of the IID model.** This figure plots the posterior probabilities of the IID model using the historical annual data of stock returns and the three payout yields as predictors from 1926-2005. The prior is set at $\mu_0 = 0.5$. The log dividend yield (ldy): solid line; the log total payout yield (ltp): dashdot line; the log net payout yield (lnp): dashed line.

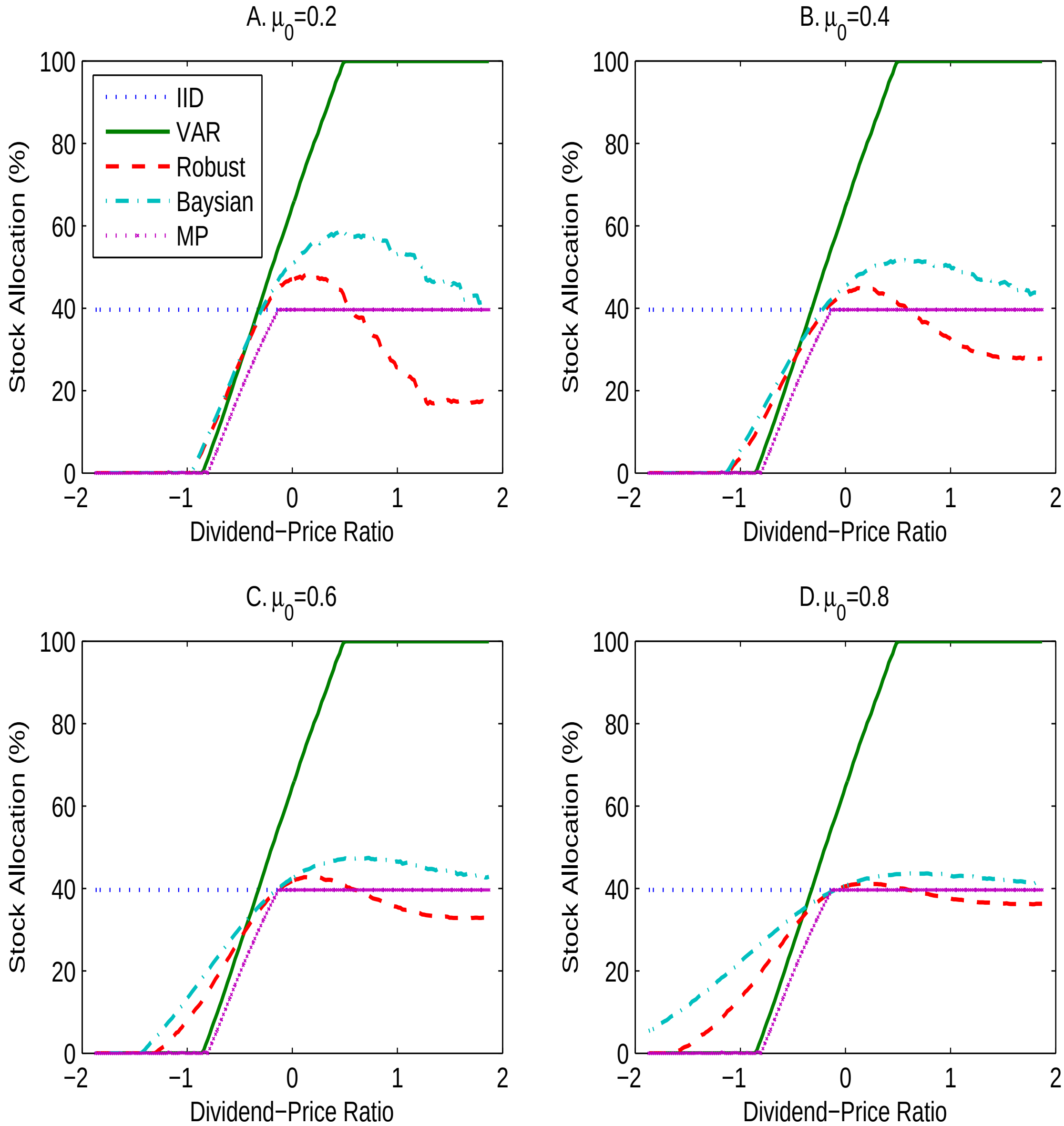

Figure 3: **Alternative investment strategies: market-timing effect.** This figure plots the initial portfolio weights on the stock for five alternative investment strategies as functions of the initial observation of the dividend yield given four different values of initial beliefs about the IID model. We set the risk aversion parameter $\gamma = 5$ and the investment horizon $T = 40$. Other parameter values are estimated using annual data as reported in Table 2. Bayesian strategy: dashdot line; Robust strategy for $\eta = 60$: dashed line; VAR strategy: solid line; IID strategy: horizontal dot line; Multiple-prior (MP) strategy: dotted line with x-marks.

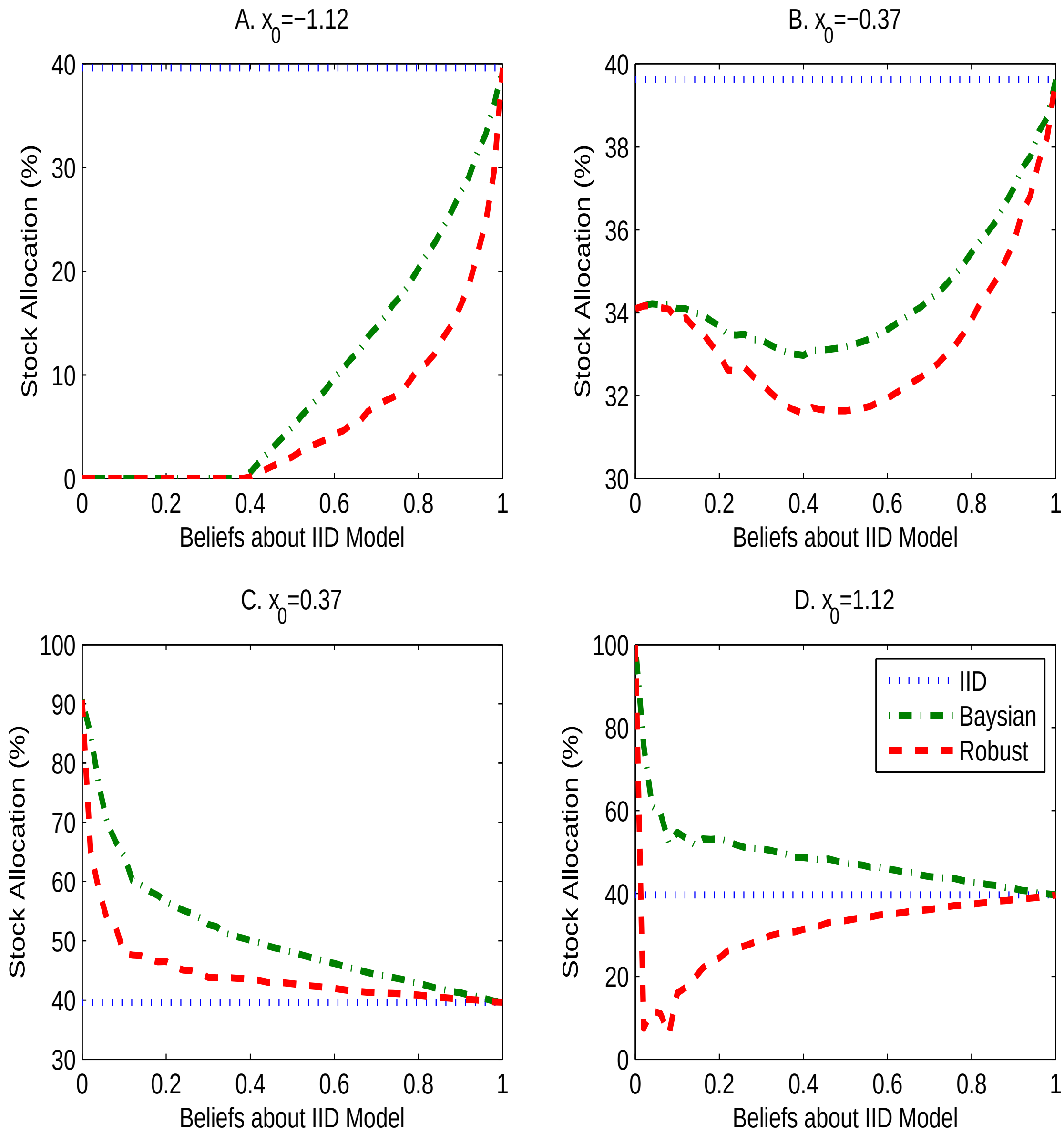

Figure 4: **Alternative investment strategies: uncertainty effect.** This figure plots the portfolio weights on the stock for three alternative investment strategies as functions of the initial beliefs about the IID model given four different initial values of the dividend yield. One standard deviation of x_t is $\sigma_3/\sqrt{1-\rho^2} = 0.4677$. We set the risk aversion parameter $\gamma = 5$ and the investment horizon $T = 40$. Other parameter values are estimated using annual data as reported in Table 2. Bayesian strategy: dashdot line; Robust strategy for $\eta = 60$: dashed line; IID strategy: horizontal dotted line.

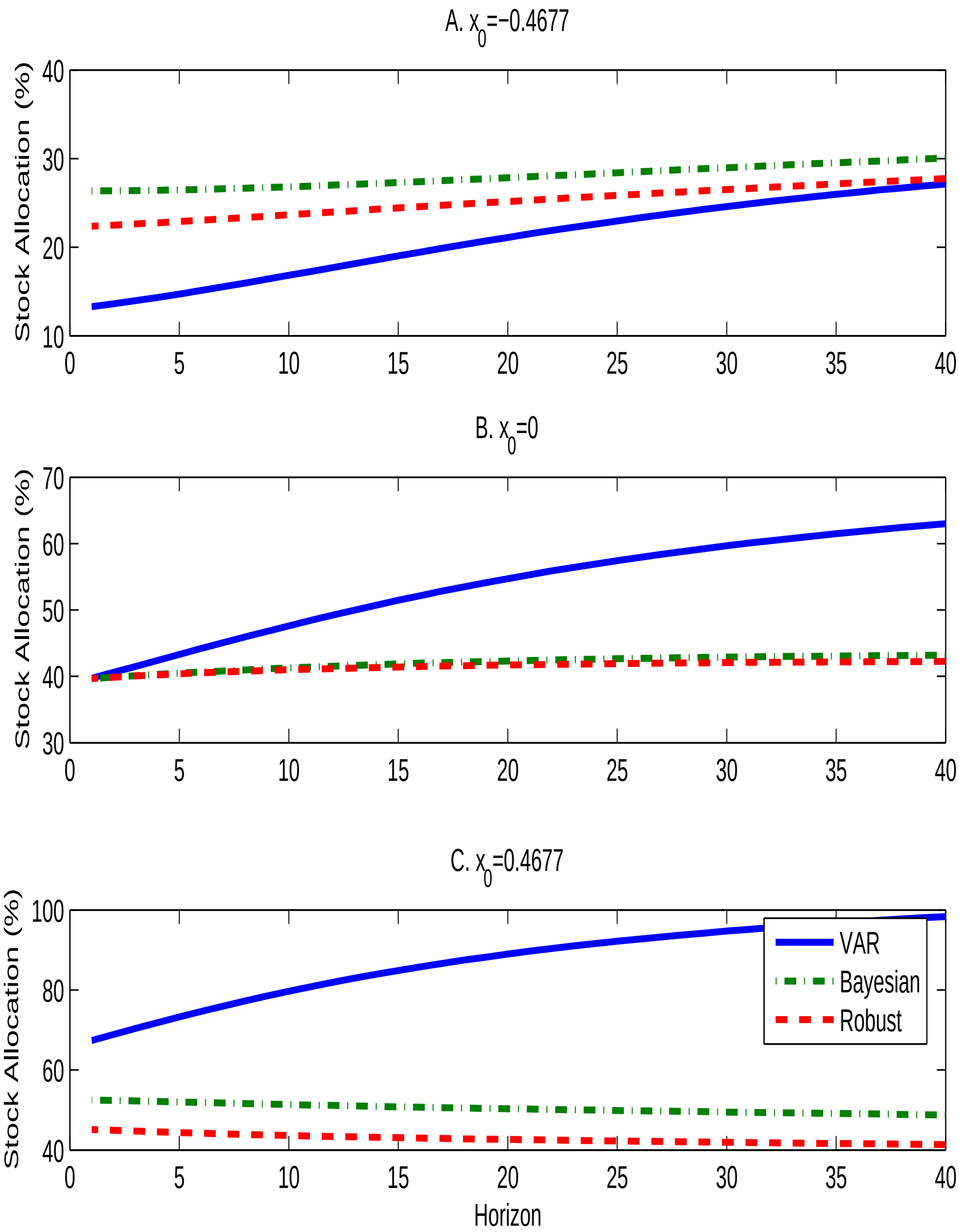

Figure 5: **Alternative investment strategies: horizon effect.** This figure plots the portfolio weights on the stock for three alternative investment strategies as functions of the investment horizon. One standard deviation of x_t is $\sigma_3/\sqrt{1-\rho^2} = 0.4677$. We set the risk aversion parameter $\gamma = 5$ and the initial belief $\mu_0 = 0.5$. Other parameter values are estimated using annual data as reported in Table 2. Bayesian strategy: dashdot line; Robust strategy for $\eta = 60$: dashed line; VAR strategy: solid line.

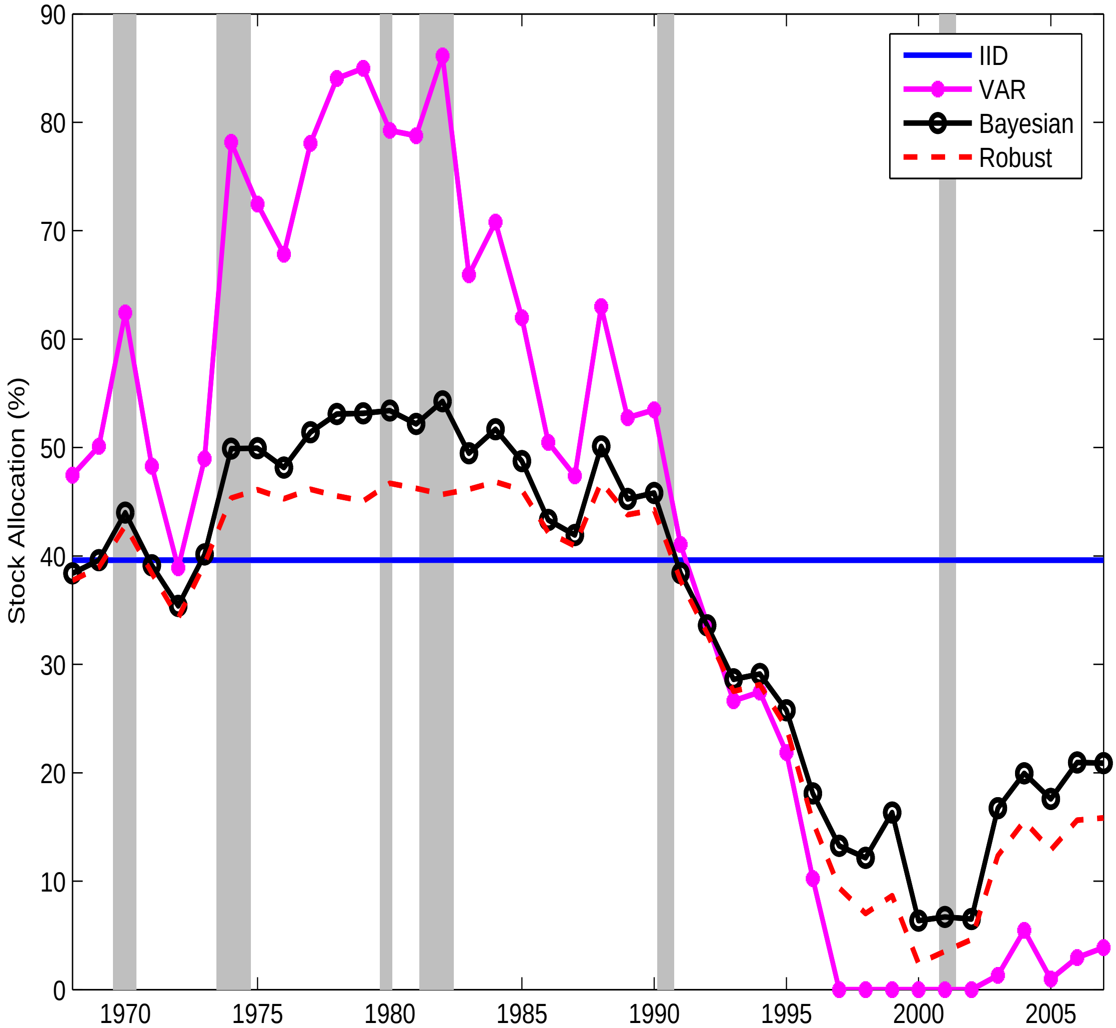

Figure 6: **Simulated stock allocations.** This figure plots the simulated portfolio weights on the stock for four alternative investment strategies using the historical data of returns and dividend price ratios from 1926-2007. The investor starts in 1968 with initial prior $\mu_0 = 0.5$. We set the risk aversion parameter $\gamma = 5$ and the investment horizon $T = 40$. Other parameter values are estimated using annual data as reported in Table 2. IID strategy: horizontal solid line; VAR strategy: solid line with star; Bayesian strategy: solid line with circle; Robust strategy for $\eta = 60$: dashed line.

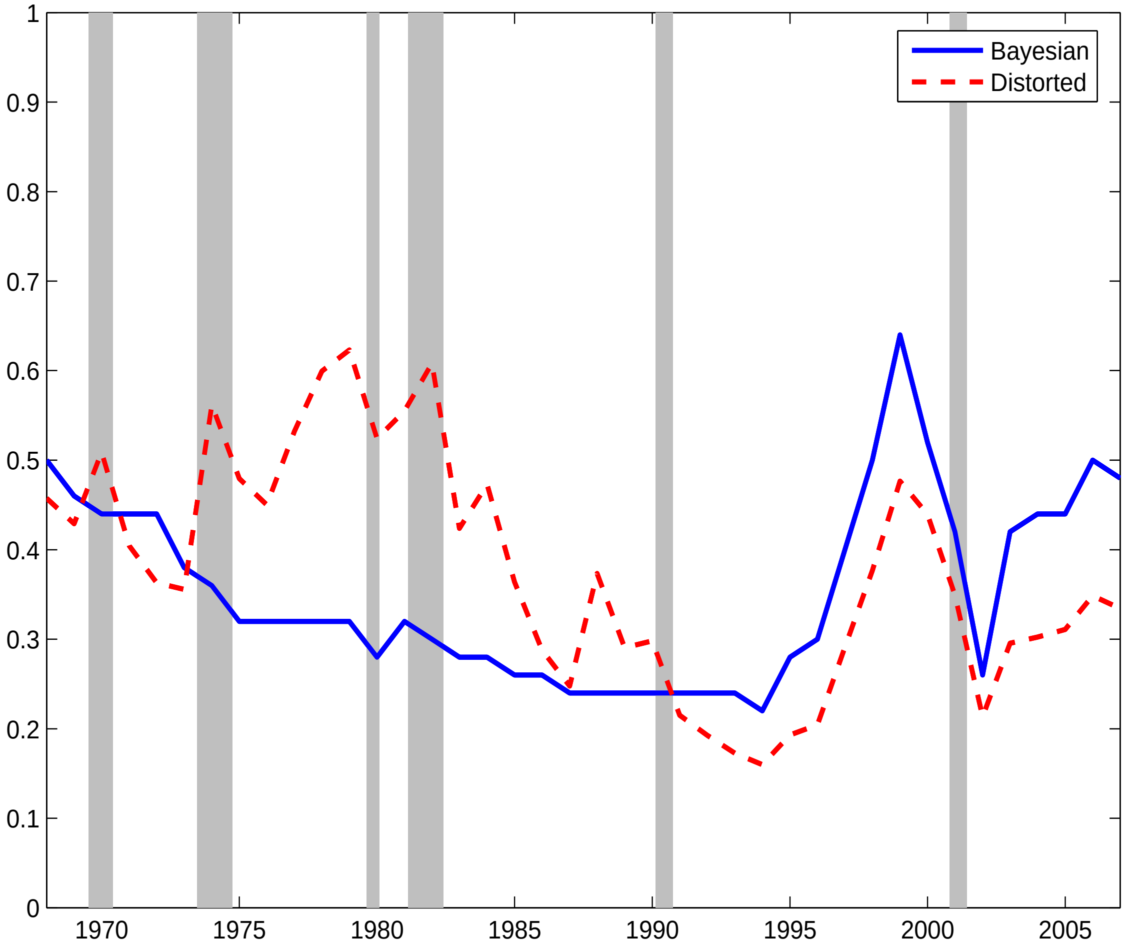

Figure 7: **Distorted beliefs about the IID model.** This figure plots the beliefs of a Bayesian investor (μ_t) and the distorted beliefs of an ambiguity-averse investor ($\hat{\mu}_t$) about the IID model using historical annual data of stock returns and dividend yields from 1968-2007. We set the initial belief $\mu_0 = 0.5$, the risk aversion parameter $\gamma = 5$, and the investment horizon $T = 40$. Other parameter values are estimated using annual data as reported in Table 2. Bayesian beliefs: solid line; distorted beliefs for $\eta = 60$: dashed line.