Ambiguity in the Cross-Section of Expected Returns: An Empirical Assessment

Julian Thimme[*]   Clemens Völkert[‡]

This version: September 9, 2012

Abstract

This paper estimates and tests the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005, 2009) based on stock market data. We introduce a novel methodology to estimate the conditional expectation which characterizes the impact of a decision maker's ambiguity attitude on asset prices. Our point estimates of the ambiguity parameter are between 25 and 40, whereas our risk aversion estimates are considerably lower. The substantial difference indicates that market participants are ambiguity averse. Furthermore, we evaluate if ambiguity aversion helps explaining the cross-section of expected returns. Compared with Epstein and Zin (1989) preferences, we find that incorporating ambiguity into the decision model improves the fit to the data while keeping relative risk aversion at more reasonable levels.

Keywords: Ambiguity aversion, asset pricing, cross-section of returns

JEL: C52, D81, E21, E44, G12

[*] Finance Center Münster, Westfälische Wilhelms-Universität Münster, Universitätsstr. 14-16, 48143 Münster, Germany. E-mail: julian.thimme@wiwi.uni-muenster.de.

[‡] Finance Center Münster, Westfälische Wilhelms-Universität Münster, Universitätsstr. 14-16, 48143 Münster, Germany. E-mail: clemens.voelkert@wiwi.uni-muenster.de.

1 Introduction

Whereas it has a long tradition to model preferences with subjective expected utility (SEU), researchers nowadays consider more sophisticated preference representations. If a decision maker (DM) has vague information about the model that determines the distribution of outcomes, uncertainty does not only appear as risk, i.e. fluctuations with a known probability distribution, but also as ambiguity about the model itself. Ambiguity may cause a loss of utility to a DM. The resulting bias in portfolio allocations, mirroring the aspiration for robust decision making, might have a perceptible impact on asset prices.

This paper investigates if ambiguity aversion is present in investors' decision patterns by looking at the cross-section of stock returns and macroeconomic variables. We estimate a set of preference parameters assuming that investors act in line with the smooth ambiguity (SA) model of preference as developed by Klibanoff et al. (2005, 2009). We compare the pricing performance of the SA model with the recursive preference model of Epstein and Zin (1989), EZ in the following, to evaluate the impact of ambiguity on asset prices. To highlight the major difficulties in estimating the preference parameters and to justify our estimation technique, we use a long-run risks (LRR) asset pricing model, similar to Bansal and Yaron (2004).

Intuitively, a DM with SA preferences considers a whole set of different economic models. For each, she calculates a certainty equivalent with respect to expected utility. Her decisions are finally based on the expected utility of the set of certainty equivalents with respect to a second utility function. This function displays the DM's ambiguity attitude and is characterized by the ambiguity parameter η. One goal of this paper is to estimate this parameter and thus gauge the ambiguity attitude of investors. An alternative approach to introduce ambiguity aversion is the multiple priors model of Gilboa and Schmeidler (1989). They assume that a DM does not consider all certainty equivalents belonging to different candidate

1

models, but only the "worst case". Hansen and Sargent (2001) suggest the relation of this approach to the robustness theory of Andersen et al. (2000) and Hansen and Sargent (2008). Ju and Miao (2012) point out that the SA framework contains these and further preference specifications as special and limiting cases.

Halevy (2007) investigates a variety of decision models using extensions of the Ellsberg (1961) experiment. Bossaerts et al. (2010) and Ahn et al. (2011) analyze the impact of ambiguity in portfolio choice experiments. In contrast to experimental studies, our results are based on historical stock market data. Epstein and Schneider (2010) review the literature on ambiguity and asset markets. They conclude that ambiguity has important implications for the pricing of financial assets. General equilibrium asset pricing applications of the SA approach include Collard et al. (2011), Ju and Miao (2012), and Miao et al. (2012). In these papers, the risk aversion parameter γ is set to a low value, while the ambiguity parameter η is calibrated to match important asset pricing moments. The assumed value varies significantly between the asset pricing applications in the literature. However, as for γ, there also has to be a reasonable range for η. The findings of Halevy (2007) are interpreted by Chen et al. (2011), who infer an ambiguity parameter between 50 and 90. Our point estimates of η are between 25 and 40, while γ is clearly lower and within the range considered plausible by Mehra and Prescott (1985). The substantial difference to the risk aversion parameter indicates that market participants are ambiguity averse.

The consumption-based asset pricing model of Lucas (1978) and Breeden (1979) has severe problems in explaining the large equity premium and the cross-sectional variation in expected returns.[1] We investigate whether accounting for ambiguity helps explaining these phenomena and compare the pricing performances of the EZ and SA models. We find that it is difficult to discriminate between these two decision models solely based on pricing errors. The SA model achieves a slightly

---

[1]See Lettau and Ludvigson (2001b), Parker and Julliard (2005), and the references therein.

better fit to the data with lower relative risk aversion.

To estimate preference parameters, we use the generalized method of moments (GMM) of Hansen (1982). Hansen and Singleton (1982) employ GMM to estimate the consumption-based capital asset pricing model, while Epstein and Zin (1991) estimate the EZ model. GMM relies on Euler equations to test the fit of candidate pricing kernels. Compared with EZ preferences, the pricing kernel and hence the Euler equation of the SA model contains an additional term, which characterizes the impact of an investor's ambiguity attitude on asset prices. An ambiguity averse agent puts more weight on economic models that yield a low expected continuation value. Estimation of this expected value, conditional on the economic model, imposes technical difficulties. Motivated by a LRR model, we show how to overcome these.

Another difficulty in estimating consumption-based asset pricing models with recursive preferences is that it requires the return on the wealth portfolio, which is not observable. Several approximations have been proposed in the literature. Epstein and Zin (1991) use the return on a broad stock market index. However, Chen et al. (2012) and Lustig et al. (2012) study the properties of the return on wealth and find that it is less volatile and only weakly correlated with the return on the stock market. Among others, Campbell (1996) and Jagannathan and Wang (1996) account for the large fraction of human wealth in total wealth. As in Zhang (2006), we use a proxy for the return on wealth based on the variable *cay* of Lettau and Ludvigson (2001a), which includes human wealth and total asset holdings.

The remainder of this paper is organized as follows. Section 2 reviews SA preferences and the pricing kernel. In Section 3, we discuss the estimation technique. In Section 4, we perform a simulation study to investigate the finite sample behavior of our estimation approach. The preference models are estimated based on post-war consumption and stock market data in Section 5. Section 6 concludes.

2 Smooth Ambiguity Preferences

In this section, we introduce SA preferences. We start with a static setting as developed by Klibanoff et al. (2005) and generalize to a dynamic setting as done by Klibanoff et al. (2009) and Ju and Miao (2012). The preference representation is a generalization of recursive preferences as developed by Kreps and Porteus (1978) and Epstein and Zin (1989). Since our goal is to provide a strong intuition for the nature of SA preferences, we set technicalities aside and refer the interested reader to the papers named above. An axiomatic foundation of SA preferences can be found in Hayashi and Miao (2011).

2.1 The static setting

Consider a state space S that contains all states of the nature. The preference function is supposed to evaluate all *acts*, i.e. random variables, that map states in S to the set $\mathcal{C}$ of (financial) consequences, which can be thought of as a subset of the real numbers. For instance, the payoff resulting from an investment decision is an act, since the realized payoff depends on the realized state of the nature. Aggregate consumption in the economy, i.e. the number of purchased goods, is also an act. The amount of utility a certain consequence provides to the DM is quantified by her utility function u. The SEU approach proposed by Savage (1954) assumes that the DM evaluates acts f by the help of the functional $\int_S u(f)d\pi \equiv \mathbb{E}_\pi[u(f)]$. To take expectations, it is necessary to quantify how probable it is that different states of nature occur, i.e. pin down *one* probability measure π on the state space S. The thought experiments of Ellsberg (1961) give reason to believe that individuals do not base their decisions on expected utility. They may have vague information about the probability measure π and rather deem a set Π of measures possible. Ambiguity aversion is characterized by a rejection of mean-preserving spreads in

expected utilities corresponding to different measures in Π.

In case of a risk averse DM, the concavity of u makes the SEU-functional $\mathbb{E}_\pi[u(f)]$ weight mean-preserving spreads in consequences of f down, due to Jensen's inequality. To model ambiguity aversion, Klibanoff et al. (2005) take advantage of the same mechanism and define the functional

$$F(f) = \int_\Pi \phi \left(\int_S u(f)d\pi \right) d\mu(\pi) \equiv \mathbb{E}_\mu[\phi(\mathbb{E}_\pi[u(f)])].$$

μ denotes a probability measure on the set Π of all probability measures on S and the function ϕ characterizes the DM's attitude towards ambiguity. Defining a utility function $v \colon \mathcal{C} \to \mathbb{R}$ via $v = \phi \circ u$, we may interpret $F(f)$ as $\mathbb{E}_\mu[v(CE(\pi, f)]$, where we define the conditional certainty equivalent CE corresponding to the measure $\pi \in \Pi$ as $CE(\pi, f) = u^{-1}(\mathbb{E}_\pi[u(f)])$. Hence, a DM who acts in line with SA preferences maximizes expected utility of the conditional certainty equivalents that belong to different probability measures. The ambiguity attitude of the DM is characterized by the curvature of the function ϕ. A convex ϕ displays ambiguity loving decision behavior, while a concave ϕ leads to an ambiguity averse decision pattern. This is equivalent to v being a concave transformation of u. If ϕ is linear, SA preferences reduce to subjective expected utility.

As pointed out by Ju and Miao (2012), the interpretation of the set Π links the SA approach to the robust control literature and provides valuable intuition. Each probability measure $\pi \in \Pi$ corresponds to an economic model. The measure μ determines how probable the DM believes the different models are. The less concentrated the mass under μ, the higher is the *model uncertainty*, a term that we use interchangeably with the term *ambiguity*.

5

2.2 The dynamic setting

Observations of realized outcomes today might cause the DM to change her perception of the likelihood of future economic models. While the set Π of all possible measures does not vary, the distribution μ on Π hence changes over time, so we add a time index to μ.[2] We also add a time index to the measure π to point out that π_{t+1} corresponds to the economic model at hand at time $t+1$.

Let $C = (C_t)_{t \in \mathbb{N}}$ be the DM's consumption plan, i.e. a series of acts, where each act is a time t-measurable random variable. As in Epstein and Zin (1989), we assume that the DM's time t value function is given recursively by

$$V_t(C) = \mathcal{W}(C_t, \mathcal{R}_t(V_{t+1}(C))),$$

where $\mathcal{W}$ denotes the time aggregator and $\mathcal{R}_t$ the uncertainty aggregator. The latter can be interpreted as an unconditional certainty equivalent and is specified as

$$\mathcal{R}_t(x) = v^{-1} \left(\mathbb{E}_{\mu_t} \left[\phi \left(\mathbb{E}_{\pi_{t+1}} [u(x)] \right) \right] \right).$$

More precisely, we assume that $\mathcal{W}$ is a constant elasticity of substitution time aggregator and that u and v are of the power utility type

$$
\begin{aligned}
\mathcal{W}(x, y) &= \left[\left(1 - e^{-\delta} \right) x^{1-\rho} + e^{-\delta} y^{1-\rho} \right]^{\frac{1}{1-\rho}}, \\
u(x) &= \frac{x^{1-\gamma}}{1 - \gamma}, \quad \gamma > 0, \neq 1, \\
v(x) &= \frac{x^{1-\eta}}{1 - \eta}, \quad \eta > 0, \neq 1.
\end{aligned}
$$

---

[2]Other authors, e.g. Ju and Miao (2012), characterize Π as the set of those measures the DM thinks possible. From their point of view, the size of Π determines the amount of ambiguity in an economy and therefore has to be time-varying. We obtain this set by excluding all μ_t-zero sets from our Π, which is the set of *all* probability measures on S and therefore constant over time.

The parameter vector $\Theta = (\rho, \delta, \gamma, \eta)$ describes the DM's preferences and is therefore the core of this paper. ρ denotes the reciprocal of the DM's elasticity of intertemporal substitution (EIS) and δ the DM's subjective time discount rate. The remaining parameters describe the DM's attitudes towards risk (γ) and ambiguity (η). She is ambiguity averse if $\eta > \gamma$. Summing up, at time t the DM evaluates consumption plans C according to

$$V_t(C) = \left[\left(1 - e^{-\delta}\right) C_t^{1-\rho} + e^{-\delta} \left\{ \mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_{\pi_{t+1}} \left[V_{t+1}^{1-\gamma}(C) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right\}^{\frac{1-\rho}{1-\eta}} \right]^{\frac{1}{1-\rho}}. \tag{1}$$

At time t, the DM observes the current economic model π_t. However, the future economic model π_{t+1} is ambiguous, meaning that $\mathbb{E}_{\pi_{t+1}}[V_{t+1}^{1-\gamma}(C)]$ is a random variable on Π. Equation (1) nests the value functions of EZ preferences ($\eta = \gamma$) and constant relative risk aversion (CRRA) utility ($\rho = \eta = \gamma$).

2.3 The pricing kernel

The pricing kernel ξ links preferences to asset returns via the relation

$$\mathbb{E}_t \left[\xi_{t,t+1} R_{t,t+1}^i - 1 \right] = 0,$$

where $\mathbb{E}_t$ is an abbreviation for $\mathbb{E}_{\mu_t} \mathbb{E}_{\pi_{t+1}}$ and $R_{t,t+1}^i$ denotes the gross return on money invested at time t for one period in an arbitrary asset i. Alternatively, the Euler equation for asset i can be expressed in terms of excess returns

$$\mathbb{E}_t \left[\xi_{t,t+1} \left(R_{t,t+1}^i - R_{t,t+1}^f \right) \right] = 0,$$

where R^f denotes the return on the risk-free asset. In complete markets, $(\xi_{t,t+1})_{t \in \mathbb{N}}$ is a unique series of random variables. It can be expressed in terms of continuation values with the help of the value function described in Section 2.2. Following Duffie

and Skiadas (1994) and Hansen et al. (2007), it satisfies

$$\xi_{t,t+1} \;=\; e^{-\delta} \left(\frac{C_{t+1}}{C_t}\right)^{-\rho} \left(\frac{V_{t+1}(C)}{\mathcal{R}_t(V_{t+1}(C))}\right)^{\rho-\gamma} \left(\frac{(\mathbb{E}_{\pi_{t+1}}[V_{t+1}^{1-\gamma}(C)])^{\frac{1}{1-\gamma}}}{\mathcal{R}_t(V_{t+1}(C))}\right)^{\gamma-\eta}, \quad (2)$$

as reported in Hayashi and Miao (2011), Proposition 8. The first three terms are the EZ pricing kernel, which collapses to the CRRA pricing kernel for $\gamma = \rho$. The last term displays the impact of the DM's ambiguity attitude on asset prices. Its numerator is the conditional certainty equivalent $CE(V_{t+1}(C))$ of the continuation value as defined in Section 2.1. Hence, the DM considers the conditional certainty equivalent corresponding to a certain economic model π_{t+1} relative to the unconditional certainty equivalent. Depending on her ambiguity attitude, she puts more (if ambiguity averse, i.e. $\gamma < \eta$) or less (if ambiguity loving, i.e. $\gamma > \eta$) weight on economic models that yield a low expected utility.

The continuation value is unobservable and in applications it is usually more convenient to work with the pricing kernel in terms of the return on wealth. We define $\theta_1 := \frac{1-\gamma}{1-\rho}$ and $\theta_2 := \frac{1-\eta}{1-\gamma}$. The pricing kernel then is

$$\xi_{t,t+1} = e^{-\delta\theta_1} \left(\frac{C_{t+1}}{C_t}\right)^{-\rho\theta_1} (R_{t,t+1}^w)^{\theta_1-1} \left(\mathbb{E}_{\pi_{t+1}} \left[e^{-\delta\theta_1} \left(\frac{C_{t+1}}{C_t}\right)^{-\rho\theta_1} (R_{t,t+1}^w)^{\theta_1}\right]\right)^{\theta_2-1}, \quad (3)$$

where R^w denotes the return on the wealth portfolio, i.e. the claim on aggregate consumption. The parameter θ_2 expresses the concavity of ϕ and therefore the ambiguity attitude of the DM. Hence, the bias through the additional last term in the pricing kernel (compared to the EZ pricing kernel) causes the impact of the DM's ambiguity attitude on asset prices.

3   Estimation Technique

In this section, we introduce the econometric methodology to infer attitudes towards ambiguity from financial market data. We use GMM to estimate the preference parameters. From Equation (3) it is clear that two central components of the pricing kernel are the return on the wealth portfolio, which cannot be observed at the market, and the expected value, conditional on the economic model, that distinguishes SA from EZ preferences. We discuss a return proxy in Section 3.2 and provide an estimation technique for the conditional expectation in Section 3.3.

3.1   GMM estimation

Euler equations link asset returns to consumption growth and the return on the wealth portfolio. The imposed population moment restrictions can be employed to test the fit of candidate pricing kernels.

The returns used in the estimation are those of a number of test assets. We use returns on the 3-month Treasury bill, a broad stock market index, and 30 additional equity portfolios. To weight the moment conditions, we employ the identity matrix. The test assets are selected based on economically interesting characteristics and this choice of the weighting matrix guarantees that the candidate pricing kernels are evaluated on how they price these assets, rather than linear combinations of them. The moment condition for the 3-month Treasury bill forces the mean of the pricing kernel to equal the inverse of the gross return on the risk-free asset. The other moment conditions require the model to fit the equity premium and the cross-section of average returns. Minimizing the sum of squared pricing errors makes the results comparable to asset pricing tests using OLS cross-sectional regressions. In addition, the identity matrix is suitable for model comparison, as it is invariant across all models tested. According to Altonji and Segal (1996), first stage GMM

estimates are more robust in finite samples. Cochrane (2005, Ch. 11) explains several additional advantages of using a prespecified weighting matrix.

There are more moment conditions than unknown parameters, i.e. the system is overidentified. The null hypothesis that all moment conditions are zero can be tested using Hansen's J-test. If we acknowledge that all models are misspecified, hypotheses tests of the null of correct model specification against the alternative of incorrect specification are of limited value. Following the idea that we are looking for the least misspecified model, we compare root mean squared errors (RMSE) and Hansen and Jagannathan (1997) distances (HJD) of different preference specifications and parameter vectors. We test if these performance measures are zero using the methodology proposed by Jagannathan and Wang (1996) and Parker and Julliard (2005). To evaluate parameter restrictions for time-separability ($\gamma = \rho$) and ambiguity neutrality ($\gamma = \eta$), we employ Wald tests. Details on the estimation procedure and on testing hypotheses are provided in Appendix A.

Ferson and Foerster (1994), Hansen et al. (1996), Smith (1999), and Ahn and Gadarowski (2004) point out that commonly employed specification tests reject too often in finite samples. Thus, relying solely on these tests to evaluate the goodness of fit of candidate asset pricing models is problematic. Lewellen et al. (2010) show that focusing to closely on high cross-sectional R^2s and small pricing errors can be misleading. We follow their advice to expand the set of test portfolios beyond size and book-to-market sorted portfolios and evaluate if the decision models produce plausible preference parameter estimates. Allowing all parameters to be estimated freely focuses solely on model fit. Restricting certain preference parameters to economically reasonable values, balances the objective between minimizing pricing errors and the plausibility of the parameter estimates.

For a large part of the analysis, we follow Bansal et al. (2007) and fix the EIS at economically reasonable values. There is considerable debate about the correct

value of the EIS. Hall (1988), Campbell and Mankiw (1989), and Yogo (2004) find an EIS close to zero, while Vissing-Jorgensen and Attanasio (2003), Bansal and Yaron (2004), Guvenen (2006), and Chen et al. (2012) argue for a higher value. Hansen et al. (2008) and Malloy et al. (2009) set the EIS to one. This choice simplifies the analysis considerably. However, it implies that the wealth-consumption ratio is constant. Lettau and Ludvigson (2001a) and Lustig et al. (2012) show that this contradicts empirical evidence. In the LRR model of Bansal and Yaron (2004), a drop in volatility and a rise in expected consumption growth increase the wealth-consumption ratio if the EIS is greater than one. Bansal et al. (2005) support the negative relation between volatility and asset prices and Lustig et al. (2012) show that the LRR model produces a wealth-consumption ratio that fits the data.

As suggested by Constantinides and Ghosh (2011), we report results for several values of the EIS. Prefixing the EIS is beneficial for several reasons. First, it is very difficult to estimate the EIS reliably. As our main object of interest is the DM's attitude towards ambiguity and not the magnitude of the EIS, setting it to economically reasonable values simplifies the estimation of the ambiguity parameter. Second, it facilitates the comparison of parameter estimates of the EZ and SA models, i.e. the effect that differences in the estimated EIS cause large changes in the estimated values of risk and ambiguity aversion are avoided. Furthermore, fixing the EIS at reasonable levels may provide valuable guidance on the magnitude of risk and ambiguity aversion for researchers in calibrating the SA model.

3.2   Return on wealth

Testing candidate pricing kernels corresponding to EZ or SA preferences presumes that either the continuation value of the future consumption plan in Equation (2) or the return on the wealth portfolio in Equation (3) is observable. The wealth portfolio is an asset that pays aggregate consumption as dividends. Although ag-

gregate consumption is observable, neither the return on aggregate wealth nor the continuation value can be observed at the market. This causes severe problems for estimating consumption-based asset pricing models, as pointed out by Ludvigson (2012).[3] Approximating the continuation value is discussed in Hansen et al. (2008), Ju and Miao (2012), and Chen et al. (2012). The latter estimate the functional form of the continuation value ratio, defined as the continuation value divided by time t consumption, using a profile sieve minimum distance procedure.

Approximating the return on wealth with a suitable function of observable variables is another alternative. Epstein and Zin (1991) use the return on a broad stock market index as a proxy for the return on aggregate wealth. Among others, Stock and Wright (2000) and Yogo (2006) follow this approach. However, a stock market index is only a good proxy for the return on aggregate wealth if human capital and other non-tradable assets are minor components of aggregate wealth. Critique of this approach goes back to Roll (1977). Lustig et al. (2012) show that human capital makes up the largest fraction of aggregate wealth. Campbell (1996) and Jagannathan and Wang (1996) include human capital. However, other components of wealth, such as total household asset holdings, should also be accounted for. We discuss an approach that incorporates all kinds of wealth by using the *cay* variable.

Lettau and Ludvigson (2001a) define *cay* as

$$cay_t := c_t - \omega a_t - (1 - \omega)y_t,$$

where c denotes log consumption, a log asset holdings, and y log aggregate labor income.[4] The variable ω is the relative share of asset holdings in total wealth, which is assumed to be constant over time. Lettau and Ludvigson (2001a) assume that

---

[3]The pricing kernel can be expressed in terms of observables by restricting the EIS to one or by imposing tight restrictions on either the consumption dynamics or the joint distribution of consumption and asset returns.

[4]See Appendix C for the precise specification of the variables used.

asset holdings and human capital sum up to total wealth and that human wealth is approximately proportional to labor income. The variable cay is a proxy for innovations in the time t log consumption-wealth ratio. As in Zhang (2006), we use cay to construct the return on the wealth portfolio.[5]

Let W_t denote aggregate wealth at time t. Using the budget constraint $W_{t+1} = (W_t - C_t)R^w_{t,t+1}$, the return on wealth is given by

$$R^w_{t,t+1} = \frac{C_{t+1}}{C_t} \cdot \frac{C_t/W_t}{C_{t+1}/W_{t+1}} \cdot \frac{1}{1 - \frac{C_t}{W_t}}.$$

Assume that $\frac{C_t}{W_t} = \kappa \cdot \exp(cay_t)$, i.e. the consumption-wealth ratio fluctuates around its steady state value κ. A proxy for the return on wealth is

$$R^{cay}_{t,t+1} := \frac{C_{t+1}}{C_t} \frac{\exp(cay_t - cay_{t+1})}{1 - \kappa \cdot \exp(cay_t)}.$$

The constant κ is of minor relevance, since it is the timing of innovations to the consumption-wealth ratio rather than its level that is important for the estimation.

Table 1 contains summary statistics of two proxies for the return on wealth. R^{CRSP} denotes the gross return on the CRSP value weighted index. The mean of $r^{cay} = \log R^{cay}$ is 1.48% per quarter, which is close to the mean of $r^{CRSP} = \log R^{CRSP}$ (1.52%). However, its standard deviation is only 0.81% per quarter and therefore less than one tenth of the standard deviation of the return on the CRSP index (8.57%). r^{cay} has similar statistical properties as the return on wealth in Chen et al. (2012). They find that the return on aggregate wealth is less volatile compared with the return on the CRSP stock market index and the correlation between the two is rather low. In our sample, the correlation between r^{cay} and r^{CRSP} is about 0.5.

---

[5]Lettau and Ludvigson (2001a) propose that c, a, and y are cointegrated and estimate the coefficients ω and $1 - \omega$ using OLS. Since the estimates do not perfectly sum up to 1, we proceed as Zhang (2006) and divide the estimate of ω by the sum of both estimates.

3.3 Estimation of the conditional expectation

Using Euler equations to estimate the SA model requires an empirical estimate of the conditional expectation

$$\mathbb{E}_{\pi_{t+1}}\left[e^{-\delta\theta_1} \left(\frac{C_{t+1}}{C_t} \right)^{-\rho\theta_1} (R^w_{t,t+1})^{\theta_1} \right],$$

which distinguishes SA preferences from EZ utility. To estimate the conditional expected value, we employ local relative least squares estimation on the realizations

$$Y_{t+1} := \left(e^{-\delta} \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R^w_{t,t+1} \right)^{\theta_1}, \tag{4}$$

i.e. we minimize the relative error

$$\sum_{t=0}^{T-1} \left\{ \frac{Y_{t+1} - M(X_{t+1})}{Y_{t+1}} \right\}^2.$$

We assume that the conditional expectation of Y_{t+1} is a function M of a standardized vector X_{t+1} of time $t+1$ regressors. For the approach to be valid, the economic model needs to be explicable by these regressors. More precisely, there has to be a bijective relation between the set of economic models Π and the image of X, i.e. all possible realizations of the regressor variables. We discuss adequate choices of the regressors in Section 4.2.

The conditional expectation to be estimated is multiplied by other terms in the Euler equations. This makes minimizing the relative error a sensible choice. Furthermore, minimizing relative least squares instead of ordinary least squares is motivated by commonly utilized affine asset pricing models. In these models, the log of Equation (4) is affine in the innovations to log consumption growth and the state variables. Following Park and Stefanski (1998), we exploit the approximation

$$\log(x) \approx x - 1$$

$$\sum_{t=0}^{T-1} \left\{ \frac{Y_{t+1} - M(X_{t+1})}{Y_{t+1}} \right\}^2 \approx \sum_{t=0}^{T-1} \left\{ \log(Y_{t+1}) - \log(M(X_{t+1})) \right\}^2.$$

Defining $m := \log \circ M$, we run the following regression

$$\log(Y_{t+1}) = m(X_{t+1}) + \Sigma(X_{t+1}) \varepsilon_{t+1}, \quad t \in \{0, \dots, T-1\},$$

where Σ is the conditional volatility and ε denotes an i.i.d. zero mean, unit variance disturbance term. We allow for nonlinearities in m and approximate the conditional expectation locally by a linear function.[6] Fan (1992) and Ruppert and Wand (1994) show that the local linear estimator has several advantages compared to other non-parametric estimators. For each $i \in \{0, \dots, T-1\}$, an estimate of the conditional expectation at the data point X_{i+1} is

$$\hat{m}(X_{i+1}) = \hat{\alpha}_{i+1} + \hat{\beta}_{i+1} \cdot X_{i+1},$$

where $([\hat{\alpha}_{i+1}, \hat{\beta}_{i+1}])_{i \in \{0,\dots,T-1\}}$ is given by

$$[\hat{\alpha}_{i+1}, \hat{\beta}_{i+1}] = \operatorname*{argmin}_{[\alpha,\beta]} \sum_{t=0}^{T-1} \left\{ \log(Y_{t+1}) - (\alpha + \beta \cdot X_{t+1}) \right\}^2 w(X_{i+1}, X_{t+1}).$$

The weighting function w assigns more weight to observations close to the current data point and less weight to observations farther away. If l denotes the number of explanatory variables, i.e. the length of the vector X_{t+1} for $t \in \{0, \dots, T-1\}$, the weighting function is defined as

$$w(X_1, X_2) := \prod_{j=1}^{l} K \left(\frac{X_{1,j} - X_{2,j}}{h_j} \right).$$

---

[6]Nagel and Singleton (2011) use a similar approach to estimate conditional moments.

The weighting depends on the specification of the kernel function K, which assigns local weights to the linear estimator. The vector of bandwidths $h = (h_1, \ldots, h_l)$ controls the neighborhood of the current point. We use the Epanechnikov kernel $K(u) = \frac{3}{4}(1 - u^2)1_{(|u| \leq 1)}$ and employ an individual bandwidth for each regressor. A large bandwidth h_i results in a smooth estimate that might neglect important features of the data contained in the i-th regressor, while for a small bandwidth the estimate follows the data very closely. The optimal vector of bandwidths h is chosen by minimizing the cross-validation criterion

$$CV(h) = \frac{1}{T} \sum_{i=0}^{T-1} \left\{ \log(Y_{i+1}) - (\hat{\alpha}_{i+1}^- + \hat{\beta}_{i+1}^- X_{i+1}) \right\}^2,$$

where $[\hat{\alpha}_{i+1}^-, \hat{\beta}_{i+1}^-]$ denotes the estimate computed excluding the $(i+1)$-th data point. The optimal bandwidths depend on the explanatory variables and the EIS.

4   Long-Run Risks Model

In this section, we estimate the EZ and SA models based on simulated consumption and return data to investigate the performance of our estimation technique in finite samples. As in Bansal and Yaron (2004), the laws of motion of consumption and dividend growth are driven by two long-run risk factors.

4.1   Data generating process

In Bansal and Yaron (2004), the distribution of consumption and dividend growth, hence the time $t + 1$ economic model, depends on the realizations of two state variables that characterize the economic circumstances at time $t + 1$. We use a similar endowment process and assume that the representative investor is ambiguous about the state variables that drive consumption and dividend growth in the long-run.

Short-run consumption and dividend uncertainty are treated as risk and evaluated with the risk aversion coefficient γ. The long-run growth factor and the conditional volatility of consumption growth are perceived as ambiguous. The parameter η determines the investor's attitude towards these sources of uncertainty.

The laws of motion of log consumption growth $\Delta c_{t+1} = \log C_{t+1} - \log C_t$, log dividend growth $\Delta d_{t+1} = \log D_{t+1} - \log D_t$, and of the state variables are

$$
\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_{t+1} + \sigma_{t+1} w^c_{t+1}, \\
\Delta d_{t+1} &= \mu_d + \lambda x_{t+1} + \sigma_{t+1} \phi_{d,\sigma} \left(\rho_{cd} w^c_{t+1} + \sqrt{1 - \rho^2_{cd}} w^d_{t+1} \right), \\
x_{t+1} &= \varphi_x x_t + \phi_x \sigma_{t+1} w^x_{t+1}, \\
\sigma^2_{t+1} &= \sigma^2 + \varphi_\sigma(\sigma^2_t - \sigma^2) + \phi_\sigma w^\sigma_{t+1},
\end{aligned}
\tag{5}
$$

where $w^c_{t+1}, w^d_{t+1}, w^x_{t+1}, w^\sigma_{t+1} \sim i.i.d.\ \mathcal{N}(0,1)$. Consumption and dividend growth contain a persistent long-run growth component x_{t+1} and the conditional volatilities are driven by a time-varying uncertainty factor σ^2_{t+1}.[7]

4.2 Conditional expectation

Using the dynamics in Equation (5) and the coefficients of the wealth-consumption ratio in Equation (8) of Appendix B, the conditional expectation of Y_{t+1} equals

$$
\begin{aligned}
\mathbb{E}_{\pi_{t+1}}[Y_{t+1}] &= \mathbb{E}_\pi \left[e^{\theta_1(q - B_x x_t + (k_1 B_x + (1-\rho))x_{t+1} - B_\sigma \sigma^2_t + k_1 B_\sigma \sigma^2_{t+1} + (1-\rho)\sigma_{t+1} w^c_{t+1})} \right] \\
&= e^{\theta_1(q - B_x x_t + (k_1 B_x + (1-\rho))x^\pi_{t+1} - B_\sigma \sigma^2_t + k_1 B_\sigma (\sigma^\pi_{t+1})^2)} \; e^{\frac{1}{2}(1-\gamma)^2(\sigma^\pi_{t+1})^2},
\end{aligned}
\tag{6}
$$

with $q = -\delta + (1-\rho)\mu_c + k_0 + (k_1 - 1)(A - B_\sigma \sigma^2)$. In this equation, the measure π_{t+1}, i.e. the economic model, is fully characterized by the realizations x^π_{t+1} and

---

[7]The solution of the model is sketched in Appendix B. Note that in Bansal and Yaron (2004), consumption growth depends on x_t and σ^2_t and therefore, at time t, the investor knows that $\Delta c_{t+1} \sim \mathcal{N}(\mu_c + x_t, \sigma^2_t)$. Consequently, she cannot be ambiguous about the distribution. In contrast, we assume that consumption growth depends on x_{t+1} and σ^2_{t+1}.

σ_{t+1}^{π}. Hence, taking expectations conditional on the measure π means conditional on the state variables. In general, imposing affine dynamics on the endowment process in Equation (5) implies that the conditional expectation is an exponential affine function of the underlying state variables.

In order to use the methodology outlined in Section 3.3, we need to identify a set of predictor variables whose realizations map one-to-one and onto the set Π of economic models. For the endowment process in Equation (5), the vector of state variables and their first lags has this property

$$\log(Y_{t+1}) = \alpha + \beta_1 x_{t+1} + \beta_2 x_t + \beta_3 (\sigma_{t+1}^2 - \sigma^2) + \beta_4 (\sigma_t^2 - \sigma^2) + (1 - \gamma)\sigma_{t+1}\varepsilon_{t+1}.$$

In reality, we are not able to observe the underlying state vector. Thus, it is necessary to identify which observable variables describe the dynamics of the economy. The risk-free rate and the log price-dividend ratio are observable at the market, show a clear business cycle pattern, and have a long tradition as predictors of stock and bond returns.[8] Furthermore, in standard affine asset pricing models these variables are approximately affine in the state vector.[9] Among others, Constantinides and Ghosh (2011) and Bansal et al. (2012b) exploit this relation to estimate LRR models. In the models they consider, these two variables span the state space. This also holds if the representative investor is ambiguous about the distribution of consumption and dividend growth. They invert the expressions for the log price-dividend ratio and the risk-free rate to express the state variables in terms of observables. Because of their economic relevance and motivated by the relation in affine asset pricing models, we use these two quantities and their first lags as predictor variables.

---

[8] Cochrane (2005, Ch. 20) provides a detailed review of the literature.
[9] See Drechsler and Yaron (2011), Lustig et al. (2012), and the references therein.

4.3 Finite sample evidence

Bansal and Yaron (2004) and Bansal et al. (2012a) calibrate their models assuming a monthly decision interval. However, we use quarterly data over the post-war sub-period to estimate the preference parameters in Section 5. In order to be consistent with the empirical sampling frequency, we use a quarterly calibration and convert the model parameters from a monthly into a quarterly frequency. Table 2 contains the parameters used in our simulation study. The risk aversion coefficient γ is assumed to be 10 for EZ preferences. For the SA model, we choose a lower value of the risk aversion coefficient ($\gamma = 5$) and set the ambiguity parameter η to 20. The subjective time discount rate is fixed at 0.0033. The models are estimated based on simulated consumption and return data with the EIS restricted to 1.5. We simulate the models 1,000 times at a quarterly frequency with a sample size approximately equivalent to the actual data (60 years). To investigate the finite sample behavior, we also estimate the parameters based on longer samples. We start the simulations at the unconditional means of the state variables and discard the first 10 years of each simulated path. In addition to the return on the dividend claim, we generate 30 additional equity test assets to represent the cross-section of returns.[10]

The results of the estimations are summarized in Table 3. First, we discuss the results assuming that the investor has EZ preferences. Panel 1 of Table 3 shows that the median estimated parameter vector $\hat{\Theta}^{\text{EZ}} = (0.667, 0.003, 10.200)$ is very close to the assumed preference parameters. The subjective time discount rate is estimated very precisely, while the (median) standard error is larger for the risk aversion coefficient. For the hypotheses tests, we report the median test statistics and p-values (in parentheses). The model is not rejected by any of the three specification

---

[10]The log dividend growth of asset i follows from

$$\Delta d^i_{t+1} = \mu_{d^i} + \lambda^i x_{t+1} + \sigma_{t+1}\phi_{d^i,\sigma}\left(\rho^i_{cd}w^c_{t+1} + \sqrt{1-(\rho^i_{cd})^2}w^{d^i}_{t+1}\right).$$

tests and the Wald test rejects the null hypothesis of time-separability.

As commonly done in empirical tests of the EZ model, we substitute the return on the consumption claim by the return on the dividend claim. In the LRR model, the return on the consumption claim has a mean of 0.73% and a standard deviation of 1.52% per quarter. In contrast, the mean and volatility of the return on the dividend claim are 1.15% and 8.94%. The standard deviation is about 6 times larger. The correlation between the two return series is 0.53.[11] Panel 2 shows that using the return on the dividend claim as a proxy for the return on wealth biases the parameter estimates dramatically. The median estimated parameter vector is $\hat{\Theta}^{EZ} = (0.667, 0.000, 1.752)$. The subjective time discount rate is close to zero and the estimated risk aversion coefficient is far too low, which is mainly due to the high volatility of the return on the dividend claim. All specification tests reject the model. This indicates that even in our model economy, substituting the return on a stock market index for the return on wealth is inappropriate.

Panels 3 and 4 of Table 3 present the results for the SA model. While Panel 3 shows the estimated preference parameters and hypotheses tests with the analytic formula for the conditional expectation in Equation (6), Panel 4 displays results for the approximation technique outlined in Section 3.3. The estimated parameters are close to the true values. Furthermore, the approximated conditional expectation delivers almost identical results compared to the true conditional expectation. The standard errors of the estimated parameters are quite large, indicating that it is rather difficult to estimate the risk aversion and ambiguity parameters jointly in small samples. The specification tests do not reject the model. Due to the lower value of γ, as well as its large standard error, the Wald test does not reject time-separability ($\rho = \gamma$). The large standard error of the risk aversion coefficient also implies that ambiguity neutrality ($\gamma = \eta$) is difficult to reject. Even though the

---

[11]The numbers are obtained from a single simulation run of 100,000 years.

20

model is true, the median p-value of the Wald test is above 10%.

In addition, we investigate the bias if the parameters are estimated on the basis of EZ preferences but the data was generated by the SA model. Our main conclusion from Panel 5 is that risk aversion is estimated with an upward bias if ambiguity aversion is present. The p-values of the specification test based on the HJD and the J-test are slightly below 5%. The RMSE is close to the one in Panel 3. Even if ambiguity aversion is present, it is rather difficult to discriminate between the models solely based on their pricing errors in small samples. Considering the plausibility of the estimated preference parameters is also important.

Table 4 reports results for larger sample sizes N (100, 200, 500, and 1000 years). As expected, the estimated parameters are close to the true values and their standard errors decrease in N. Concerning the hypotheses tests for the SA model, the Wald test for ambiguity neutrality is rejected for a sample size of 100 years and above, while the Wald test rejects time-separability for 500 and 1000 years. Both the EZ and SA models are not rejected by the specification tests. Concerning the finite sample properties of these tests, we are able to draw several conclusions. Table 5 shows the rejection rates of the specification tests, i.e. the proportion of estimations where the empirical p-value is smaller than the asymptotic size. It is well-known that the J-test has poor finite sample properties. We observe that in our simulation study the rejection rates are far too large for samples smaller than 500 years. This confirms that the test rejects too often for sample sizes typically used in empirical tests of consumption-based asset pricing models. Similar to Ahn and Gadarowski (2004), we find that the specification test based on the second moment matrix also performs poorly in small samples. In contrast to this, the test based on the RMSE behaves superior in finite samples. The rejection rates for a sample size of 60 years differ only slightly from those with 1000 years and are close to the asymptotic values.

5 Empirical Evidence

In this section, we estimate the preference parameters based on consumption and stock market data. Furthermore, we analyze the evolution of the estimated pricing kernels and investigate the in-sample and out-of-sample pricing performances of the alternative preference models. The sample period is from the first quarter of 1952 to the third quarter of 2011. The set of test assets includes the 3-month Treasury bill, the CRSP value weighted stock market index, 10 portfolios formed on size, 10 book-to-market value sorted portfolios, and 10 industry portfolios. R^{cay} is used as proxy for the return on wealth. The data is described in Appendix C. Table 1 contains descriptive statistics of the variables used in the estimation.

5.1 Parameter estimates

The estimated parameters and their standard errors are reported in Table 6. First, we discuss the results if the preference parameters are estimated under the restriction $\gamma = \eta$, i.e. investors are assumed to be ambiguity neutral. All results are reported for three values of the EIS, 1.5, 2, and 2.5, which are in line with the ones typically used in asset pricing studies with recursive preferences. Nevertheless, it is informative which value of the EIS minimizes the objective function. Our point estimate of the EIS is 1.78, which is in line with these values. However, the objective function is very flat in ρ and the confidence interval contains negative values of the EIS. The problem of getting precise estimates of the EIS has also been reported by Bansal et al. (2007) and Constantinides and Ghosh (2011), among others.[12]

Restricting the values of ρ to 2.5^{-1}, 2^{-1}, and 1.5^{-1}, the estimated parameters $\hat{\Theta}^{\text{EZ}}$ are $(0.400, 0.010, 36.292)$, $(0.500, 0.011, 31.075)$, and $(0.667, 0.011, 21.843)$. The subjective time discount rate is approximately 4% per annum and estimated with

---

[12]Also see the discussion in Section 3.1.

great precision. The point estimates of relative risk aversion are clearly above the values considered plausible by Mehra and Prescott (1985). Malloy et al. (2009) argue that the EIS has little impact on the risk aversion estimate if the estimation is solely based on the cross-sectional variation in returns. In contrast to their study, we force the models to also match the equity premium. We observe that higher values of the EIS lead to larger estimates of the risk aversion parameter. For the intuition of this result, consider the standard LRR model of Bansal and Yaron (2004). In this model, the market prices of the long-run risk factors are proportional to $(\theta_1 - 1)$. When estimating the preference parameters, returns are given exogenously. Consequently, the representative investor's preferences just influence the pricing kernel and thus the market prices of risk. To fit the equity premium, a lower value of ρ requires a higher value of γ, as long as $\rho < 1$.

The Wald test rejects the null hypothesis of time-separability, which indicates that breaking the link between the EIS and risk aversion imposed by CRRA preferences is beneficial. EZ preferences price the cross-section of expected returns rather well and a RMSE of zero cannot be rejected. The other two specification tests, the J-test and the test based on the HJD, reject the model. However, we have seen in Section 4.3 that the finite sample properties of these two tests are poor.

For the SA model, the parameter vectors $\hat{\Theta}^{\text{SA}}$ are $(0.400, 0.012, 4.283, 40.545)$, $(0.500, 0.012, 2.295, 35.091)$, and $(0.667, 0.011, 0.819, 24.817)$. While the subjective time discount rate is similar to the one of the EZ model, we obtain lower values of risk aversion. The point estimates of the ambiguity parameter η are considerably larger than the risk aversion estimates. The ambiguity parameter increases in the value of the EIS. In the LRR model with ambiguity about the state variables, Equation (9) of Appendix B reveals that the market prices of the long-run risk factors are proportional to $(\theta_1 \theta_2 - 1)$. To match the equity premium in the data, an increase in ρ is compensated by a lower value of η, as long as $\rho < 1$.

We find the relation $\hat{\gamma}^{SA} < \hat{\gamma}^{EZ} < \hat{\eta}$ throughout all investigated cases. As predicted by our simulation study, risk aversion is estimated with an upward bias if ambiguity aversion is present and $\gamma = \eta$ is imposed in the estimation. The result that the ambiguity parameter is estimated above the risk aversion coefficient indicates the presence of ambiguity aversion in the cross-section of expected returns. The null hypothesis of ambiguity neutrality is not rejected by the Wald test. However, this has to be put into perspective to the finite sample evidence in Section 4.3, where the Wald test did not reject ambiguity neutrality even in the presence of ambiguity aversion. Compared with EZ preference, the performance measures R^2, RMSE, and HJD of the SA model are slightly lower. Based on the magnitude of relative risk aversion, the SA model delivers more plausible results.

As in our simulation study, the large standard errors of the risk aversion and ambiguity parameters show that these parameters are hard to estimate precisely. An economic interpretation why it is difficult to identify the two parameters separately is that once the agent knows the economic model, the remaining uncertainty of the distribution of returns might be of minor relevance. If returns are well described by the economic model, ambiguity accounts for a major part of the overall uncertainty and the impact of risk aversion on asset prices is relatively small.

5.2   Estimated pricing kernel

Figure 1 shows the realized pricing kernels of the EZ and SA models. The shaded areas represent NBER recessions. The estimated pricing kernels are always positive and thus satisfy the no arbitrage condition. Economic theory suggests that an investor evaluates payoffs more highly when economic conditions are bad, i.e. during recessions. Figure 1 shows that the realized pricing kernels have a clear business cycle pattern. As consumption growth and the return on wealth are low during recessions, the realized pricing kernels are highest during these periods.

The two realized pricing kernels show a similar behavior over time. The corre-lation between the two time-series is 0.81. In the estimation, we force the mean of the realized pricing kernel to match the inverse of the average real quarterly gross return on the risk-free asset, which is 1.0029 in our sample. Thus, the average pricing kernels are both close to one. Figure 1 shows that the peaks in the pricing kernel are more pronounced for the EZ model. An ambiguity averse investor pays relatively little attention to single extreme outcomes in consumption growth and the return on wealth. She rather cares about the expected utility conditional on the economic model at hand, respectively its certainty equivalent, which leads to less extreme values of the pricing kernel.

To provide deeper insights into how ambiguity distorts the pricing kernel, we decompose Equation (3) into three parts $\xi_{t,t+1} = \xi_{t,t+1}^{CRRA} \times \xi_{t,t+1}^{EZ} \times \xi_{t,t+1}^{SA}$, with

$$
\begin{aligned}
\xi_{t,t+1}^{CRRA} &= e^{-\delta} \left(\frac{C_{t+1}}{C_t} \right)^{-\rho}, \\
\xi_{t,t+1}^{EZ} &= \left(e^{-\delta} \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R_{t,t+1}^w \right)^{\theta_1 - 1}, \\
\xi_{t,t+1}^{SA} &= \left(\mathbb{E}_{\pi_{t+1}} \left[\left(e^{-\delta} \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R_{t,t+1}^w \right)^{\theta_1} \right] \right)^{\theta_2 - 1}.
\end{aligned}
\tag{7}
$$

If investors know the economic model, the conditional expectation is a constant and does not contain any additional information. If ambiguity is present and investors care about it, the question is whether the conditional expectation matters for the pricing of assets. In order to improve the fit to the cross-section of expected returns, ξ^{SA} has to carry additional information compared with ξ^{CRRA} and ξ^{EZ}. The sample correlation between ξ^{CRRA} and ξ^{SA} is about 0.25. If ξ^{EZ} and ξ^{SA} have a correlation close to one, the introduction of ambiguity is basically relabeling risk as ambiguity. For the realized pricing kernel of the SA model, the sample correlation between ξ^{EZ} and ξ^{SA} is 0.53. This shows that ambiguity matters for asset prices.

5.3 Pricing performance

Figure 2 displays realized versus predicted returns of the 3-month Treasury bill, the CRSP market return, and 30 portfolios formed on size, book-to-market, and industry. Predicted returns are calculated using the covariance decomposition (see Appendix A). Table 7 reports cross-sectional relative pricing errors (in percent), which are the square-roots of the mean squared differences between realized and predicted returns divided by the square-roots of the mean squared returns.

How can ambiguity help explaining the equity premium and the cross-sectional variation in expected returns? The EZ model only accounts for the covariation of returns with consumption growth and the return on wealth. The pricing kernel of the SA model contains the additional term ξ^{SA}. Hence, it also accounts for the covariance between returns and the continuation value of the time $t + 1$ economic model. Consider a portfolio which has low returns whenever the economic model is unfavorable, i.e. when it yields a low continuation value. Ambiguity averse investors command a premium for bearing this uncertainty (ambiguity premium). Compared to the EZ model, the expected return on such an asset is higher. Thus, the SA model may help explaining the returns of portfolios which are highly exposed to ξ^{SA}. If consumption growth and the return on wealth already characterize the economic model rather well, i.e. $\xi^{CRRA} \times \xi^{EZ}$ and ξ^{SA} are highly correlated, the ambiguity premium can be replicated by amplifying the risk factors of the EZ model. This can be achieved by using a high value of relative risk aversion. In Section 5.2, we have seen that the correlation is 0.53. Thus, we expect the risk factors of the EZ model to replicate the ones of the SA model to some extent, but not entirely.

Figure 2 shows that both models perform similarly in matching the equity premium. In the data, it is 1.61% per quarter, while it amounts to 2.05% in the EZ model and 2.01% in the SA model. To quantify the contribution of ambiguity, we

decompose the equity premium into a risk premium and an ambiguity premium.[13] We find that the risk premium accounts for 31.19% of the equity premium, while the ambiguity premium makes up the remaining 68.81%. For lower values of the EIS, the relative contribution of the ambiguity premium is smaller, but still above 50%.

Table 7 and Figure 2 show that both models have a remarkable fit to the 10 size sorted portfolios, with the SA model slightly outperforming. Both have difficulties in accurately pricing book-to-market portfolios. Concerning the industry sorted portfolios, although the average pricing performance is similar across models, there are some noteworthy differences for the individual industries. Consider for instance the industry portfolios 1 (non-durables) and 4 (energy). The pricing error of portfolio 1 reduces from 0.76% in the EZ model to 0.52% in the SA model, while for portfolio 4 it increases from 0.20% to 0.46%. Consistent with the arguments above, the correlation between the return on portfolio 1 and ξ^{SA} is relatively large in absolute terms, while it is low for portfolio 4. In line with this, the ambiguity premium contributes more to the total premium for portfolio 1 compared with portfolio 4.

We also investigate the fit with respect to 10 long-term reversal portfolios, portfolios formed on dividend yield, and portfolios based on two corporate profitability measures, the earnings to price ratio and the cash-flow to price ratio. As these portfolios were not used in the estimation, pricing these assets constitutes a test of the out-of-sample performance of the preference models. Table 7 shows that the SA model prices the portfolios more accurately than the EZ model, in particular the long-term reversal and the dividend yield sorted portfolios. Overall, we find that accounting for ambiguity aversion is useful in explaining the cross-section of expected returns. Even thought the pricing performance of the SA model is only

---

[13]Exploiting the approximation $\log(x) \approx x - 1$ yields

$$-\frac{\text{Cov}(\xi^{CRRA}\xi^{EZ}\xi^{SA}, \mathbf{R})}{\mathbb{E}[\xi]} \approx -\frac{\text{Cov}(\xi^{CRRA}\xi^{EZ}, \mathbf{R})}{\mathbb{E}[\xi]} - \frac{\text{Cov}(\xi^{SA}, \mathbf{R})}{\mathbb{E}[\xi]}.$$

The first summand is labeled *risk premium* and the second *ambiguity premium*.

slightly better for some of the test portfolios, it is able to explain the cross-sectional variation in expected returns with a more reasonable level of risk aversion.

6   Conclusion

Several recent studies show that ambiguity may have a significant impact on asset prices. However, there is little research investigating whether ambiguity aversion is actually present in the prices of traded assets and how consumption-based asset pricing models that account for ambiguity perform in explaining the cross-section of expected returns. To the best of our knowledge, this is the first study which estimates the SA model based on financial market data. Our point estimates of the ambiguity parameter are between 25 and 40, while relative risk aversion is clearly lower and within the range considered plausible by Mehra and Prescott (1985). This indicates that market participants are ambiguity averse.

We analyze whether the SA model is able to explain the cross-section of expected returns and if it improves upon EZ preferences. We find that ambiguity helps explaining the cross-sectional variation in expected returns. However, solely based on pricing errors and commonly employed model specification tests, it is difficult to discriminate between the two decision models. Our simulation study shows that even in an economy where ambiguity has a perceptible impact on asset prices, the pricing performances are similar. In the SA model, there is an additional priced factor which compensates for bearing model uncertainty. Thus, the total equity premium constitutes a risk premium and an ambiguity premium. If ambiguity is neglected, matching the equity premium and the cross-section of expected returns requires a high level of relative risk aversion to make up for the missing ambiguity premium. The SA model can account for the patterns in expected stock returns with lower relative risk aversion and thus provides a more reasonable explanation of asset prices.

A   GMM Estimation and Model Evaluation

Let $\mathbf{R}_{t,t+1} = (R^1_{t,t+1}, \ldots, R^n_{t,t+1})'$ denote the vector of gross returns. The population moment conditions are

$$\mathbb{E}\left[\mathbf{f}(\Theta, \mathbf{R}_{t,t+1})\right] = 0.$$

For gross returns, we have $\mathbf{f}(\Theta, \mathbf{R}_{t,t+1}) = \xi_{t,t+1}(\Theta)\mathbf{R}_{t,t+1} - \mathbf{1}$ and for excess returns $\mathbf{f}(\Theta, \mathbf{R}_{t,t+1}) = \xi_{t,t+1}(\Theta)\left(\mathbf{R}_{t,t+1} - R^f_{t,t+1}\mathbf{1}\right)$, where we make the dependence of the pricing kernel on the parameter vector Θ explicit. The sample equivalent of the population orthogonality conditions is

$$\mathbf{g}_T(\Theta, \mathbf{R}) = \frac{1}{T}\sum_{t=0}^{T-1}\mathbf{f}(\Theta, \mathbf{R}_{t,t+1}).$$

The GMM estimator of Θ, denoted by $\hat{\Theta}$, minimizes the criterion function

$$\hat{\Theta} = \mathrm{argmin}_\Theta \; \mathbf{g}_T(\Theta, \mathbf{R})' \, \mathbf{W}_T \, \mathbf{g}_T(\Theta, \mathbf{R}),$$

where $\mathbf{W}_T$ is a positive semi-definite weighting matrix which converges in probability to a positive definite matrix of constants $\mathbf{W}$.

Let $\mathbf{S}$ denote the variance covariance matrix of the moment conditions. Hansen (1982) shows that the weighting matrix that minimizes the asymptotic variance of the parameter estimates is $\mathbf{W} = \mathbf{S}^{-1}$. Even if $\mathbf{S}$ is not used as weighting matrix, we still need to estimate $\mathbf{S}$ to construct standard errors and to test hypotheses. We estimate the long-run covariance matrix $\mathbf{S}$ according to Newey and West (1987b). As in Yogo (2006), the lag length is set to one to account for time aggregation in consumption data.

Hansen (1982) shows that the variance of the parameter estimates and the estimated moment conditions are

$$\mathrm{Var}(\hat{\Theta}) \;\;=\;\; \frac{1}{T}\left(\mathbf{D}'\mathbf{W}\mathbf{D}\right)^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}\left(\mathbf{D}'\mathbf{W}\mathbf{D}\right)^{-1},$$

$$\mathrm{Var}\left(\mathbf{g}_T(\hat{\Theta}, \mathbf{R})\right) \;\;=\;\; \frac{1}{T}\left(\mathbf{I}_n - \mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\right)'\mathbf{S}\left(\mathbf{I}_n - \mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\right),$$

where $\mathbf{D} = \mathbb{E}\left[\partial\mathbf{f}(\Theta, \mathbf{R}_{t,t+1})/\partial\Theta'\right]$ and $\mathbf{I}_n$ is an n-dimensional identity matrix.

To perform inference on the parameters, we use Wald tests. Let $\mathbf{C}$ denote a $k \times k$ matrix and $\mathbf{c}$ a $k \times 1$ vector, where k denotes the number of estimated parameters. Newey and West (1987a) show that the restrictions

$$\mathbf{C}\Theta - \mathbf{c} = 0$$

can be tested using the following test statistic

$$(\mathbf{C}\hat{\Theta} - \mathbf{c})'\left(\mathbf{C}\,\mathrm{Var}(\hat{\Theta})\,\mathbf{C}'\right)^{-1}(\mathbf{C}\hat{\Theta} - \mathbf{c}),$$

which has a χ^2 distribution with degrees of freedom equal to the rank of $\mathbf{C}$. We use this test to check for time-separability ($\gamma = \rho$) and ambiguity neutrality ($\gamma = \eta$).

We perform Hansen's J-test for overidentifying restrictions

$$\mathbf{g}_T(\hat{\Theta}, \mathbf{R})' \operatorname{Var}\left(\mathbf{g}_T(\hat{\Theta}, \mathbf{R})\right)^{-1} \mathbf{g}_T(\hat{\Theta}, \mathbf{R}),$$

which has a χ^2 distribution with n (number of moments) minus k (number of estimated parameters) degrees of freedom under the null of correct model specification. Since the variance covariance matrix of the moment conditions is singular, we use the Moore-Penrose pseudoinverse.

We compare root mean squared errors (RMSE) and Hansen and Jagannathan (1997) distances (HJD) of different preference specifications and parameter vectors. The RMSE is the square-root of the objective function that is minimized if the identity matrix is used for weighting

$$\operatorname{RMSE}(\Theta, \mathbf{R}) = \sqrt{\mathbf{g}_T(\Theta, \mathbf{R})' \, \mathbf{I}_n \, \mathbf{g}_T(\Theta, \mathbf{R})}.$$

The Hansen and Jagannathan (1997) distance is given by

$$\operatorname{HJD}(\Theta, \mathbf{R}) = \sqrt{\mathbf{g}_T(\Theta, \mathbf{R})' \, G(\mathbf{R})^{-1} \, \mathbf{g}(\Theta, \mathbf{R})},$$

i.e. the square-root of the objective function if the second moment matrix $G(\mathbf{R})$, with $G(\mathbf{R})_{i,j} := \frac{1}{T} \sum_{t=0}^{T-1} R_{t,t+1}^i R_{t,t+1}^j$, is used for weighting.

We use the methodology of Jagannathan and Wang (1996) and Parker and Julliard (2005) to test if the RMSE or the HJD is zero. Let

$$\mathbf{A} = \mathbf{S}^{1/2} \left(\mathbf{W} - \mathbf{W}\,\mathbf{D}\,(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\,\mathbf{D}'\,\mathbf{W}\right)(\mathbf{S}^{1/2})',$$

where $\mathbf{S}^{1/2}$ is the upper-triangular matrix from the Cholesky decomposition of $\mathbf{S}$. The matrix $\mathbf{A}$ has $n - k$ positive eigenvalues, denoted by $\lambda_1, ..., \lambda_{n-k}$. The asymptotic sampling distribution of the distance is

$$T\,\mathbf{g}_T(\hat{\Theta}, \mathbf{R})' \, \mathbf{W}_T \, \mathbf{g}_T(\hat{\Theta}, \mathbf{R}) \xrightarrow{d} u = \sum_{j=1}^{n-k} \lambda_j v_j \text{ as } T \to \infty,$$

where $v_1, ..., v_{n-k}$ are independent χ^2 random variables with one degree of freedom. The empirical p-value of the statistic can be computed by drawing $T(n - k)$ independent random numbers from a $\chi^2(1)$ distribution and counting the number of cases where u exceeds the test statistic. If the efficient weighting matrix is used, i.e. $\mathbf{W} = \mathbf{S}^{-1}$, all eigenvalues are unity and this test coincides with the J-test.

The cross-sectional R^2 is computed as

$$R^2 \;\; = \;\; 1 - \frac{\operatorname{Var}\left(\mathbb{E}\left[\mathbf{R}_{t,t+1}\right] - \mathbb{E}_p\left[\mathbf{R}_{t,t+1}\right]\right)}{\operatorname{Var}\left(\mathbb{E}\left[\mathbf{R}_{t,t+1}\right]\right)},$$

where the predicted returns are calculated using the covariance decomposition

$$\mathbb{E}_p\left[\mathbf{R}_{t,t+1}\right] \;\; = \;\; \frac{1 - \operatorname{Cov}\left[\xi_{t,t+1}, \mathbf{R}_{t,t+1}\right]}{\mathbb{E}\left[\xi_{t,t+1}\right]}.$$

B Model Solution

We solve the model in the same manner as Bansal and Yaron (2004), Bansal et al. (2012a,b), and Beeler and Campbell (2012) using analytical approximations. We assume that the log wealth-consumption ratio z is affine in the state variables

$$z_t = A + B_x x_t + B_\sigma(\sigma_t^2 - \sigma^2).$$

For the log return on the consumption claim $r_{t,t+1}^w = \log R_{t,t+1}^w$, we use the log-linear return approximation of Campbell and Shiller (1988)

$$r_{t,t+1}^w = k_0 + k_1 z_{t+1} - z_t + \Delta c_{t+1},$$

where k_0 and k_1 are linearizing constants. It holds that $k_0 = \log(1 + e^{\bar{z}}) - k_1 \bar{z}$, where $k_1 = \frac{e^{\bar{z}}}{1+e^{\bar{z}}}$ and $\bar{z}$ is the long-run mean of the log wealth-consumption ratio. Using the Euler equation $\mathbb{E}_t\left[e^{\log \xi_{t,t+1} + r_{t,t+1}^w}\right] = 1$ yields the following coefficients of the wealth-consumption ratio

$$
\begin{aligned}
A &= \frac{1}{1-k_1}\left(-\delta + k_0 + (1-\rho)\mu_c + (1-k_1\varphi_\sigma)\frac{B_\sigma}{\varphi_\sigma}\sigma^2 + \frac{1-\eta}{2(1-\rho)}\left(\frac{B_\sigma}{\varphi_\sigma}\phi_\sigma\right)^2\right), \\
B_x &= \frac{1-\rho}{1-k_1\varphi_x}\varphi_x, \\
B_\sigma &= \frac{(1-\gamma)(1-\rho)}{2(1-k_1\varphi_\sigma)}\left(1 + \frac{1-\eta}{1-\gamma}\left(\frac{\phi_x}{1-k_1\varphi_x}\right)^2\right)\varphi_\sigma.
\end{aligned}
\tag{8}
$$

By substituting the return on the consumption claim into Equation (3), we obtain an expression of the log pricing kernel in terms of the state variables

$$\log \xi_{t,t+1} = s_0 + s_x x_t + s_\sigma(\sigma_t^2 - \sigma^2) - \Lambda_c \sigma_{t+1} w_{t+1}^c - \Lambda_x \phi_x \sigma_{t+1} w_{t+1}^x - \Lambda_\sigma \phi_\sigma w_{t+1}^\sigma,$$

with the drift characterized by the coefficients

$$
\begin{aligned}
s_0 &= -\delta - \rho\mu_c - \frac{(1-\eta)(\rho-\eta)}{2(1-\rho)^2}\left(\frac{B_\sigma}{\varphi_\sigma}\phi_\sigma\right)^2 + \frac{s_\sigma}{\varphi_\sigma}\sigma^2, \\
s_x &= -\rho\varphi_x, \\
s_\sigma &= \frac{1}{2}(1-\gamma)(\gamma-\rho)\varphi_\sigma + \frac{1}{2}(1-\eta)(\eta-\rho)\left(\frac{\phi_x}{1-k_1\varphi_x}\right)^2\varphi_\sigma.
\end{aligned}
$$

The market prices of uncertainty in consumption, expected consumption growth, and volatility are determined by

$$
\begin{aligned}
\Lambda_c &= \gamma, \\
\Lambda_x &= \frac{\eta-\rho}{1-\rho}k_1 B_x + \eta, \\
\Lambda_\sigma &= \frac{\eta-\rho}{1-\rho}k_1 B_\sigma + \frac{1}{2}(\eta-\gamma)(1-\gamma).
\end{aligned}
\tag{9}
$$

Given the log pricing kernel, the continuously compounded risk-free rate is

$$
\begin{aligned}
r^f_{t,t+1} &= -\log \mathbb{E}_t \left(e^{\log \xi_{t,t+1}} \right) \\
&= r^f_0 + r^f_x x_t + r^f_\sigma (\sigma^2_t - \sigma^2),
\end{aligned}
$$

with

$$
\begin{aligned}
r^f_0 &= -s_0 - \frac{1}{2} \left(\Lambda_\sigma - \frac{1}{2} \left(\Lambda^2_c + \Lambda^2_x \phi^2_x \right) \right)^2 \phi^2_\sigma - \frac{1}{2} \left(\Lambda^2_c + \Lambda^2_x \phi^2_x \right) \sigma^2, \\
r^f_x &= -s_x, \\
r^f_\sigma &= -s_\sigma - \frac{1}{2} \left(\Lambda^2_c + \Lambda^2_x \phi^2_x \right) \varphi_\sigma.
\end{aligned}
$$

To solve for the price-dividend ratio z^d, we rely on the log-linear approximation of the log return on the dividend claim $r^d_{t,t+1} = \log R^d_{t,t+1}$

$$
r^d_{t,t+1} = k^d_0 + k^d_1 z^d_{t+1} - z^d_t + \Delta d_{t+1},
$$

where $k^d_0 = \log(1 + e^{\bar{z}^d}) - k^d_1 \bar{z}^d$, $k^d_1 = \frac{e^{\bar{z}^d}}{1+e^{\bar{z}^d}}$, and $\bar{z}^d$ denotes the long-run mean of the log price-dividend ratio. We conjecture that the log price-dividend ratio z^d is affine in the state variables

$$
z^d_t = A^d + B^d_x x_t + B^d_\sigma (\sigma^2_t - \sigma^2).
$$

The coefficients of the log price-dividend ratio follow by applying the Euler equation to the log return on the dividend claim

$$
\begin{aligned}
A^d &= \frac{s_0 + k^d_0 + \mu_d + \frac{1}{2} \left(\frac{B^d_\sigma}{\varphi_\sigma} - \frac{s_\sigma}{\varphi_\sigma} - \Lambda_\sigma \right)^2 \phi^2_\sigma + \left((1 - k^d_1 \varphi_\sigma) \frac{B^d_\sigma}{\varphi_\sigma} - \frac{s_\sigma}{\varphi_\sigma} \right) \sigma^2}{(1 - k^d_1)}, \\
B^d_x &= \frac{\lambda - \rho}{1 - k^d_1 \varphi_x} \varphi_x, \\
B^d_\sigma &= \frac{s_\sigma + \frac{1}{2} \left(\left(k^d_1 B^d_x + \lambda - \Lambda_x \right)^2 \phi^2_x + \gamma^2 + \phi^2_{d,\sigma} - 2\gamma \rho_{cd} \phi_{d,\sigma} \right) \varphi_\sigma}{1 - k^d_1 \varphi_\sigma}.
\end{aligned}
$$

The conditional expected return on the dividend claim is given by

$$
\begin{aligned}
\mathbb{E}_t \left[r^d_{t,t+1} \right] = \ & k^d_0 + (k^d_1 - 1)A^d + \mu_d \\
& + ((k^d_1 \varphi_x - 1)B^d_x + \lambda \varphi_x)x_t + (k^d_1 \varphi_\sigma - 1)B^d_\sigma (\sigma^2_t - \sigma^2).
\end{aligned}
$$

The equity premium follows by subtracting the risk-free rate from the expected return on the dividend claim

$$
\begin{aligned}
\mathbb{E}_t \left[r^d_{t,t+1} \right] - r^f_{t,t+1} = \ & \frac{1}{2} \left(\left(\Lambda_\sigma - \frac{1}{2} (\Lambda^2_c + \Lambda^2_x \phi^2_x) \right)^2 - \left(\Lambda_\sigma - \frac{B^d_\sigma - s_\sigma}{\varphi_\sigma} \right)^2 \right) \phi^2_\sigma \\
& + \left(\Lambda_x \left(k^d_1 B^d_x + \lambda \right) \phi^2_x - \frac{1}{2} \left(k^d_1 B^d_x + \lambda \right)^2 \phi^2_x + \Lambda_c \phi_{d,\sigma} \rho_{cd} - \frac{1}{2} \phi^2_{d,\sigma} \right) \sigma^2 \\
& + \left(\Lambda_x \left(k^d_1 B^d_x + \lambda \right) \phi^2_x - \frac{1}{2} \left(k^d_1 B^d_x + \lambda \right)^2 \phi^2_x + \Lambda_c \phi_{d,\sigma} \rho_{cd} - \frac{1}{2} \phi^2_{d,\sigma} \right) \varphi_\sigma (\sigma^2_t - \sigma^2).
\end{aligned}
$$

C   Data

Risk-free rate: We use the 3-month secondary market Treasury bill rate from the H.15 release of the Federal Reserve Board of Governors (`http://www.federalreserve.gov/releases/h15/data.htm`) as risk-free rate.

Stock returns: All stock returns are taken from Kenneth French's homepage (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`), including the CRSP value weighted stock return index, which we use as proxy for the return on the stock market. As test assets, we employ the return on the 3-month Treasury bill, the CRSP value weighted stock return, and the returns on 30 additional equity portfolios. Among these, 10 value weighted portfolios are formed on size (market equity) at the end of each June using NYSE breakpoints, 10 value weighted portfolios formed on BE/ME (book equity at the last fiscal year end of the prior calendar year divided by market equity at the end of December of the prior year) at the end of each June using NYSE breakpoints, and 10 industry portfolios (the sectors are Consumer Nondurables, Consumer Durables, Manufacturing, Energy, Business Equipment, Telecommunication and Television, Retail, Healthcare, Utilities, and Other) also formed at the end of each June. In Section 5.3, we also use the returns on 10 portfolios formed on long-term reversal, 10 dividend yield sorted portfolios, 10 portfolios formed on earnings to price ratios, and 10 cash-flow to price sorted portfolios. For a detailed description of the return data, see the URL above.

Inflation: All returns are deflated using the seasonally adjusted Consumer Price Index (CPI). We obtain the CPI from the Bureau of Labor Statistics (`http://www.bls.gov/cpi`). Quarterly inflation is the growth rate of the CPI in the final month of the current quarter over the final month of the previous quarter.

Consumption and return on wealth: We use the same definitions of consumption, labor income, asset holdings, and cay as in Lettau and Ludvigson (2001a). The updated data is available on Martin Lettau's homepage (`http://faculty.haas.berkeley.edu/lettau/data_cay.html`). Lettau and Ludvigson (2001a) define aggregate consumption as expenditures on nondurables and services, excluding shoes and clothing. The quarterly data is seasonally adjusted at annual rates, in billions of chain-weighted dollars. Labor income is defined as wages and salaries plus transfer payments plus other labor income minus personal contributions for social insurance minus taxes. Asset holdings is household net worth in billions of current dollars. We refer to Lettau and Ludvigson (2001a) for a more detailed description of the data.

Price-dividend ratio: The price-dividend ratio (based on the S&P Composite index) is taken from Robert Shiller's homepage (`http://www.econ.yale.edu/~shiller/data.htm`).

References

AHN, D., S. CHOI, D. GALE, AND S. KARIV (2011): "Estimating Ambiguity Aversion in a Portfolio Choice Experiment," Working Paper.

AHN, S. AND C. GADAROWSKI (2004): "Small Sample Properties of the GMM Specification Test Based on the Hansen-Jagannathan Distance," *Journal of Empirical Finance*, 11, 109–132.

ALTONJI, J. AND L. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14, 353–366.

ANDERSEN, E., L. HANSEN, AND T. SARGENT (2000): "Robustness, Detection, and the Price of Risk," Working Paper.

BANSAL, R., A. GALLANT, AND G. TAUCHEN (2007): "Rational Pessimism, Rational Exuberance, and Asset Pricing Models," *Review of Economic Studies*, 74, 1005–1033.

BANSAL, R., V. KHATCHATRIAN, AND A. YARON (2005): "Interpretable Asset Markets?" *European Economic Review*, 49, 531–560.

BANSAL, R., D. KIKU, AND A. YARON (2012a): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2012b): "Risks for the Long Run: Estimation with Time Aggregation," Working Paper.

BANSAL, R. AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

BEELER, J. AND J. CAMPBELL (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

BOSSAERTS, P., P. GHIRARDATO, S. GUARNASCHELLI, AND W. ZAME (2010): "Ambiguity in Asset Markets: Theory and Experiment," *Review of Financial Studies*, 23, 1325–1359.

BREEDEN, D. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

CAMPBELL, J. (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104, 298–345.

CAMPBELL, J. AND N. MANKIW (1989): "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," *NBER Macroeconomics Annual*, 4, 185–246.

CAMPBELL, J. AND R. SHILLER (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

CHEN, H., N. JU, AND J. MIAO (2011): "Dynamic Asset Allocation with Ambiguous Return Predictability," Working Paper.

CHEN, X., J. FAVILUKIS, AND S. LUDVIGSON (2012): "An Estimation of Economic Models with Recursive Preferences," Working Paper.

COCHRANE, J. (2005): *Asset Pricing*, New Jersey: Princeton University Press.

COLLARD, F., S. MUKERJI, K. SHEPPARD, AND J. TALLON (2011): "Ambiguity and the Historical Equity Premium," Working Paper.

CONSTANTINIDES, G. AND A. GHOSH (2011): "Asset Pricing Tests with Long-Run Risks in Consumption Growth," *Review of Asset Pricing Studies*, 1, 96–136.

DRECHSLER, I. AND A. YARON (2011): "What's Vol Got to Do with It," *Review of Financial Studies*, 24, 1–45.

DUFFIE, D. AND C. SKIADAS (1994): "Continuous-Time Security Pricing: A Utility Gradient Approach," *Journal of Mathematical Economics*, 23, 107–131.

ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

EPSTEIN, L. AND M. SCHNEIDER (2010): "Ambiguity and Asset Markets," *Annual Review of Financial Economics*, 2, 315–346.

EPSTEIN, L. AND S. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

——— (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

FAN, J. (1992): "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87, 998–1004.

FERSON, W. AND S. FOERSTER (1994): "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," *Journal of Financial Economics*, 36, 29–55.

GILBOA, I. AND D. SCHMEIDLER (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

GUVENEN, F. (2006): "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective," *Journal of Monetary Economics*, 53, 1451–1472.

HALEVY, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.

HALL, R. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–356.

HANSEN, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1084.

HANSEN, L., J. HEATON, J. LEE, AND N. ROUSSANOV (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier.

HANSEN, L., J. HEATON, AND N. LI (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116, 260–302.

HANSEN, L., J. HEATON, AND A. YARON (1996): "Finite Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics*, 14, 262–280.

HANSEN, L. AND R. JAGANNATHAN (1997): "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance*, 52, 557–590.

HANSEN, L. AND T. SARGENT (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

——— (2008): *Robustness*, New Jersey: Princeton University Press.

HANSEN, L. AND K. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

HAYASHI, T. AND J. MIAO (2011): "Intertemporal Substitution and Recursive Smooth Ambiguity Preferences," *Theoretical Economics*, 6, 423–472.

Jagannathan, R. and Z. Wang (1996): "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3–53.

Ju, N. and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80, 559–591.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 76, 1849–1892.

——— (2009): "Recursive Smooth Ambiguity Preferences," *Journal of Economic Theory*, 144, 930–976.

Kreps, D. and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Lettau, M. and S. Ludvigson (2001a): "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815–849.

——— (2001b): "Resurrecting the (C)CAPM: A Cross-Sectional Test when Risk Premia are Time-Varying," *Journal of Political Economy*, 109, 1238–1287.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A Skeptical Appraisal of Asset Pricing Tests," *Journal of Financial Economics*, 96, 175–194.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 48, 1149–1168.

Ludvigson, S. (2012): "Advances in Consumption-Based Asset Pricing: Empirical Tests," in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. Stulz, Amsterdam: Elsevier.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2012): "The Wealth-Consumption Ratio," Working Paper.

Malloy, C., T. Moskowitz, and A. Vissing-Jorgensen (2009): "Long-Run Stockholder Consumption Risk and Asset Returns," *Journal of Finance*, 64, 2427–2479.

Mehra, R. and E. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Miao, J., B. Wei, and H. Zhou (2012): "Ambiguity Aversion and Variance Premium," Working Paper.

Nagel, S. and K. Singleton (2011): "Estimation and Evaluation of Conditional Asset Pricing Models," *Journal of Finance*, 66, 873–909.

NEWEY, W. AND K. WEST (1987a): "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, 777–787.

——— (1987b): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PARK, H. AND L. STEFANSKI (1998): "Relative-Error Prediction," *Statistics and Probability Letters*, 40, 227–236.

PARKER, J. AND C. JULLIARD (2005): "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 113, 185–222.

ROLL, R. (1977): "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics*, 4, 129–176.

RUPPERT, D. AND M. WAND (1994): "Multivariate Locally Weighted Least Squares Regression," *Annals of Statistics*, 22, 1346–1370.

SAVAGE, L. (1954): *The Foundations of Statistics*, New York: Dover.

SMITH, D. (1999): "Finite Sample Properties of Tests of the Epstein-Zin Asset Pricing Model," *Journal of Econometrics*, 93, 113–148.

STOCK, J. AND J. WRIGHT (2000): "GMM with Weak Identification," *Econometrica*, 68, 1055–1096.

VISSING-JORGENSEN, A. AND O. ATTANASIO (2003): "Stock-Market Participation, Intertemporal Substitution, and Risk Aversion," *American Economic Review*, 93, 383–391.

YOGO, M. (2004): "Estimating the Elasticity of Intertemporal Substitution when Instruments are Weak," *Review of Economics and Statistics*, 86, 797–810.

——— (2006): "A Consumption-Based Explanation of Expected Stock Returns," *Journal of Finance*, 61, 539–580.

ZHANG, Q. (2006): "Human Capital, Weak Identification, and Asset Pricing," *Journal of Money, Credit, and Banking*, 38, 879–899.

	Mean	Min	Max	Std. Dev.	Skew	Kurt	AC1
Log Consumption Growth	0.46	-1.25	2.10	0.47	-0.37	4.25	0.45
Risk-Free Rate	0.28	-1.75	3.55	0.67	0.44	5.52	0.47
Log Price-Dividend Ratio	3.50	2.78	4.50	0.40	0.51	2.59	0.97
Market Return ($CRSP$)	1.52	-31.38	20.17	8.57	-0.83	4.29	0.10
Market Return (cay)	1.48	-2.01	4.86	0.81	-0.09	5.44	0.14
Size 1	1.96	-40.79	34.43	12.78	-0.34	3.71	0.06
Size 2	1.87	-39.65	35.28	12.10	-0.50	3.89	0.01
Size 3	2.04	-37.79	32.33	11.54	-0.60	4.02	-0.01
Size 4	1.94	-40.97	29.91	11.18	-0.68	4.24	-0.01
Size 5	2.02	-32.97	29.63	10.76	-0.66	3.95	0.02
Size 6	1.94	-34.17	26.75	10.05	-0.70	3.95	0.03
Size 7	1.93	-31.09	28.49	10.00	-0.67	3.99	0.06
Size 8	1.83	-26.89	24.47	9.61	-0.72	3.81	0.04
Size 9	1.74	-28.64	19.60	8.83	-0.89	4.15	0.08
Size 10	1.39	-34.41	20.33	8.05	-0.88	4.74	0.13
BM 1	1.27	-44.18	23.00	10.05	-0.91	5.04	0.10
BM 2	1.42	-31.10	23.05	8.93	-0.77	4.52	0.02
BM 3	1.64	-29.20	20.36	8.69	-0.65	3.79	0.05
BM 4	1.57	-28.40	22.00	8.78	-0.79	4.24	0.10
BM 5	1.82	-28.57	18.89	8.07	-0.79	3.93	0.09
BM 6	1.81	-29.49	20.61	8.33	-0.74	4.10	0.13
BM 7	1.80	-23.30	24.66	8.41	-0.50	3.44	0.15
BM 8	2.20	-42.00	30.44	8.99	-0.78	5.68	0.12
BM 9	2.25	-25.26	27.46	9.21	-0.55	3.85	0.06
BM 10	2.35	-37.26	39.23	11.02	-0.33	4.50	0.13
Industry 1 (NoDur)	2.03	-37.67	26.30	8.51	-0.89	5.44	0.04
Industry 2 (Durbl)	1.26	-47.01	35.42	11.88	-0.64	4.60	0.12
Industry 3 (Manuf)	1.60	-34.94	24.58	9.40	-0.83	4.43	0.09
Industry 4 (Enrgy)	1.87	-31.51	23.38	8.92	-0.70	3.90	0.10
Industry 5 (HiTec)	1.67	-43.07	32.98	12.27	-0.83	5.01	0.08
Industry 6 (Telcm)	1.41	-26.44	23.04	8.39	-0.47	3.57	0.16
Industry 7 (Shops)	1.73	-43.66	36.09	10.03	-0.73	5.79	0.03
Industry 8 (Hlth)	2.02	-43.76	21.95	9.54	-1.01	5.82	0.04
Industry 9 (Utils)	1.55	-26.88	23.20	7.30	-0.69	4.34	0.10
Industry 10 (Other)	1.39	-31.66	21.63	10.34	-0.84	3.93	0.09

Table 1: Summary Statistics

The table shows descriptive statistics of log consumptions growth, the risk-free rate, the log price-dividend ratio, two proxies for the market return, and the 30 equity test assets. The log consumption growth rate and the continuously compounded returns are expressed in percentage terms. The sample period is from the first quarter of 1952 to the third quarter of 2011. The data is described in Appendix C.

Preferences

ρ	δ	γ	η
0.6667	0.0033	$5 - 10$	$5 - 20$

Consumption and Dividends

μ_c	μ_d	λ	$\phi_{d,\sigma}$	ρ_{cd}
0.0045	0.0045	2.5	6.5024	0.40

Long-Run Growth Rate

φ_x	ϕ_x
0.9269	0.1140

Volatility

φ_σ	ϕ_σ	σ
0.9970	$4.84 \cdot 10^{-6}$	0.0125

Table 2: Parameters of the LRR Model

The table shows preference and model parameters in quarterly terms. The monthly values are taken from Bansal et al. (2012a).

N	δ	γ	η	R^2	RMSE	HJD	$\gamma = \rho$	$\eta = \gamma$	J-test
Panel 1: EZ with R^w ($\rho = 0.6667$)									
60	0.003	10.080		0.498	0.030	0.379	10.110		36.929
	(0.000)	(2.947)		(–)	(0.459)	(0.220)	(0.001)		(0.179)
Panel 2: EZ with R^d ($\rho = 0.6667$)									
60	0.000	1.704		0.165	0.054	0.538	4.676		45.869
	(0.010)	(0.484)		(–)	(0.002)	(0.000)	(0.031)		(0.032)
Panel 3: SA with R^w and true conditional expectation ($\rho = 0.6667$)									
60	0.003	6.107	19.294	0.603	0.031	0.406	0.700	2.647	39.751
	(0.000)	(6.528)	(3.809)	(–)	(0.411)	(0.183)	(0.403)	(0.104)	(0.088)
Panel 4: SA with R^w and approximated conditional expectation ($\rho = 0.6667$)									
60	0.003	6.016	19.245	0.603	0.031	0.406	0.627	2.626	39.770
	(0.000)	(6.622)	(3.800)	(–)	(0.407)	(0.185)	(0.428)	(0.105)	(0.088)
Panel 5: EZ with R^w, data based on SA model ($\rho = 0.6667$)									
60	0.004	14.293		0.435	0.037	0.441	20.091		46.546
	(0.000)	(3.037)		(–)	(0.126)	(0.048)	(0.000)		(0.028)

Table 3: Parameter Estimates (Simulated Data)

The table shows GMM estimates of the preference parameters. EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. HAC standard errors are in parentheses. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The table also reports the cross-sectional R^2, the Wald tests for time-separability ($\gamma = \rho$) and ambiguity neutrality ($\eta = \gamma$), and the J-test for overidentifying restrictions (p-values in parentheses). Details on the tests are provided in Appendix A. All estimates are based on 1,000 simulation runs of approximately equivalent length to the data ($N = 60$ years). The true parameters are $\Theta^{\mathrm{EZ}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 10)$, for Epstein and Zin (1989) preferences, and $\Theta^{\mathrm{SA}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 5, \eta = 20)$, for the smooth ambiguity model.

N	δ	γ	η	R^2	RMSE	HJD	$\gamma = \rho$	$\eta = \gamma$	J-test
Panel 1: EZ with R^w ($\rho = 0.6667$)									
100	0.003	10.121		0.630	0.024	0.282	17.089		32.825
	(0.000)	(2.281)		(−)	(0.486)	(0.319)	(0.000)		(0.330)
200	0.003	9.930		0.768	0.017	0.195	33.865		31.356
	(0.000)	(1.597)		(−)	(0.479)	(0.376)	(0.000)		(0.398)
500	0.003	10.075		0.892	0.011	0.123	87.600		30.638
	(0.000)	(1.007)		(−)	(0.453)	(0.380)	(0.000)		(0.433)
1000	0.003	10.069		0.944	0.008	0.086	174.601		29.913
	(0.000)	(0.712)		(−)	(0.484)	(0.400)	(0.000)		(0.470)
Panel 2: SA with R^w and true conditional expectation ($\rho = 0.6667$)									
100	0.003	6.145	19.762	0.720	0.024	0.303	1.171	4.853	34.376
	(0.000)	(5.039)	(2.986)	(−)	(0.454)	(0.283)	(0.279)	(0.028)	(0.226)
200	0.003	5.179	19.912	0.840	0.017	0.209	1.571	11.200	31.524
	(0.000)	(3.544)	(2.109)	(−)	(0.492)	(0.346)	(0.210)	(0.001)	(0.341)
500	0.003	5.278	20.168	0.928	0.011	0.132	4.197	27.905	30.249
	(0.000)	(2.248)	(1.342)	(−)	(0.473)	(0.367)	(0.041)	(0.000)	(0.402)
1000	0.003	5.139	20.192	0.963	0.008	0.093	7.911	56.956	29.135
	(0.000)	(1.594)	(0.954)	(−)	(0.476)	(0.397)	(0.005)	(0.000)	(0.458)

Table 4: Parameter Estimates (Simulated Data and Larger Sample Size)

The table shows GMM estimates of the preference parameters. EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. HAC standard errors are in parentheses. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The table also reports the cross-sectional R^2, the Wald tests for time-separability ($\gamma = \rho$) and ambiguity neutrality ($\eta = \gamma$), and the J-test for overidentifying restrictions (p-values in parentheses). Details on the tests are provided in Appendix A. The estimates are based on 1,000 simulation runs with sample size equal to N years. The true parameters are $\Theta^{\mathrm{EZ}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 10)$, for Epstein and Zin (1989) preferences, and $\Theta^{\mathrm{SA}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 5, \eta = 20)$, for the smooth ambiguity model.

Test	N	EZ			SA		
		1%	5%	10%	1%	5%	10%
RMSE	60	1.13	6.26	10.77	1.40	7.40	13.90
HJD	60	8.31	21.44	34.56	10.20	25.40	36.40
J-test	60	11.49	27.28	38.36	23.40	42.50	53.10
RMSE	100	1.20	4.50	10.70	0.70	6.90	12.20
HJD	100	4.20	14.00	22.40	4.80	17.20	26.20
J-test	100	6.00	16.70	25.50	10.10	24.70	33.10
RMSE	200	1.30	5.80	11.00	1.30	5.90	11.00
HJD	200	3.10	11.40	19.00	3.60	11.80	20.50
J-test	200	3.20	10.90	17.10	4.60	14.40	22.00
RMSE	500	1.20	6.10	12.10	2.30	8.20	13.80
HJD	500	2.50	8.20	16.20	3.40	10.20	17.50
J-test	500	1.90	7.00	12.20	3.10	9.60	14.80
RMSE	1000	1.10	6.10	11.10	0.90	5.80	10.90
HJD	1000	2.00	8.40	14.80	2.60	8.60	16.50
J-test	1000	0.80	6.30	11.10	1.70	6.50	12.40

Table 5: Finite Sample Behavior of Specification Tests

The table shows the proportion of the empirical p-values of the specification tests being smaller than the asymptotic size. EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The J-test refers to Hansen's test for overidentifying restrictions. Details on the tests are provided in Appendix A. All estimates are based on 1,000 simulation runs. The results are reported for three significance levels (1%, 5%, and 10%) and for different values of the sample size N in years.

ρ	δ	γ	η	R^2	RMSE	HJD	$\gamma = \rho$	$\eta = \gamma$	J-test
Panel 1: EZ (ρ unrestricted)									
0.563	0.011	27.688		0.404	0.021	0.690	0.116		85.434
(1.406)	(0.006)	(78.368)		(–)	(0.143)	(0.000)	(0.734)		(0.000)
Panel 2: EZ (ρ restricted)									
0.400	0.011	36.292		0.402	0.021	0.691	8.052		90.750
(–)	(0.001)	(12.649)		(–)	(0.242)	(0.000)	(0.005)		(0.000)
0.500	0.011	31.075		0.403	0.021	0.690	8.130		91.679
(–)	(0.001)	(10.723)		(–)	(0.239)	(0.000)	(0.004)		(0.000)
0.667	0.010	21.843		0.402	0.021	0.690	8.211		92.921
(–)	(0.001)	(7.390)		(–)	(0.233)	(0.000)	(0.004)		(0.000)
Panel 3: SA (ρ restricted)									
0.400	0.012	4.283	40.545	0.475	0.020	0.616	0.016	1.037	87.505
(–)	(0.000)	(31.009)	(13.703)	(–)	(0.088)	(0.000)	(0.900)	(0.308)	(0.000)
0.500	0.012	2.295	35.091	0.477	0.020	0.612	0.004	1.125	85.991
(–)	(0.000)	(26.836)	(11.744)	(–)	(0.088)	(0.000)	(0.947)	(0.289)	(0.000)
0.667	0.011	0.819	24.817	0.474	0.020	0.608	0.000	1.184	84.046
(–)	(0.000)	(19.217)	(8.146)	(–)	(0.084)	(0.000)	(0.994)	(0.276)	(0.000)

Table 6: Parameter Estimates

The table shows GMM estimates of the preference parameters. EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. HAC standard errors are in parentheses. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The table also reports the cross-sectional R^2, the Wald tests for time-separability ($\gamma = \rho$) and ambiguity neutrality ($\eta = \gamma$), and the J-test for overidentifying restrictions (p-values in parentheses). Details on the tests are provided in Appendix A.

	In-Sample				Out-of-Sample			
ρ	Size	BM	Industry	All	Reversal	D/P	E/P	CF/P
Panel 1: EZ								
0.400	7.53	19.38	20.57	16.36	21.24	32.28	28.85	25.46
0.500	7.51	19.54	20.36	16.34	20.86	32.12	28.91	25.72
0.667	7.50	19.85	20.06	16.36	20.23	31.86	29.04	26.18
Panel 2: SA								
0.400	6.56	17.77	20.11	15.41	15.24	24.89	26.71	24.68
0.500	6.50	17.83	20.08	15.41	15.35	25.12	26.81	24.74
0.667	6.61	18.26	20.04	15.58	15.90	26.21	27.27	25.08

Table 7: Pricing Errors

The table reports relative cross-sectional pricing errors (in percent), which are computed by taking the square-roots of the mean squared differences between realized and predicted returns divided by the square-roots of the mean squared returns. EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. The construction of the individual portfolios is described in Appendix C. "All" contains 10 size, 10 book-to-market, and 10 industry sorted portfolios.
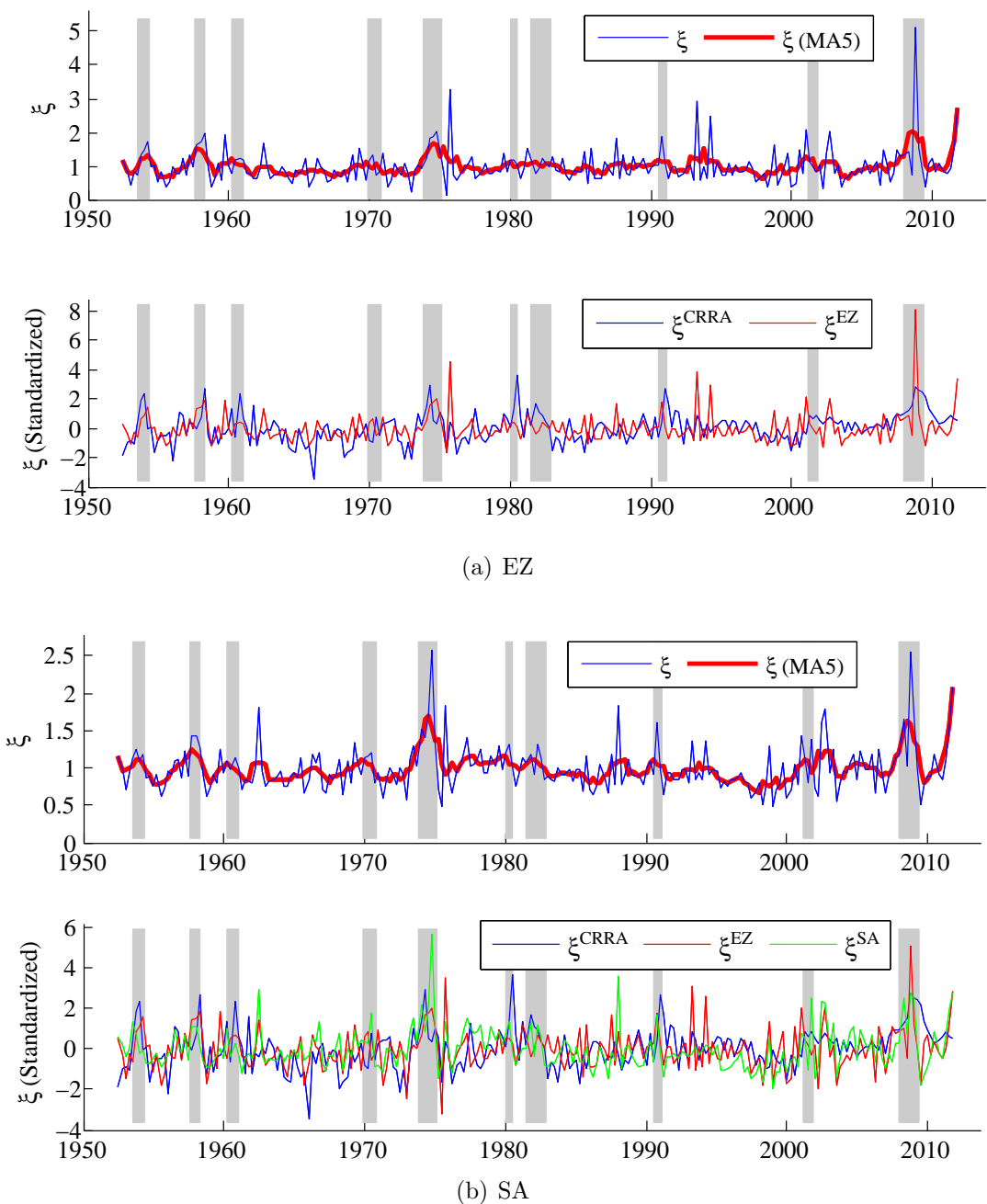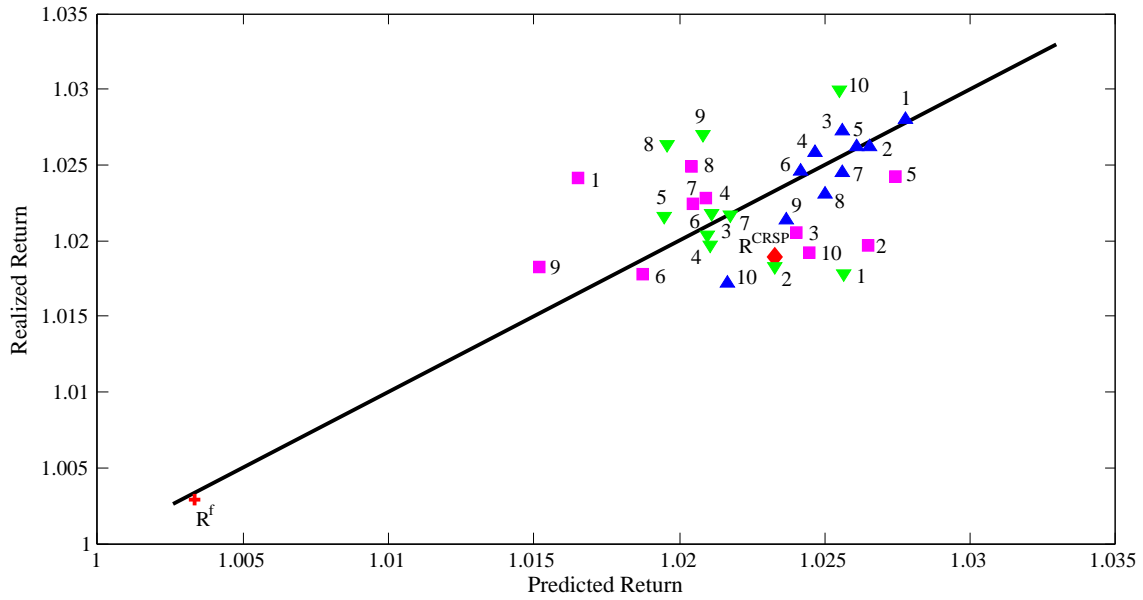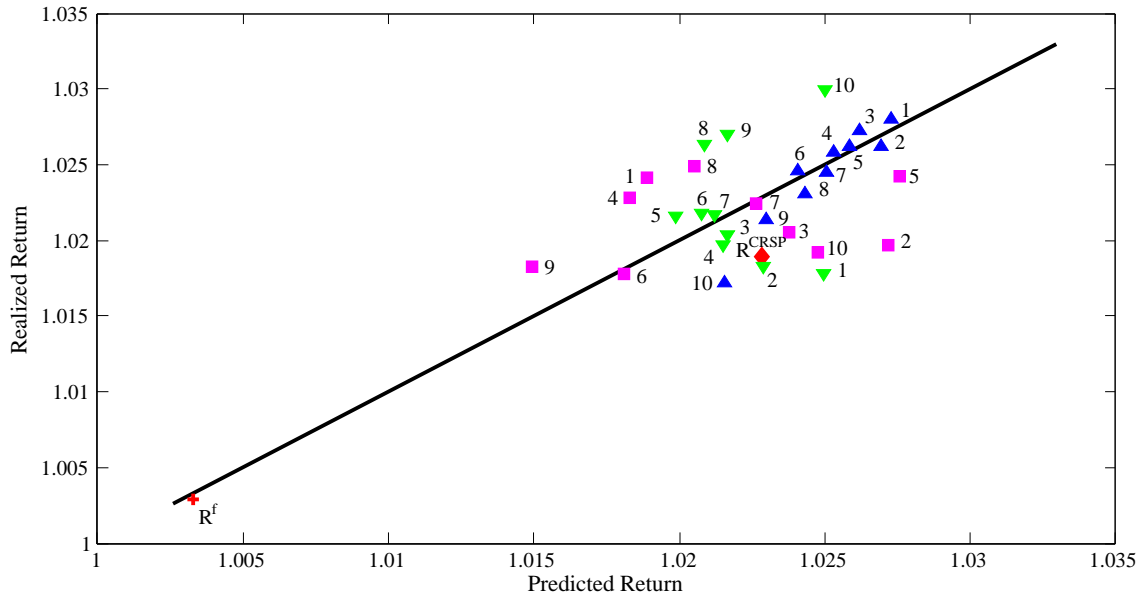
Figure 1: Realized Pricing Kernels

The figure shows time-series of realized pricing kernels and their five-quarter moving averages. It also displays the standardized components of the decomposition in Equation (7). EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. The EIS is set to 2. The shaded areas represent NBER recessions.

Figure 2: Realized versus Predicted Returns

The figure shows realized versus predicted returns on the 3-month Treasury bill (+), the CRSP market return ($\diamond$), and portfolios formed on size ($\triangle$), book-to-market ($\triangledown$), and industry ($\square$). EZ refers to Epstein and Zin (1989) preferences, SA to the smooth ambiguity model. The EIS is set to 2.