# Safety First, Learning Under Ambiguity, and the Cross-Section of Stock Returns

**Ariel M. Viale**
Florida Atlantic University, College of Business

**Luis Garcia-Feijoo**
Florida Atlantic University, College of Business

**Antoine Giannetti**
Florida Atlantic University, College of Business

---

We are especially grateful to the editor, Wayne Ferson, and an anonymous referee for their extensive comments that substantially improved the article. We also appreciate the valuable comments of Damian Damianov, Andrei Nikoforov, Jaime Casassus, and participants at the 2011 regional meetings of the Econometric Society in Santiago, Chile, the 2010 FMA meetings in New York City, and seminar at UTPA, Texas. Send correspondence to Ariel M. Viale, Florida Atlantic University, Department of Finance, College of Business, 777 Glades Road, Boca Raton, FL 33431; telephone: (561) 297-2914. E-mail: aviale@fau.edu.

**Abstract**

We examine the empirical implications of learning under ambiguity for the cross-section of stock returns. We introduce a theoretically-motivated ambiguity measure and find that ambiguity is priced in the cross-section of average stock returns. Ambiguity is not subsumed by state variables known to predict stock returns, nor by value, size, and momentum factors. In R-squared comparative tests, a model that takes ambiguity into account performs better than empirical implementations of the Bayesian learning model, the Intertemporal CAPM, and the four-factor model of Fama and French (1993) and Carhart (1997).

***JEL Classification:*** G12.
***Keywords:*** Dynamic Asset Pricing; Ambiguity; Knightian uncertainty; Learning; Cross-section of expected returns.

Conditional asset pricing models based on rational expectations perform poorly empirically (Lewellen and Nagel 2006; Kan, Robotti, and Shanken 2013). While the rational expectations hypothesis assumes investors know the probability law governing asset returns, authors such as Knight (1921), Keynes (1921), Shackle (1949), and Roy (1952) have emphasized that investors form expectations based on vague information that cannot be quantified precisely. Related evidence from experimental studies (Ellsberg 1961) has confirmed that individuals are averse not only to uncertainty regarding the outcome of events with known probabilities (risk), but also uncertainty regarding the outcome of events with unknown probabilities (Knightian uncertainty or ambiguity).[1] Consequently, a body of literature has emerged that examines the implications of ambiguity for portfolio selection and asset pricing.[2] So far, most of the literature has been theoretical, partly due to the difficulty in measuring ambiguity empirically.

We investigate the asset pricing implications of ambiguity regarding the state of the economy. Our goal is two-fold. First, we investigate whether ambiguity is priced in the cross-section of expected stock returns. To this end, we derive a conditional asset pricing model assuming investors do not know the process driving the evolution of the investment opportunity set. Second, we examine the performance of the model compared to the Bayesian learning model of Ozoguz (2009), the intertemporal capital asset pricing (ICAPM) model of Petkova (2006), and the four-factor model of Fama-French (1993) and Carhart (1997).

---

[1] Following the literature, we use Knightian uncertainty and ambiguity interchangeably. The use of the word ambiguity helps distinguish Knightian uncertainty from the common use of uncertainty as risk.

[2] A partial list of the growing asset pricing literature includes articles by Epstein and Wang (1994), Kogan and Wang (2003), Chen and Epstein (2002), Epstein and Schneider (2007, 2008), Leippold, Trojani, and Vanini (2008), Hansen (2008), Hansen and Sargent (2008, 2010, 2011), and Anderson, Ghysels, and Juergens (2009). A partial list of portfolio selection articles includes Uppal and Wang (2003), Maenhout (2004), and Garlappi, Uppal, and Wang (2007). Recent surveys include Epstein and Schneider (2010) and Guidolin and Rinaldi (2013).

The theoretical motivation for our investigation comes from the recent literature on learning under ambiguity. In this literature, investors seek to simultaneously learn the hidden state of the time-varying investment opportunity set, and the model driving the evolution of the hidden state. Investors' beliefs are represented by a set of conditional distributions across future states of the world. As it is the case in the learning literature, investors use a signal to discover the hidden state but, under ambiguity, investors evaluate the signal under the likelihood that puts the lowest possible conditional probability of being in the good economic state in the next period. That is, ambiguity averse investors always discount bad news more heavily than good news, which leads to an ambiguity premium that is time-varying and dependent upon the worst consumption state.[3]

Our particular theoretical framework is a specialization of the general *recursive multiple-priors* approach of Epstein and Wang (1994) and Epstein and Schneider (2003). We consider the investor consumption-investment problem when there is ambiguity about the probability of a shift in the economy from the "good" to the "bad" state. The solution to this problem leads to a fundamental pricing equation in which investors form expectations relative to the worst case probability measure. Empirically, we estimate the pricing equation as a conditional asset pricing model that includes three-factors: the market portfolio, and two additional factors that reflect the impact of systematic ambiguity on asset returns.

A major contribution of this paper is the introduction of an ambiguity measure that is theoretically motivated and empirically tractable. Anderson, Ghysels, and Juergens (2009) measure ambiguity using the dispersion of market return forecasts constructed from professional

[3] The literature distinguishes between cases in which the source of uncertainty is the signal investors use to learn the hidden state and the data generating process ("structured" uncertainty), and cases in which the source of uncertainty cannot be specified ("unstructured" uncertainty). Taking learning into account is important as Barillas, Hansen, and Sargent (2009) indicate that under unstructured uncertainty, ambiguity averse beliefs are equivalent to standard Bayesian beliefs with augmented risk aversion.

forecasts of aggregate corporate profits. Williams (2009) uses changes in the VIX index to measure changes in ambiguity. Our measure is based on the intuition in Epstein and Schneider (2008) that when investors are concerned with model misspecification, they 1) simultaneously consider multiple likelihoods for the "good" state of the economy, 2) form an interval of plausible values for the a priori imprecise probability of good economic times, and 3) select the minimum value within the interval.

The construction of our ambiguity measure involves three steps. First, we make an assumption about the investor's ex-ante minimum confidence level for his forecast of the state of the economy in the next period. Second, we estimate a regime-switching model to obtain the "reference" transition probability driving the dynamics of the time-varying investment opportunity set. Finally, we compute the asymptotic interval around the regime-switching forecast given the investor's minimum confidence level. Our ambiguity measure, denoted as *KUNC*, is equal to the difference between the reference probability estimate and the lower bound probability estimate (the "distorted" probability), scaled by the reference probability. As we explain in a later section, *KUNC* is the first order approximation of the log-likelihood ratio between the reference probability and the distorted probability and displays properties desirable in an empirical measure of ambiguity.

Using our proposed ambiguity measure, we find that ambiguity is priced in the cross-section of stock returns and constitutes an extra dimension of systematic uncertainty distinct from time-varying systematic risk. This result holds for the long sample period from 1927 to 2007, as well as for the post-COMPUSTAT period 1962-2007, and it is robust to errors-in-variables and to model misspecification. The test assets are the Fama-French 25 portfolios sorted by size and book-to-market or the augmented set that also includes 30 industry portfolios.

We also contrast our approach with the Bayesian approach to model uncertainty. Zhang (2003) and Ozoguz (2009) propose a conditional capital asset pricing model to test the Bayesian learning model of Veronesi (1999). Empirically, both Zhang (2003) and Ozoguz (2009) find support for the models and report that the Fama-French factors become insignificant after controlling for (Bayesian) model uncertainty. In contrast, we find that the Bayesian proxy for uncertainty is not significant once model misspecification is taken into account. Furthermore, using R-squared comparative tests (Kan, Robotti, and Shanken, 2013), we find that the Bayesian proxy for uncertainty has no incremental explanatory power relative to the learning under ambiguity asset pricing model.

In additional tests, we find that the learning under ambiguity model: 1) survives model misspecification, unlike the empirical Bayesian learning model and the ICAPM of Petkova (2006), and 2) performs better than the four-factor model of Fama-French (1993) and Carhart (1997) to explain the cross-section of average stock returns.

Before proceeding, we should note that Roy's (1952) principle of safety-first provides further motivation and meaning to our investigation.[4] Roy introduces his safety-first principle in the context of a portfolio problem under ambiguity. The principle of safety-first assumes that investors seek to minimize the effects of the worst case scenario. Thus, investors solve a *MaxMin* portfolio problem maximizing utility under the lowest plausible consumption state. This same argument is the main intuition of the ambiguity literature in economics and finance. Consequently, we adopt the safety-first principle as a plausible explanation for investors' conservative behavior when facing ambiguity in financial markets.

---

[4] Lintner (1965, footnote 21, page 243) acknowledges the pioneering work of Roy in asset pricing under Knightian uncertainty.

The rest of the paper proceeds as follows. In Section 1, we derive the fundamental dynamic asset pricing equation in beta regression form under ambiguity and learning. In Section 2, we introduce an empirically tractable ambiguity measure and compare it with the Bayesian proxy for uncertainty. In Section 3, we present the data and empirical methodology used in the empirical analyses. Section 4 investigates whether ambiguity is priced, while Section 5 uses robust comparative tests to evaluate model performance. We discuss the results of robustness checks in Section 6 and provide our conclusions in Section 7. Technical details and proofs are collected in appendices.

## 1. The Consumption-Investment Problem of the Ambiguity-Averse Investor

In this section, we first revisit the investor consumption-investment problem when there is ambiguity concerning the probability of a shift in the economy from the "good" to the "bad" state. The solution to this problem leads to a fundamental pricing equation in which investors form expectations relative to the worst case probability measure. Our theoretical framework is a specialization of the general *recursive multiple-priors* approach[5] introduced by Epstein and Wang (1994) and formalized by Epstein and Schneider (2003). Epstein and Schneider (2010) survey the related literature.

Additionally, we link the investor problem and the multiple-priors set to a classic result in statistics by Goodman (1965). This link explains the construction of the ambiguity measure *KUNC* that we use in the empirical analysis. Finally, we present the fundamental pricing equation in beta regression form. To obtain a testable unconditional model, we follow the

[5] A major advantage of the *recursive multiple-priors* modeling approach is that it is designed to ensure consistency in dynamic programming: plans made at date $t$ for decisions at subsequent dates remain optimal when those dates arrive.

conditional asset pricing literature and impose a linear factor restriction on the stochastic discount factor; the result is a three-factor model of learning under ambiguity.

### 1.1 The fundamental asset pricing equation

We assume that the economy evolves switching between the good state, one, and the bad state, zero. For each path $\omega_t$ (a random sequence of 0 and 1 moves), the investor chooses his level of consumption $C_t$, as well as the proportions $\xi_{i,t}$ of his wealth $W_t$, to allocate to $i$ risky assets ($i = 1, \ldots, n$ ) and to the risk-free asset. At time $t+1$, the investor's conditional wealth is $W_{t+1}^s$, where $s_t = \{0,1\}$ represents the bad (and, respectively, good) transition of the economy from prior state $\omega_t$.

The investor is ambiguous about the true *one-step-ahead* conditional probabilities that he should attach to the good and bad state. [6] He holds some prior beliefs $\pi_t(\omega_t)$, possibly derived from historical analysis, which we refer to as his *reference model*. However, he also acknowledges that this reference model $\pi_t(\omega_t)$ may be surrounded by a cloud of *alternative models* $\pi_t^*(\omega_t)$, also known in the literature as the multiple-priors set. The investor only considers alternative models that are close to the reference model. Thus, we restrict the statistical "distance" between any alternative model $\pi_t^*$ and the reference model $\pi_t$ by imposing:

$$D_t(\pi_t^* \| \pi_t) \leq \eta_t, \tag{1}$$

[6] For dynamic consistency, it is critical that the one-step-ahead conditional probability be considered. Epstein and Schneider (2003) pp. 16-17 discuss the technical requirements needed to ensure dynamic consistency in the Entropy framework. We thank the referee for bringing this to our attention. We simplify notation by setting $\pi_t^*(\omega_t) \equiv \pi_t^*$ and $\pi_t(\omega_t) \equiv \pi_t$.

where $D_t(\pi_t^* \| \pi_t)$ denotes the Kullback-Leibler divergence or "Relative Entropy" and $\eta_t$ is an exogenous state-dependent ambiguity parameter that restricts what Epstein and Schneider (2010) call the *entropy-constrained ball* containing all alternative measures $\pi_t^*$ that are statistically close to the reference model $\pi_t$.[7] Intuitively, the ambiguity coefficient $\eta_t$ is related to the level of confidence that the investor places in his model. This is the interpretation that underlies the construction of the ambiguity measure *KUNC* we use in our empirical analysis.

Next, the ambiguity averse investor solves the following version under ambiguity of the standard Bellman backward dynamic recursion subject to the usual budget constraint:[8]

$$J(W_t, t) = Max_{C_{t,\{\xi_{i,t}\}}}\left(U(C_t) + Min_{\pi_t^*} E_t[\pi_t^* J(W_{t+1}^1, t+1) + (1-\pi_t^*)J(W_{t+1}^0, t+1)]\right), \quad (2)$$

where $E_t(.)$ is the conditional expectation operator at time $t = (0,1,..,T-1)$, $U(C_t)$ is increasing and concave, and $J(W_t, t)$ denotes the usual derived utility of wealth function.

The economic interpretation of the above dynamic *MaxMin* problem is as follows. The ambiguity averse investor first solves an inner constrained minimization problem to identify the worst case scenario among all of the alternative models $\pi_t^*$. The solution is a conditional measure that intuitively allows the investor to calculate the *ambiguity certainty equivalent* of the continuation value function $J(W_{t+1}^s, t+1)$. We refer to this conditional measure as the *distorted probability* $\pi_t^L$. The constrained minimization problem reflects the notion that while the ambiguity averse investor attempts to determine the worst case scenario by taking the most

7 The Kullback-Leibler divergence arises in statistics as the expected logarithm of the likelihood ratio between two distributions (see Cover and Thomas, 1991, Chapter 2). It is not a true distance as it is not symmetric and does not satisfy the triangular inequality.

8 Details are noted in Appendix A.

pessimistic view of the transition probabilities, he must respect the entropy constraint specified by (1).

We emphasize that our use of Relative Entropy is not an arbitrary formalism. In the field of information theory, Topsøe (1979) considers a density estimation game between two players: the malevolent "Nature" and a decision maker. Nature is allowed to choose any distribution $\pi_t^*$ that satisfies a set of constraints. The decision maker only knows of the constraints, but not the distribution chosen by Nature. As such, his best strategy consists of selecting the distribution $\pi_t^L$ that maximizes the worst log likelihood with respect to his reference model $\pi_t$. Topsøe (1979) determines that this distribution is also the one that minimizes the relative entropy $D_t(\pi_t^* \| \pi_t)$ under constraints. We will return to this interpretation when we introduce our empirical proxy for ambiguity.

Once the worst case scenario measure $\pi_t^L$ has been identified, the investor proceeds to solve the usual outer utility maximization problem, albeit under the distorted probability measure $\pi_t^L$. Solving the *MaxMin* problem yields the fundamental pricing equation under ambiguity. Proposition 1 states the solution to the *MaxMin* problem (2). Appendix A includes the proof.

**Proposition 1:** The fundamental asset pricing equation under ambiguity is given by:

$$1 = E_t^{\pi^L}\left(m_{t,t+1} R_{i,t+1}\right), \tag{3}$$

where $R_{i,t+1}$denote the gross return of asset $i$ and $\mathrm{m_{t,t+1}} \equiv \frac{\mathrm{U_C(C_{t+1}^*,t+1)}}{\mathrm{U_C(C_t^*,t)}}$ is the marginal rate of substitution under the worst case scenario probability $\pi^{\mathrm{L}}$ with $\pi_{\mathrm{t}}^{\mathrm{L}} = \pi_{\mathrm{t}} - \sqrt{2\eta_{\mathrm{t}}\pi_{\mathrm{t}}(1-\pi_{\mathrm{t}})}$.

According to Proposition 1, the ambiguity averse investor forms a conditional expectation $E_t^{\pi^L}$ across future states of the economy using the state probabilities associated with the worst case scenario probability measure $\pi_t^L$. In contrast, a Bayesian investor would first estimate the conditional probability $\pi_t$ of being in the good state, and then evaluate the conditional expectation under this very same measure. For a Bayesian investor, there is no ambiguity; hence, $\eta_t = 0$ resulting in $\pi_t^L = \pi_t$.

To clarify, investors must first solve a signal extraction problem using past data to estimate the conditional probability of the good state $\pi_t$. This step is common to both the Bayesian and ambiguity literature with learning.[9] Empirically, the Bayesian literature typically uses variants of the Markov regime-switching model of Hamilton (1989, 1990) for the hidden state of the economy (e.g., Ozoguz, 2009). However, in the ambiguity literature, which puts the investor and the econometrician on the same footing, the regime-switching model (RSM) is only a reference model. The separation of the step of solving the signal extraction problem using a reference model and the step of using the estimated signal as an input into the solution to the investor problem is common in the literature.

### 1.2 Recovering the worst case scenario probability measure

As described previously, the ambiguity averse investor solves the minimization problem before maximizing utility. To solve the minimization problem, the ambiguity averse investor acts as if he follows a three-step process. First, he selects, ex-ante, a minimum confidence level for his forecast of the state of the economy in the next period, $\hat{\pi}_t$. The minimum confidence level determines how big the "cloud" of alternative models is. Next, he simultaneously considers

[9] See, for example, Veronesi (1999), David and Veronesi (2001), Ozoguz, (2009), Hansen and Sargent (2010, 2011) and Epstein and Schneider (2008).

several possible empirical likelihoods of the state of the economy $\pi_t^*$. Additionally, he distorts the conditional expectation for the good state of the economy $\hat{\pi}_t$ conservatively (i.e., slanting his beliefs toward the bad economic state, $\pi_t^L$). The uniform interval of possible values for the state of the economy is time-varying conditional upon the realization of the signal.

More formally, investors have an irreducible set of priors $\Pi_0$ for the objective transition probability $\pi_t$ of a good state represented by the uniform interval $[\pi_t^L, \pi_t^H]$ constructed around the quasi-maximum likelihood estimate (Q-MLE) $\hat{\pi}_t$ obtained from the reference model. We define $(1-\varsigma)$ as the investors' ex-ante minimum confidence level on the Q-MLE forecast $\hat{\pi}_t$. Thus, we interpret $(1-\varsigma)$ as an inverse proxy for the theoretical coefficient $\eta_t$; an increase in $\varsigma$ implies lesser confidence in the reference model.[10] In Appendix B, we show that minimization of the Kullback-Leibler divergence (i.e., the expected log ratio) is equivalent to a statistical result in Goodman (1965) with solution given by a quadratic equation with the following roots:

**Proposition 2:** Let $\hat{\pi}_t$ be the subjective Q-MLE of the probability of transitioning to the good economic state; then, given some minimum confidence level $100(1-\varsigma)\%$ on $\hat{\pi}_t$, investors will entertain a uniform interval of alternative possible values for the objective probability defined by:

$$[\pi_t^L, \pi_t^H] = \left[\frac{B-\Delta^{\frac{1}{2}}}{2A}, \frac{B+\Delta^{\frac{1}{2}}}{2A}\right], \tag{4}$$

[10] The parameter α in Epstein and Schneider (2007, p. 1294) may also be interpreted as a confidence interval related to the interval within which a theory would be excluded from the set of multiple beliefs. Kogan and Wang (2003) and Garlappi, Uppal, and Wang (2007) also characterize the level of ambiguity in terms of a “conservative” confidence interval.

where $A = \mathcal{X}^2(1-\varsigma/I,1)+T$ denotes the quantile of order $1-\varsigma/I$ of a chi-squared distribution with one degree of freedom with I number of intervals and sample size $T = \sum_{i=1}^{I} t_i$; $B = \mathcal{X}^2(1-\varsigma/I,1)+2t_i$ ; and $\Delta = B^2 - 4AC$ with $C = t_{\{\omega=g\}}^2/T$.

Proposition 2 provides a feasible approach to empirically estimate the distorted probability $\pi_t^L$. The approach involves i) forming a uniform interval around the reference model forecast and ii) selecting the lower bound. This is the approach we take to measure ambiguity empirically, as further explained in Section 2.

**1.3 Model specification in beta regression form**

While Proposition 1 delivers an intuitive asset pricing equation under ambiguity, it does not place enough restrictions on the marginal rate of substitution $m_{t-1,t}$ for $t = (1,2,..,T)$ to proceed empirically. Therefore, we follow the conditional asset pricing literature and impose an affine factor structure on $m_{t-1,t}$ before "conditioning down" using the worst case scenario probability measure $\pi_t^L$.[11] More specifically, we assume:

$$m_{t-1,t} = \phi_{t-1}^0 + \phi_{t-1}^f{}' f_t,$$

where the column vector of fundamental asset pricing factors $f_t = \left(R_{MKT,t}^e, R_{d\pi_t^L}^e, R_{dKUNC_t}^e\right)$ includes the excess market (*MKT*) return and excess returns of tracking portfolios proxying for

[11] We incorporate conditioning information into the pricing kernel using "scaling factors". Ludvigson (2012), Section 5.2, discusses various approaches for conditioning in the empirical asset pricing literature. Following Ferson and Harvey (1999), we use excess returns as factors tracking orthogonalized innovations in order to address the "spurious regression" problem discussed in Ferson et al. (2003).

innovations in the distorted conditional probability ($\pi_t^L$) and the measure of ambiguity (*KUNC*). We follow Cochrane (1996, section 6.3) and Singleton (2006, section 11.4 p. 296) and assume that the time-varying coefficients $\tilde{\phi}_{t-1}' = \left(\phi_{t-1}^0, \phi_{t-1}^{f}{}'\right)$ are affine functions of innovations in state variables lagged one period $z_{t-1}$ driving the investment opportunity set:

$$\phi_{t-1}^0 = a^0 + b^0 dz_{t-1}, \text{ and}$$

$$\phi_{t-1}^f = a^f + b^f dz_{t-1}.$$

In our setting, innovations in the worst case scenario probability measure $dz_{t-1} = d\pi_{t-1}^L$ are used as the “scaling” variable to “condition down” the conditional multi factor model under ambiguity. Substitution into Equation (3) leads to the following unconditional moment conditions:

$$E\left[(a^0 + b^0 d\pi_{t-1}^L + a^f f_t + b^f d\pi_{t-1}^L f_t)R_{i,t}\right] = 1, \textit{for all risky assets } i = 1, \cdots, n.$$

When pricing factors are orthogonal and the scaling variable is white noise,[12] then, it is straightforward to show that the above moment conditions jointly applied to the risk-free asset as well as the three portfolios tracking the pricing factors[13] yield the following unconditional multifactor representation in beta regression form for all risky assets i $= 1, \cdots, n$:

[12] More specifically, we require $E(d\pi_{t-1}^L) = 0$ and $d\pi_{t-1}^L$ be independent of all factors $f_t$, hence $Cov\left(d\pi_{t-1}^L, \varphi(f_t)\right) = 0$ where $\varphi(f_t)$ denotes any measurable function of factor $f_t$.

[13] The previous result is intuitive since, under our assumptions, the three factors along with risk-free rate exactly span the expected return space of all risky assets.

$$E\left[R_{i,t+1}\right] = \lambda_0 + \beta_{i,MKT}\lambda_{MKT} + \beta_{i,\pi^L}\lambda_{\pi^L} + \beta_{i,KUNC}\lambda_{KUNC}, \tag{5}$$

where $E\left[R_{i,t+1}\right]$ is the expected return on the $i$th risky asset, the β's denote factor loadings defined as before, $\lambda_0 = E\left[R_{f,t-1}\right]$ is the constant return of the zero beta portfolio (excess returns are relative to the risk free rate), and the rest of the λ's are the market prices of the factors measuring the three dimensions of systematic uncertainty: 1) systematic risk, 2) uncertainty regarding the state of the economy (i.e., learning under ambiguity), and 3) uncertainty regarding the data generating process driving stock returns (i.e., the direct effect of ambiguity on the equity premium).

The learning under ambiguity model described by Equation (5) predicts that an ambiguity averse investor will demand an additional premium to hold stocks that are heavily affected by negative news from an ambiguous signal, independent of the state of the economy. Furthermore, the investor prefers stocks whose payoffs co-vary negatively with his level of confidence in the reference model. Thus, he will require higher compensation to hold stocks that perform relatively worse when his confidence in the model is lower. A major goal of this paper is to provide empirical evidence regarding these predictions.

We highlight the following special cases of Equation (5). When the signal is not ambiguous (but there is learning), the scaling variable is the reference model $\hat{\pi}_t$. Equation (5) becomes the Bayesian C-CAPM of Veronesi (1999), Zhang (2003), and Ozoguz (2009). We empirically compare the Bayesian and ambiguity approaches to uncertainty in Section 4.3.

Further, if we replace the distorted probability $\pi_t^L$ for economic-wide state variables driving $\hat{\pi}_t$, Equation (5) becomes a version under ambiguity of Petkova's (2006) implementation of the ICAPM of Merton (1973). This empirical specification can be understood as a modified version of the learning model of Epstein and Schneider (2007, 2008) when the investor uses news on a set of ambiguous macroeconomic state variables as the signal. We provide empirical evidence on this implementation in Section 4.4.

In the next section, we describe the empirical version of the reference model $\hat{\pi}_t$, and introduce our proxies for the learning and ambiguity factors in (5).

## 2. Measuring Ambiguity

### 2.1 Estimating the conditional probability of the good state of the economy

We adopt the Smooth Transition Autoregressive (STAR) model to measure the investor's beliefs regarding the future state of the economy (i.e., the reference model $\hat{\pi}_t$). A STAR model is a type of regime switching model particularly appropriate when the transition variable is observable and forward looking, switching between states smoothly and endogenously, given some threshold values for the good/bad state.[14] Our adoption of a STAR model is consistent with the assumption that the investor-econometrician "learns" the hidden state of the economy using stock market returns as signals of the future state of the economy.

The standard Bayesian setting assumes investors know the data generating process but not the future state of the economy (i.e., the transition variable is unknown and driven by a hidden Markov process). In this setting, the Markov switching model of Hamilton (1989, 1990) is the natural choice to model state dependent risk and expected returns. For example, in Ozoguz's (2009) empirical implementation of Veronesi's (1999) learning model, the economy

[14] STAR models have been used to describe the asymmetric behavior of macroeconomic time series, such as output and unemployment, conditional on the phase of the business cycle. STAR models are discussed in detail in Teräsvirta (1998) and surveyed in Granger and Teräsvirta (1993) and van Dijk, Teräsvirta, and Franses (2002).

switches discretely between good and bad states driven by a hidden Markov process. The STAR model, however, is more appropriate in our case because we assume the investor learns the future state of the economy using an observable forward looking transition variable that switches smoothly between states.

More formally, we model the market return as a self-exciting ARMA(1,1)-GARCH(1,1) dynamic stochastic process with smooth "double" transition function in the interval $[0,1]$. Regime 1 denotes the good state, while Regime 0 represents the bad state. The transition function is defined as a continuous logistic function defined in the interval bounded between zero and one:

$$G(\omega_t;\ \psi, \boldsymbol{\delta}) = \frac{1}{1+exp\{-\psi(R^e_{MKT,t-1}-\delta_{1,0}-\boldsymbol{\delta}'_{\mathbf{1,1}}dz_{t-1})(R^e_{MKT,t-1}-\delta_{0,0}-\boldsymbol{\delta}'_{\mathbf{0,1}}dz_{t-1})\}}, \tag{6}$$

where $\psi \geq 0$ is the smoothness parameter, and $dz_{t-1}$ denotes innovations in a parsimonious group of state variables $z_t = [DIV_t, RF_t, TERM_t, DEF_t]'$ from the return predictability literature (defined in the next section).

Consistent with the notion that stock returns are forward looking, the switching indicator variable defining the state $\omega_t$ is the one-period lagged market return $R^e_{MKT,t-1}$. This specification is the self-exciting (SETAR) model, which is a special case in the class of STAR models discussed in detail by Tong (1990). When $\psi \to 0$, the logistic function approaches a constant equal to one-half, and when $\psi = 0$, the STAR model collapses to a linear ARMA process. The double transition function with a state dependent vector of coefficients $\boldsymbol{\delta}$ defines an interval of values for $r^e_{MKT,t-1}$, where the model is predominantly in Regime 1; or in Regime 0 otherwise.

The self-exciting smooth transition GARCH (ST-GARCH) model for the excess market return $R^e_{MKT,t}$, which is observed by investors at $t = 0,1,\cdots,T-1,T$ is given by:

$$R^e_{MKT,t} = \left(\phi_{1,0} + \phi_1 R^e_{MKT.t-1}\right)\left(G(\omega_t;\psi,\boldsymbol{\delta})\right) + \left(\phi_{0,0} + \phi_1 R^e_{MKT.t-1}\right)\left(1 - G(\omega_t;\psi,\boldsymbol{\delta})\right) + \cdots$$
$$\cdots + (\varepsilon_t - \zeta\varepsilon_{t-1}), \tag{7}$$

where $\phi_{1,0}, \phi_{0,0}$ are state dependent intercepts, $\phi_1$ and $\zeta$ are the autoregressive and moving average parameters of the data generating process, respectively, and $\varepsilon_t = \sqrt{\hbar_t}v_t$ with $\hbar_t = \left(G(\omega_t;\psi,\boldsymbol{\delta})\right)\hbar_{1,t} + \left(1 - G(\omega_t;\psi,\boldsymbol{\delta})\right)\hbar_{0,t}$, $\hbar_{1,t} = a_{1,0} + a_1\varepsilon^2_{t-1} + b\hbar_{1,t-1}$, $\hbar_{0,t} = a_{0,0} + a_1\varepsilon^2_{t-1} + b\hbar_{0,t-1}$, with $a_0$ as the only state dependent parameter in the variance return equation.

We emphasize that under ambiguity, the STAR model is only a reference model. The next section connects the reference model, the worst-case probability, and the ambiguity measure.

### 2.2 A theoretically-motivated and empirically tractable ambiguity measure

The learning under ambiguity model described by Equation (5) predicts there are three priced dimensions of uncertainty: 1) systematic risk, 2) uncertainty regarding the state of the economy (i.e., learning under ambiguity), and 3) uncertainty regarding the return generating process (i.e., ambiguity or Knightian uncertainty). In this section, we introduce our empirical proxies for the two dimensions of uncertainty related to ambiguity.

Our proxy for learning under ambiguity is the distorted probability $\pi^L_t$. Asset pricing models that include learning often lead to a "learning factor" that is measured by the probability forecast of the good economic state, such as the one derived from a regime switching model (i.e.,

$\hat{\pi}_t$). However, when investors are ambiguity averse, they will doubt their forecast, consider a set of plausible alternative forecasts and associated likelihoods, and assume the most pessimistic forecast to make investment decisions. Therefore, we use the distorted probability rather than the reference forecast as our measure of ambiguity regarding the future state of the economy.

Our proxy for ambiguity regarding the stock return data generating process, denoted by *KUNC* (for Knightian uncertainty), is the first order approximation of the log-likelihood ratio between the reference probability and the distorted probability:

$$KUNC = -\ln\left(\frac{\pi_t^L}{\hat{\pi}_t}\right) \approx (\hat{\pi}_t - \pi_t^L)/\hat{\pi}_t. \tag{8}$$

This ambiguity measure has both intuitive appeal and desirable technical properties. Intuitively, the distance between the reference and distorted probabilities is a reasonable measure of investor's uncertainty about the true model for $\pi_t$.[15] Indeed, the minimization problem of the ambiguity averse investor may be equivalently stated as the problem of maximizing the expected negative of the log likelihood ratio between the worst case scenario and the reference model, $E_t\left[-\ln\left(\frac{\pi_t^L}{\hat{\pi}_t}\right)\right]$.

In Table 1, we present summary statistics for the two empirical measures associated with ambiguity, $\pi_t^L$ and $KUNC$. For comparative purposes, we also include the standard Bayesian uncertainty index defined as $UC = \hat{\pi}_t(1-\hat{\pi}_t)$ (Ozoguz, 2009). We plot the two uncertainty measures in Figure 1, for different values of $\hat{\pi}_t$ shown in the horizontal axis. Our proposed

[15] In Appendix A, we present the exact connection between the proposed proxy and the investor's constrained minimization problem. In Appendix B, we demonstrate how the empirical proxy is motivated by the frequentist interpretation of entropy introduced by Topsøe (1979), and formalized in Harremoës and Topsøe (2001), Grünwald and Dawid (2004), and Topsøe (2007).

ambiguity measure displays some desirable properties, particularly when compared to the Bayesian uncertainty index *UC*. First, *KUNC* is asymmetric, leaning toward the recession state as suggested by theory (Epstein and Schneider, 2008). Ambiguity aversion implies that stocks should be discounted more in the "bad" economic state. Intuitively, ambiguity averse investors fear stocks not because stocks do poorly in recessions, but because investors seek safety first, especially during economic recessions. Additionally, *KUNC* is well behaved, taking values between zero and one.[16] In contrast, the Bayesian index is symmetric and achieves a maximum value of 0.25 when the probability $\hat{\pi}_t$ goes to 0.5. Further, Table 1 shows that *UC* displays significantly less variability than *KUNC* (Panel A); and that *UC* and *KUNC* are negatively correlated (Panel B).[17]

Insert Table 1 and Figure 1 about here.

We further illustrate the differences between *UC* and *KUNC* with two contrasting examples. Consider first an investor with a Q-MLE of only 0.86% for the transition probability to the good economic state next period. This may be typically the case when the economy is in the trough of the business cycle, such as during the recession in in the year 1990. The investor's probability estimate is obtained by fitting a sample of 1,000 monthly market returns (approximately the size of our data sample from 1927-2007) and the STAR reference model. If the investor has a minimum confidence level in his reference model of 50%, he will form an interval of possible values for the hidden objective transition probability between 0.31% and 2.41%. The ambiguity averse investor will then select the lower bound of 0.31% rather than the

[16] In Appendix B, we note that the distorted probability moves to zero faster than the reference probability as long as $\pi_t$ is non-degenerate. Empirically, we need to force the ratio to be bounded in about ten data points around the Great Depression.

[17] The negative association is expected based on Equation (B.3) in Appendix B.

estimate of 0.86% as the conditional probability of transitioning to the good economic state next period. The value of the ambiguity measure *KUNC* obtained using Equation (8) is equal to 0.6452. In contrast, the corresponding value of the Bayesian uncertainty index *UC* is equal to 0.0086. Uncertainty is higher as measured by *KUNC* because the ambiguity averse investor doubts his initial forecast and acts conservatively.

At the other extreme, if the Q-MLE of the transition probability to the good economic state next period is 96% (i.e., the economy is at the peak of the business cycle, such as in the late 1990s), the ambiguity averse investor will now form an interval of possible values for the hidden transition probability between 93% to 97%, selecting the lower bound of 93% rather than the estimate of 96% as the relevant probability of transitioning to the good state. The value of *KUNC* is 0.0312, while the value of *UC* is 0.0384.

The distinction between *KUNC* and *UC* is further illustrated in Figure 1. Panel A confirms the result of the numerical examples. While *UC* is symmetric, *KUNC* is asymmetric and more sensitive at lower levels of the probability of transitioning to a good state. In Panel B, we plot *KUNC* at two different confidence levels regarding the reference probability. As expected, *KUNC* is lower and less sensitive when the investor has more confidence in the reference model.

In Figure 2, we plot the time series behavior of *KUNC* measured as the return of a tracking portfolio as explained in the next section. The figure shows that *KUNC* rises significantly during NBER-dated economic recessions, as conjectured in the literature (e.g., Ju and Miao, 2012). Further, there is substantial variability in both sample periods, with an indication of a negative long run trend from 1927-1961. A possible explanation for the observed negative trend may be related to the impact of the Great Depression at the beginning of the sample period, and possibly the subsequent implementation of fiscal and monetary policies to smooth out the effects of the business cycle.

Insert Figure 2 About Here.

Overall, we conclude that our ambiguity measure is not only theoretically motivated but also displays characteristics that would be desirable in an empirical measure of ambiguity.

## 3. Empirical Methodology and Asset Pricing Tests

### 3.1 Data

**3.1.1 Test assets.** The set of test assets used in the empirical analysis includes the value-weighted monthly returns on the 25 Fama-French portfolios sorted by size and book-to-market obtained from Kenneth French's website, and the 30 industry-sorted portfolios (Lewellen, Nagel, and Shanken 2010). Results are presented for two sample periods, 1927-2007 and the post-COMPUSTAT period 1962-2007.[18]

**3.1.2 Asset pricing factors.** According to Equation (5), there are three asset pricing factors: 1) systematic risk, 2) uncertainty regarding the state of the economy, and 3) uncertainty regarding the return generating model. We measure systematic risk using the excess return $r^e_{MKT,t}$ on the Center for Research in Security Prices (CRSP) value-weighted stock market index. We calculate $\pi^L_t$ and *KUNC* as explained in the previous section.

[18] Kenneth French's website is at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

In Subsection 1.2, we formally present the connection between ambiguity and the investor's minimum confidence level in his reference model. Empirically, we focus on the case of an ambiguity averse investor that has a minimum confidence level of 50% in his model and a forecasting window of *T*=36 months in his time series predictive regressions (i.e., the rolling regressions). However, in robustness checks at the end of the paper, we present results for a range of confidence levels from a minimum of 50% to a maximum of 90%, and for forecasting windows from 36 months to 180 months

Campbell (1996) emphasizes that only the unexpected component of an asset pricing factor should command a premium. Accordingly, we measure the factors as orthogonal innovations from a vector autoregressive (VAR) system of equations. More specifically, we assume the demeaned vector of state variables $z_t = [r^e_{MKT.t}, \pi^L_t, KUNC]'$ follows a first order VAR: [19]

$$z_t = Az_{t-1} + dz_t. \quad (9)$$

The residuals in the vector $dz_t$ are the innovations that proxy for asset pricing factors, and $A$ is a matrix of fixed coefficients.

We also present results for an alternative empirical specification of Equation (5) that follows Petkova's (2006) empirical implementation of the ICAPM of Merton (1973). The alternative specification replaces the distorted probability $\pi^L_t$ as a scaling factor with the vector of innovations $dz_t$ from a set of macroeconomic variables used as instruments in the estimation of the reference model $\hat{\pi}_t$. The set of state variables includes the market dividend yield, *DIV*, which serves as a proxy for time varying expected stock returns (Campbell and Shiller 1988); the

[19] Campbell and Shiller (1988) find that any high order VAR can be collapsed to its first order (companion) VAR.

one-month T-bill yield, *RF*, that proxies for the level of interest rates (Fama and Schwert 1977); the difference between the average yield of a portfolio of long-term government bonds and the one-month T-bill, *TERM*, which measures the slope of the yield curve (Campbell 1987); and the difference between the average bond yield of a portfolio of long-term corporate bonds (Aaa/Baa) and a portfolio of long-term government bonds, *DEF*, as proxy for default risk (Fama and French 1989).

In the alternative specification, the first order VAR is given by the following modification to Equation (9):

$$z_t = [r^e_{MKT.t}, RF_t, DIV_t\ , TERM_t, DEF_t, KUNC_t]'. \tag{10}$$

The alternative specification serves two purposes. First, it provides evidence as to whether any possible explanatory power of *KUNC* is subsumed by the information contained in variables associated with the evolution of systematic risk. Additionally, the alternative specification is consistent with a version of the model in Epstein and Schneider (2008) in which investors use (ambiguous) macroeconomic news to learn the hidden state of the investment opportunity set. Thus, the results of the alternative specification provide additional evidence concerning the relevance of the model in Epstein and Schneider (2008) to explain stock returns.

Furthermore, we also investigate whether the explanatory power of *KUNC* is subsumed by the risk factors *SMB* and *HML* of Fama and French (1993), and the momentum risk factor *UMD* of Jegadeesh (1990) and Jegadeesh and Titman (1993). We measure *SMB*, *HML*, and *UMD* as innovations obtained from an augmented version of the VAR in the alternative implementation in Equation (10).

**3.1.3 Tracking portfolios.** According to Lamont (2001), returns from tracking portfolios may be used to measure risk premia and reveal which state variables are important determinants of expected returns. Additionally, tracking portfolios may be interpreted as hedging tools against unexpected changes in economic state variables. Ferson, Seigel, and Xu (2006) discuss the benefits of the approach suggested by Lamont (2001), compared to the standard mimicking portfolio approach, when the return generating process is not known. Lamont's (2001) economic tracking portfolios are different from mimicking portfolios and from the maximum correlation portfolios of Breeden, Gibbons, and Litzenberger (1989). In particular, Lamont's (2001) approach is designed to track future rather than current changes in economic state variables.

Asset pricing tests based on two-pass cross-sectional regressions do not require asset pricing factors to be tradable portfolios. However, the factors described in the previous subsection are based on variables of three very different types. The ambiguity proxies are not only not tradable, but also not observable. In contrast, the variables from the return predictability literature are observable and have an intuitive economic interpretation. The Fama-French factors are mimicking portfolios whose meaning is still controversial. Thus, we transform all factors into returns following the procedure in Lamont (2001).[20] Measuring the factors in the common return space has the added benefit that the GLS cross-sectional regression is identical to a time-series regression asset pricing test (Cochrane, 2001, page 244); hence, we can directly assess the economic significance of the CSR estimates without any further calculation.

[20] In robustness checks, we report results using the state variables as innovations, rather than tracking portfolios. We also report results based on the raw measures of the state variables.

Specifically, we construct time-series of (unexpected) portfolio returns tracking the orthogonalized innovations of the state variables $\pi_t^L$, *KUNC*, *DIV*, *RF*, *TERM*, and *DEF*. The procedure involves running the following OLS regression for each state variable:[21]

$$dz_{t+12} = \boldsymbol{bR}_{t-1,t} + \boldsymbol{cY}_{t-1} + e_{t,t+12}, \tag{11}$$

where $dz_{t+12}$ is the 12-month ahead realized future value of orthogonalized innovations of each economic state variable, $\boldsymbol{R}_{t-1,t}$ is a vector of period *t* excess returns of 30 industry portfolios, and $\boldsymbol{Y}_{t-1}$ is a vector of control variables that includes the period *t*-1 inflation, the market excess return, and the innovations in the state variables lagged one period.

Alternatively, we also run the OLS regressions using *SMB*, *HML*, and *UMD* portfolios instead of the 30 industry portfolios. The results are quantitatively and qualitatively similar.[22] To save space, we report the results for tracking portfolios based on *SMB*, *HML*, and *UMD*, which is the more parsimonious approach. Lamont (2001) notes that adding variables to $\boldsymbol{R}_{t-1,t}$ and $\boldsymbol{Y}_{t-1}$ may lead to overfitting and spurious inferences.

## 4. Cross-Sectional Regression Estimates and Asset Pricing Test Results

### 4.1 The reference STAR model

In Table 2, we report the estimates of the STAR model for the excess market return. Under ambiguity, the STAR model is only a reference model. In Panel A, we present the estimates

[21] By using orthogonalized innovations, we address the potential problem of multicollinearity induced by construction of the beta regression form of the conditional CAPM.

[22] The fact that results are similar is not surprising. By construction, tracking portfolio returns have minimum variance, maximum correlation with the innovations in each economic state variable, and the highest $R^2$ among any other alternative portfolio constructed using univariate OLS regressions (Lamont, 2001).

corresponding to the stock market return Equation (7) and the corresponding variance return equation. In Panel B, we report the vector of estimates $\hat{\delta}$ in the logistic function $G$, Equation (6) that may be interpreted as the time-varying transition probability of the good state as a function of the innovations in the vector of economic state variables $[DIV_t, RF_t, TERM_t, DEF_t]$. The estimates in Panel A are consistent with well-known features of stock returns documented in the empirical asset pricing literature.[23]

Insert Table 2 About Here.

The results in Panel B confirm the expected theoretical associations between the innovation in the parsimonious set of state variables and the state transition probability for the market. The results are also consistent with previously reported results in the asset pricing literature (Petkova and Zhang 2005; Petkova 2006). News regarding *DEF* and *TERM* in the good economic state and *RF*, *DIV*, and *TERM* in the bad economic state are statistically significant at a 5% level. The standard interpretation of the *TERM* spread is that it indicates that expected market returns are low during expansions and high during recessions (Fama and French 1989). Thus, positive innovations in *TERM* are associated with bad news about the economy, which explains the negative estimated coefficient in Panel B. This is consistent with the literature that explores the correlation between the slope of the yield curve and future economic growth (Chen 1991; Estrella and Hardouvelis 1991). Bad economic times are preceded by a flattening of the yield curve, hence the higher expected market return.

[23] Researchers have found that stock returns are persistent and subject to low frequency jumps in their drift and/or volatility, which can generate fat-tailed and skewed marginal distributions (for a survey of this literature see Singleton 2006).

Negative innovations in *DEF* tend to coincide with high market returns and, as such, will predict lower expected market returns in the good economic state. In that case, even a relatively small realized stock market return is sufficient to remain in the good economic state in the next period. Alternatively, negative innovations on dividends and the level of nominal interest rates are associated with higher expected market returns during a bad economic state. Thus, the probability of transitioning to the good economic state in the next period increases/decreases depending upon the magnitude of the realized shock in the stock market return.

### 4.2 The asset pricing model under learning and ambiguity

In this section, we examine whether ambiguity is priced in the cross-section of stock returns as predicted by Equation (5). We follow the two-pass cross-sectional regression method (henceforth CSR) of Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). In the first-pass, we run the following time-series regression for each asset:

$$R^e{}_{i,t} = \alpha_i + \beta_{i,MKT} R^e{}_{MKT,t} + \beta_{i,\pi^L} R^e{}_{\pi^L,t} + \beta_{i,KUNC} R^e{}_{KUNC,t} + \varepsilon_{i,t}, \tag{12}$$

where $R^e{}_{i,t}$ represents the excess returns on asset $i$, $R^e{}_{MKT,t}$ is the stock market return, $R^e{}_{\pi^L,t}, R^e{}_{KUNC,t}$ represents the estimated tracking portfolio excess returns of the innovations in the ambiguity factors $\pi_t^L$, *KUNC*, respectively, and $\varepsilon_{i,t}$ is the error term.

In the second-pass, average excess returns on test assets are related to the factor exposures estimated in the first-pass. We run the following cross-sectional regression each month:

$$E\left[{R^e}_{i,t+1}\right] = \lambda_0 + \lambda_{MKT}\beta_{i,MKT} + \lambda_{\pi^L}\beta_{i,\pi^L} + \lambda_{KUNC}\,\beta_{i,KUNC} + v_i, \tag{13}$$

where $E\left[{R^e}_{i,t+1}\right]$ represents the time series average of excess returns ${R^e}_{i,t+1}$, the $\beta' s$ denote factor loadings obtained from the first-pass time-series regressions, $\lambda_0$ is the intercept, $\lambda_{MKT}$ is the market price of systematic risk, $\lambda_{\pi^L}, \lambda_{KUNC}$ are the market prices of uncertainty related to the innovations in the two dimensions of ambiguity, and $v_i$ are pricing errors. If loadings with respect to factors related to ambiguity are important determinants of stock returns, then the estimated coefficients for $\lambda_{\pi^L}, \lambda_{KUNC}$ should be significantly different from zero.

We estimate the first-step factor loadings every 36 months based on rolling regressions.[24] We follow Shanken (1992) to address the well-known errors-in-variables (EIV) problem associated with using estimated loadings from the first step as independent variables in the cross-sectional regressions. The correction only affects the standard errors of the CSR coefficients. As such, the point estimates may still be affected by errors in the variables. We also apply Jagannathan and Wang's (1998) adjustment for the EIV problem under heteroskedasticy and autocorrelation. Additionally, the VAR innovations and transition probabilities are also generated regressors used in the first-pass time-series regressions causing a second EIV problem. Moreover, model misspecification may be severe, particularly when measuring unobservable macroeconomic risk factors. Consequently, we also report robust standard errors and cross-sectional goodness of fit R-squared statistics $\rho^2$ adjusted following the general approach in Kan, Robotti, and Shanken (2013).

[24] In robustness checks at the end of the article, we report results for alternative window lengths.

We report the estimates from generalized least squared (GLS) cross-sectional regressions in Table 3.[25] The test assets are the monthly excess returns on 25 portfolios sorted by size and book-to-market augmented with 30 portfolios sorted by industry, as suggested by Lewellen, Nagel, and Shanken (2010). Results based on the 25 portfolios only are entirely similar. In Panel A, we report the results for the full sample period from 1927-2007, while in Panel B, we provide the results for the post-COMPUSTAT period, 1962-2007. The goal is to determine whether an asset's exposure to ambiguity in the return generating process is an important determinant of its average return. Further, we are also interested in whether an asset exposure to ambiguity in the forecast of the economic state (i.e., learning) affects its return.

Insert Table 3 About Here.

We find that *KUNC* is priced in both sample periods. Over the 1927-2007 period (Panel A), the coefficient estimate on the *KUNC* beta is 0.36%, which is statistically significant at the 1% level. Over 1962-2007 (Panel B), the coefficient is 0.46% and remains significant at the 1% level. In the lower section of both Panel A and Panel B, we provide the results of model specification robust tests, including the robust cross-sectional goodness of fit R-squared statistics $\rho^2$ of Kan, Robotti, and Shanken (2013). In both sample periods, we reject the null hypothesis $H_0: \rho^2 = 1$ as well as $H_0: \rho^2 = 0$ at the 1% level.

We conclude that exposure to ambiguity in the return generating process is a priced factor in the cross-section of expected stock returns. Further, the learning under ambiguity asset pricing model is (asymptotically) correctly specified. However, the statistical significance of

[25] Following Kan, Robotti, and Shanken (2013), we also run the empirical analyses using weighted least-squared (WLS) obtaining similar results. We do not report these results to save space.

the intercept suggests possible misspecification in finite samples. We discuss the results of a model that does not include an intercept term in robustness checks at the end of the article.

In Table 3, $\pi_t^L$ is never statistically significant. We suspect this may be the result of including the intercept in the CSR while using excess returns as dependent variables. We present evidence supporting this view in the robustness checks section. However, even if not priced, the learning factor may still help explain variation in stock returns. We examine whether this is the case in Table 4, where we follow Kan, Robotti, and Shanken (2013) and report results of the second-pass CSR using covariances rather than betas as regressors in first-pass time-series regressions. We find that coefficient estimates on the covariances of $\pi_t^L$ and *KUNC* are both statistically significant. Thus, the learning factor improves the explanatory power of the expected return model. Following Kan, Robotti, and Shanken (2013), the problematic interpretation of the learning factor using multiple regression betas is due to the dependence of the factor with the other factors included in the first-pass time-series regressions. Depending upon the severity of the model misspecification, the difference between CSR estimates using multiple regression betas or covariances may be significant, especially if the correlation of the factor with the asset returns is very low.

Insert Table 4 About Here.

Overall, we conclude that an asset pricing model that includes a market risk factor and two additional factors associated with learning under ambiguity is asymptotically correctly specified and helps explain the cross-section of expected stock returns.

### 4.3 Comparison with the three factor Bayesian asset pricing model

In this section, we assess the incremental explanatory power of *KUNC* compared to Bayesian uncertainty. Ozoguz (2009) and Zhang (2003) investigate whether investors' beliefs and investors' uncertainty about the economy are state variables that describe changes in the investment opportunity set. In particular, Ozoguz (2009) finds that an asset pricing model that includes the market portfolio, the conditional probability of the good economic state, and a factor related to Bayesian uncertainty (*UC*) help explain the cross-section of average stock returns. Ozoguz (2009) measures the two Bayesian factors using a two-state Markov-switching model; $\hat{\pi}_t$ serves as a proxy for learning and *UC* = $\hat{\pi}_t(1-\hat{\pi}_t)$ for uncertainty.

In Table 5, Panel A, we report estimates and model specification robust tests for the three-factor standard Bayesian model from 1927 to 2007. In Panel B, we include estimates and model specification robust tests for the augmented version of the model that includes the ambiguity factor *KUNC*. The set of test assets includes the augmented set of 25 Fama-French portfolios sorted by size and book-to-market and 30 portfolios sorted by industry.[26] We find that the Bayesian proxy for model uncertainty *UC* is never significantly different from zero. In contrast, the ambiguity factor *KUNC* is statistically significant at the 1% level even after a battery of corrections on the standard errors that account for errors-in-variables and model misspecification problems. The coefficient estimate of 0.27% is economically meaningful. Further, the results of the Wald test and the model misspecification test for the null $H_0\colon \rho^2 = 0$ suggest the three-factor Bayesian model cannot explain the cross-section of expected stock returns and is asymptotically misspecified. The Wald test, however, is known to reject too often in finite samples.

[26] Ozoguz (2009) does not use tracking portfolio returns on orthogonalized innovations, so results in the two articles are not directly comparable.

Insert Table 5 About Here.

The Bayesian model performs better over the recent COMPUSTAT period 1962-2007, as can be seen in Table 6. This is consistent with the results in Ozoguz (2009). Nevertheless, once *KUNC* is included, none of the Bayesian variables are significant, while *KUNC* is significant at the 1% level with an estimated monthly premium of 0.25%. The ambiguity factor *KUNC* improves the robust cross-sectional *R*-squared goodness of fit statistic $\rho^2$ of the three-factor Bayesian model from 0.14 to 0.30 over the period 1927-2007, and from 0.28 to 0.41 over 1962-2007. Furthermore, we cannot reject the null hypothesis $H_0 : \rho^2 = 1$ at the 1% level for the model that includes *KUNC* over the recent period. Thus, a four-factor Bayesian model that also includes the ambiguity factor *KUNC* has good explanatory power for the cross-section of stock returns and is asymptotically correctly specified over the 1962-2007 sample period.

Insert Table 6 About Here.

Overall, we reach two major conclusions. First, ambiguity is distinct from Bayesian uncertainty and is statistically and economically more meaningful. Additionally, an augmented Bayesian model that accounts for ambiguity aversion performs better than the standard three-factor Bayesian model of Zhang (2003) and Ozoguz (2009).

### 4.4 Is the impact of ambiguity subsumed by risk-driven intertemporal effects?

In this section, we consider an alternative empirical specification of the learning under ambiguity model. We are interested in investigating whether the ambiguity factor *KUNC* is distinct from the intertemporal risk effects associated with changes in future investment opportunities (Merton 1973). Thus, we replace the distorted transition probability $\pi_t^L$ with the state variables driving the dynamics of the investment opportunity set (Petkova 2006). We run first-pass time-series regressions of the form:

$$R^e{}_{i,t} = \alpha_i + \beta_{i,MKT} R^e{}_{MKT,t} + \beta_{i,RF} R^e{}_{RF,t} + \beta_{i,DIV} R^e{}_{DIV,t} + \beta_{i,TERM} R^e{}_{TERM,t} + \beta_{i,DEF} R^e{}_{DEF,t} + \beta_{i,KUNC} R^e{}_{KUNC,t} + \varepsilon_{i,t}, \tag{14}$$

where $R^e{}_{i,t}$ is the excess return on asset $i = 1, \cdots, n$; $R^e{}_{MKT,t}$ is the stock market return and the rest of the $R^e{}_t{}'s$ represent the estimated tracking portfolio excess returns of the innovations in the state variables *MKT, RF, DIV*, *TERM*, *DEF*, and *KUNC*; and $\varepsilon_{i,t}$ is the error term.

After estimating (14), we run the following second-pass cross-sectional regression:

$$E\left[R^e{}_{i,t+1}\right] = \lambda_0 + \beta_{i,MKT}\lambda_{MKT} + \beta_{i,RF}\lambda_{RF} + \beta_{i,DIV}\lambda_{DIV} + \beta_{i,TERM}\lambda_{TERM} + \beta_{i,DEF}\lambda_{DEF} + + \beta_{i,KUNC}\lambda_{KUNC} + v_i \tag{15}$$

where $E\left[R^e{}_{i,t}\right]$ represents the time series average of excess returns $R^e{}_{i,t}$, the $\beta's$ denote factor loadings obtained from first-pass time-series regressions, $\lambda_0$ is the intercept, $\lambda_{MKT}, \lambda_{RF}, \lambda_{DIV}, \lambda_{TERM}, and\ \lambda_{DEF}$ are market prices of risk for market beta and intertemporal effects, $\lambda_{KUNC}$ denotes the ambiguity premium, and $v_i$ are the pricing errors.

In Table 7, we present results using the set of test assets that includes 25 Fama-French portfolios sorted by size and book-to-market and 30 portfolios sorted by industry.

Insert Table 7 About Here.

Loadings on *TERM* and *KUNC* are positive and statistically significant in both sample periods. The coefficient on *RF* is negative and marginally significant in the longer sample period. Further, the results of model specification robustness tests suggest the alternative empirical implementation of the learning under ambiguity model, although asymptotically misspecified, provides a reasonable description of the cross-section of expected stock returns. The performance of the alternative specification is similar to that of the initial specification based on $\pi_t^L$ and *KUNC* provided in Table 3.

A reasonable concern regarding the results reported in Table 7 is that our use of tracking portfolios, rather than the initial orthogonal innovations, may suffer from model specification. To alleviate this concern, we repeat the analysis done in Table 7 using innovations rather than tracking returns. The results, reported in Table 8, do not change in any material way. Both *TERM* and *KUNC* are positive and significant in both sample periods, while *RF* is negative and significant, but only in the longer sample. Petkova (2006, Table V) reports the results of a specification similar to the one in Table 8, but without including *KUNC*. Over 1963-2001, she reports the coefficient on *TERM* to be significantly positive and *RF* to be significantly negative. None of the other variables are significant. We report similar findings after including *KUNC*.

Insert Table 8 About Here.

Overall, we conclude that the impact on the cross-section of stock returns of unexpected changes in ambiguity regarding the (unknown) return generating process is distinct and separate from any intertemporal effects related to unexpected changes in future investment opportunities. Stated differently, our previous result that loadings on *KUNC* can help explain the cross-section of average stock returns does not seem to stem from a spurious association between our ambiguity measure and unexpected changes in variables known to predict stock returns. Furthermore, we find that the two empirical implementations of the learning under ambiguity model perform similarly well, offering support for the broad implications of the literature (e.g., Epstein and Schneider 2008).

**4.5 Incremental explanatory power of the Fama-French-Carhart factors**

In this section, we examine whether *KUNC* remains significant after controlling for the size and value factors of Fama and French (1993) and the momentum factor of Carhart (1997). We present the results of two types of analyses. In Table 9, the asset pricing factors are based on tracking portfolios. The factors considered include both $\pi_t^L$ and *KUNC*, in addition to *MKT*, *RF*, *DIV*, *TERM*, and *DEF*, as well as *SMB*, *HML*, and *UMD*. In Table 10, the factors are based on innovations rather than tracking portfolios. The factors considered include *KUNC*, *MKT*, *RF*, *DIV*, *TERM*, *DEF*, *SMB*, *HML*, and *UMD*.

Insert Table 9 and 10 About Here.

In Table 9, *KUNC* is significantly positive in both sample periods even after including loadings on *SMB*, *HML*, and *UMD*. In addition, loadings on *HML*, *UMD*, and *TERM* are significantly positive. In Table 10, using innovations as in Petkova (2006), both *KUNC* and *TERM* remain positive and significant. In contrast, none of the other variables are significant based on the robust t-statistics of Kan, Robotti, and Shanken (2013). Petkova (2006) reports similar results with respect to *TERM* and the Fama-French factors. We conclude that the firm-specific factors of Fama and French (1993) and Carhart (1997) do not subsume the information contained in *KUNC* regarding uncertainty in the cross-section of average stock returns.

## 5. A Comparison of Empirical Asset Pricing Models

We compare the performance of the learning under ambiguity model against three prominent empirical asset pricing models: 1) the Bayesian model of Ozoguz (2009), 2) the ICAPM model of Petkova (2006), and 3) the four-factor model of Fama and French (1993) and Carhart (1997). We use the robust pair-wise comparative tests developed by Kan, Robotti, and Shanken (2013) for cross-sectional regressions $\rho^2$ of nested and non-nested models. The null hypothesis is that the difference in the $\rho^2$ statistics between the two models is not statistically significant.

In Table 11, we compare the $\rho^2$ of the empirical implementation of the Learning Under Ambiguity Model that includes loadings on *MKT*, *RF*, *DIV*, *TERM*, *DEF*, and *KUNC* versus the Augmented Model that also includes the Bayesian proxy for uncertainty *UC* = $\hat{\pi}_t(1-\hat{\pi}_t)$. The objective is to evaluate whether Bayesian uncertainty has incremental explanatory power relative to the learning under ambiguity model. The table reports the difference between the sample cross-sectional $\rho^2$ of the Augmented Model minus the Learning Under Ambiguity Model, and the associated p-values.

Insert Table 11 About Here.

Although the difference is negative, we cannot reject the null that the two models have similar explanatory power over the longer sample period. Over the shorter period, there is marginal evidence that the augmented model performs worse than the model without the Bayesian proxy. Thus, adding the Bayesian uncertainty measure does not increase the explanatory power of the learning under ambiguity model.

Kan, Robotti, and Shanken (2013) report that in the post-COMPUSTAT period, the empirical implementation of the ICAPM of Petkova (2006) does not pass asset pricing tests robust to model misspecification, in contrast to the four-factor model (FF4) of Fama-French (1993) and Carhart (1997). Therefore, in Table 12, we compare the following three models: 1) the FF4 model that includes estimated loadings on *MKT*, *SMB*, *HML*, and *UMD*; 2) the empirical ICAPM of Petkova (2006) that includes *MKT*, *RF*, *DIV*, *TERM*, and *DEF*; and 3) the empirical implementation of the learning under ambiguity model that includes *MKT*, *RF*, *DIV*, *TERM*, *DEF*, and *KUNC*.

Insert Table 12 About Here.

The empirical implementation of the learning under ambiguity model outperforms the FF4 once model misspecification is taken into account. The $\rho^2$ of FF4 is about 0.26 lower than the $\rho^2$ of the ambiguity model, and the difference is significant at the 1% level. The learning under ambiguity model performs as well as the standard ICAPM. Furthermore, when compared against the FF4 model over the 1962-2007 period, the learning under ambiguity model survives model misspecification tests, unlike the empirical model of Petkova (2006). The results of the comparative analyses confirm the findings of the previous sections that ambiguity helps explain the cross-section of stock returns better than alternative models, particularly after model misspecification is taken into account.

## 6. Robustness Checks

We present the results of four types of robustness checks. First, we previously speculated that the loading on $\pi_t^L$ was not statistically significant due to the inclusion of the intercept in the regressions. Thus, we now assess the impact that the inclusion of the intercept has on the CSR estimates of the first empirical implementation of the learning under ambiguity model.[27] To evaluate the null hypothesis of zero mispricing, we cannot use the cross-sectional *R*-squared statistic $\rho^2$. Instead, we use the composite pricing error $T\bar{\boldsymbol{\alpha}}'\boldsymbol{\Lambda}^{-1}\bar{\boldsymbol{\alpha}}$, where $T$ is the size of the sample period, $\bar{\boldsymbol{\alpha}}$ denotes the average residual vector in the second-pass cross-section regression, $\Lambda$ is a symmetric positive weighting matrix that is equal to $\Lambda = \hat{V}_{22}^{-1}$ under generalized least squares (GLS), and $\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$ is the variance-covariance matrix of the composite vector of risk factors and test asset returns. This cross-sectional test has an asymptotic chi-squared distribution. As a second robustness check, we present results using the raw values of the factors, as opposed to innovations from the VAR model or tracking portfolios of the innovations.

[27] Introducing the intercept in the CSR with excess returns as dependent variables may introduce misspecification in the CSR (Cochrane 2001, Section 12.2, page 235; Brennan, Xia, and Wang 2004). Alternatively, the absence of the intercept in the CSR will bias estimates upward. In the empirical asset pricing literature, one can find results with and without the CSR intercept.

Moreover, our theoretically motivated factors $\pi_t^L$ and $KUNC$ vary with the minimum confidence level that investors have in their Q-MLE forecasts (i.e., in the reference model). Thus, to assess the impact that different confidence levels have on our results, we repeat the econometric estimation and asset pricing tests assuming different minimum confidence levels (CL). More specifically, we consider a range of minimum confidence levels from 50% to 90%. The results in Zhang (2003) and Ozoguz (2009) correspond to the special case of a Bayesian investor with a 90% minimum confidence level in his reference model.

Finally, we also assess the impact that time-varying betas have on our results. Specifically, we repeat the estimations and asset pricing tests increasing the window length (WL) of the first-pass rolling regressions from 36 months to 180 months.

The results of the robustness checks are presented in Table 13. To save space, we report the results for the longer sample period from 1927-2007; results for 1962-2007 are entirely similar. The market price of beta risk is positive and statistically significant with an estimate of around 0.7% per month across different minimum confidence levels and rolling regression windows. The empirical asset pricing literature has shown that exclusion of the intercept in CSR leads to equity risk premium estimates closer to the historical average.

The learning component of the uncertainty premium is positive and statistically significant, but only when the ambiguity averse investor uses relatively short predictive regressions (i.e., when the forecast horizon is less than 96 months). The difference between the initial results reported in Table 3 and the results in Table 13 suggest that estimates on both $\pi_t^L$ and *MKT* are very sensitive to model misspecification.

Insert Table 13 About Here.

In contrast, the premium on *KUNC* is positive and statistically significant. Premium estimates vary from 0.70% to 3.0%, rising as minimum confidence levels decrease. The magnitude of these estimates cannot be compared with those obtained in Table 3 because the factors in Table 13 are not returns from tradable portfolios. Consequently, the estimate is not equal to the average return of the factor. We conclude that our previous results on *KUNC* are robust to model misspecification. Additionally, we provide chi-squared test results at the bottom of Table 13. We cannot reject the null hypothesis of zero mispricing at the 10% level. Although not shown, we cannot reject the null at the 1% level over the post-COMPUSTAT period.

## 7. Conclusion

We consider a dynamic framework of learning under ambiguity as a specialization of the more general recursive multiple-priors setting of Epstein and Schneider (2003). In this framework, the solution to the investor problem leads to a fundamental pricing equation in which investors form expectations relative to the worst case probability measure. In beta regression form, the result is a model of learning under ambiguity that includes three-factors: 1) systematic risk, 2) uncertainty regarding the state of the economy (i.e., learning under ambiguity), and 3) uncertainty regarding the data generating process driving stock returns (i.e., ambiguity).

To test the model empirically, we propose a novel ambiguity measure that arises theoretically when the recursive multiple-priors set is constrained based on relative entropy. Using classic statistical results on simultaneous confidence levels, the measure is defined as the first order approximation of the worst log-likelihood ratio between any alternative model and the reference model.

We find that ambiguity is priced in the cross-section of average stock returns, and that the explanatory power of the ambiguity measure is not subsumed by the introduction of other asset pricing factors. Furthermore, the learning under ambiguity model performs better than the three-factor Bayesian learning model of Ozoguz (2009), the empirical implementation of the ICAPM of Petkova (2006), and the standard Fama, French (1993) and Carhart (1997) model. Altogether, our empirical results provide support for the broad predictions of the learning under ambiguity literature (Epstein and Schneider 2008) and suggest that ambiguity aversion can help explain the cross-section of stock returns.

An interesting open question is whether ambiguity can explain well known anomalies such as the value premium. The existing literature finds it difficult to explain why the "riskier" growth stocks tend to underperform value stocks. We conjecture that an ambiguity averse investor may have a preference for growth stocks because they perform relatively better when the investor's confidence in the return generating model is lower. More broadly, ambiguity aversion has the potential to explain the many apparently unrelated anomalies that have been reported.

More technically, an important extension of the paper would recover jointly (as opposed to in separate steps) both the reference probability and the worst case scenario probability. This is left for future research.

## Appendix A

We start by modifying the standard dynamic programming consumption-investment problem to account for ambiguity in the transition probabilities describing the state of the economy.

Assume the representative agent is endowed with a standard time-additive separable Von Neumann Morgenstern utility of consumption. Let $R_{i,t+1}^{s}$denote the conditional gross return of asset $i$ over the period starting at time $t$ and ending up at time $t+1$. Also, let $R_{f,t-1}$ represent the gross return of a dollar invested in the risk-free asset. Then, the conditional wealth constraint is given by:

$$W_{t+1}^{S} = (W_t - C_t)\left(R_{f,t-1} + \sum_{i=1}^{n} \xi_{i,t}\left(R_{i,t+1}^{s} - R_f\right)\right). \tag{A.1}$$

Starting from date $T$, we iterate backward period by period to obtain the dynamically constrained Bellman Equation in (2):

$$J(W_t, t) = Max_{C_{t,\{\xi_{i,t}\}}}\left(u(C_t) + Min_{\pi_t^*} E_t\left[\pi_t^* J(W_{t+1}^{1}, t+1) + (1-\pi_t^*)J(W_{t+1}^{0}, t+1) + \frac{1}{2}\theta_t(D_t(\pi_t^* \| \pi_t) - \eta_t)\right]\right), \tag{A.2}$$

where $\theta_t \geq 0$ is a Lagrange multiplier for the ambiguity constraint in Equation (1). Next, we insert the approximation (B.3) in Appendix B into (A.2) to generate the necessary first order condition for optimality of the inner minimization problem:

$$\pi_t - \pi_t^* = \frac{\left(E_t\left(J(W_{t+1}^{1}, t+1)\right) - E_t\left(J(W_{t+1}^{0}, t+1)\right)\right)\pi_t(1-\pi_t)}{\theta_t}. \tag{A.3}$$

From Equation (1) express the complementary slackness condition of the entropy constraint as:

$$\frac{1}{2}\ \theta_t \left(\frac{(\pi_t-\pi_t^*)^2}{\pi_t(1-\pi_t)} - \eta_t\right) = 0. \tag{A.4}$$

Substitute (A.3) into (A.4) to obtain:

$$\theta_t = \frac{\left(E_t\big(J(W_{t+1}^1,t+1)\big)-E_t\big(J(W_{t+1}^0,t+1)\big)\right)\pi_t(1-\pi_t)}{\sqrt{2\eta_t}}. \tag{A.5}$$

Finally, inserting (A.5) back into (A.3) yields:

$$\pi_t^L \equiv \pi_t^* = \pi_t - \sqrt{2\eta_t\pi_t(1-\pi_t)}. \tag{A.6}$$

Hence, we must have: $0 \leq \frac{\pi_t-\pi_t^L}{\pi_t} \leq +1$. When (A.6) holds, (A.2) turns into the usual unconstrained maximization problem:

$$J(W_t,t) = Max_{C_{t,\{\xi_{i,t}\}}}\left(u(C_t) + E_t^{\pi_t^L}\big(J(W_{t+1},t+1)\big)\right), \tag{A.7}$$

where $E_t^{\pi_t^L}\big(J(W_{t+1},t+1)\big) \equiv \pi_t^L E_t\big(J(W_{t+1}^1,t+1)\big) + (1-\pi_t^L)E_t\big(J(W_{t+1}^0,t+1)\big)$ denotes the *ambiguity certainty equivalent* (see Epstein and Schneider (2010) pp. 325-326) of the expected continuation values $E_t\big(J(W_{t+1}^s,t+1)\big)$ under the worst case scenario probability $\pi_t^L$. Finally, standard first-order conditions (see Pennachi (2008)) applied to (A.7) deliver Proposition 1. Q.E.D.

## Appendix B

We provide an approximation of the Kullback-Leibler divergence between alternative model $\pi_t^*$ and reference model $\pi_t$ probability measures in our setting **(**Cover and Thomas 1991, pp.18 and 19 for technical details and interpretation). With our notation we obtain:

$$D_t(\pi_t^* \| \pi_t) = E_{\pi_t^*}\left(ln\left(\frac{\pi_t^*}{\pi_t}\right)\right) = \pi_t^* ln\left(\frac{\pi_t^*}{\pi_t}\right) + (1-\pi_t^*) ln\left(\frac{1-\pi_t^*}{1-\pi_t}\right). \qquad \text{(B.1)}$$

A first order Taylor series approximation yields:

$$ln\left(\frac{\pi_t^*}{\pi_t}\right) \approx \frac{\pi_t^* - \pi_t}{\pi_t}. \qquad \text{(B.2)}$$

Observe that the above ratio is well behaved since $-1 \leq \frac{\pi_t^* - \pi_t}{\pi_t} \leq +1$ as long $\pi_t$ is non-degenerate. Substituting (B.2) back into Equation (B.1) we obtain after algebra:

$$D_t(\pi_t^* \| \pi_t) \approx \frac{(\pi_t - \pi_t^*)^2}{\pi_t(1-\pi_t)}. \qquad \text{(B.3)}$$

In the entropy literature (Golan, Judge, and Miller 1996, pp. 31), Expression (B.3) is interpreted as a shrinkage estimator.

Define $\hat{\pi}_t$ as the Q-MLE of the transition probability of the “expansion” economic state. This is the reference model where the data generating process driving the stock market return is modeled as formally described in Appendix A. Given a sample of time-series observations with size $T$, let $p_t = t_{\{s_t=1\}}/T$ be the empirical likelihood or frequency of observing good economic times under a minimum confidence level on the forecast of $(1-\varsigma)$.

Ambiguity averse investors acknowledge that the observed empirical frequencies are drawn from a hidden distribution of multinomial proportions (given that "Nature" chooses the distribution of returns with maximum entropy or multiplicity). Thus, they will simultaneously entertain a continuum of alternative possible close likelihoods around their reference forecast. We follow Goodman (1965) and find the asymptotic interval of the reference model using the Bonferroni inequality.

Thus, $T \to +\infty$, $p_t$ is approximately normally distributed with mean $\hat{\pi}_t$ and variance $\hat{\pi}_t(1-\hat{\pi}_t)/T$ so that the empirical counterpart of Equation (B.3):

$$Z_t = \frac{\sqrt{T}(p_t - \hat{\pi}_t)}{\sqrt{\hat{\pi}_t(1-\hat{\pi}_t)}}, \tag{B.4}$$

is distributed as a standard normal variate. The interval for $\hat{\pi}_t$ has bounds equal to the solutions of the quadratic equation:

$$T(p_t - \hat{\pi}_t)^2 = \chi^2(1-\varsigma)\hat{\pi}_t(1-\hat{\pi}_t), \tag{B.5}$$

or what is the same:

$$\hat{\pi}_t{}^2\left(T + \chi^2(1-\varsigma,1)\right) - \left(2t_{\{s_t=1\}} + \chi^2(1-\varsigma,1)\right)\hat{\pi}_t + t^2_{\{s_t=1\}}/T = 0, \tag{B.6}$$

where $\chi^2(1-\varsigma, 1)$ denotes the quantile of order $(1-\varsigma)$ of a chi-square distribution with one degree of freedom. Equations (B.5) and (B.6) have two solutions that define the lower and upper bounds defining the interval $[\pi_t^L, \pi_t^H]$ around $\hat{\pi}_t$ with coverage probability $1 - T\varsigma$. An investor averse to model ambiguity that seeks safety first selects the lower bound $\pi_t^L$ of the interval as the conditional probability of a good economic regime next period. That is, the investor slants his belief toward the recession state with low consumption continuation values.

Q.E.D.

## References


Anderson, E. W., E. Ghysels, and J. L. Juergens. 2009. The impact of risk and uncertainty on expected returns. *Journal of Financial Economics* 94: 233-263.

Barillas, F., L. P. Hansen, and T. J. Sargent. 2009. Doubts or variability?" *Journal of Economic Theory* 144: 2388-2418.

Black, F., M. C. Jensen, and M. Scholes. 1972. The capital asset pricing model: Some empirical finding. In *Studies in the Theory of Capital Markets*, ed. M.C. Jensen, New York, NY: Praeger.

Breeden, D., M. Gibbons, and R. Litzenberger. 1989. Empirical tests of the consumption-oriented CAPM. *Journal of Finance* 44, 1595-1616.

Brennan, M. J., A. W. Wang, and Y. Xia. 2004. Estimation and test of a simple model of intertemporal capital asset pricing. *Journal of Finance* 59: 1743-1776.

Campbell, J. Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18: 373-399.

Campbell, J. Y. 1996, Understanding risk and return. *Journal of Political Economy* 104, 298-345.

Campbell, J. Y., and R. J. Shiller. 1988. Stock prices, earnings, and expected dividends. *Journal of Finance* 43: 661-676.

Carhart, M. M. 1997. On persistence of mutual fund performance. *Journal of Finance* 52: 57-82.

Chen, N. F. 1991. Financial investment opportunities and the macroeconomy. *Journal of Finance* 46: 529-554.

Chen, Z., and L. G. Epstein. 2002. Ambiguity, risk and asset returns in continuous time. *Econometrica* 70: 1403-1443.

Cochrane, J. H. 1996. A Cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy* 104: 572-621.

Cochrane, J. H. 2001. *Asset Pricing.* Princeton, NJ: Princeton University Press.

Cover, J., and M. Thomas. 1991. *Elements of Information Theory*. Hoboken, NJ: John Wiley & Sons.

David, A., and P. Veronesi. 2001. What ties return volatilities to price valuations and fundamentals? *Chicago Booth Research and CRSP Working Paper No. 10-05*, University of Calgary and University of Chicago.

Ellsberg, D. 1961. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75: 643-669.

Epstein, L. G., M. Schneider. 2003. Recursive multiple priors. *Journal of Economic Theory* 113: 1-13.

Epstein, L. G., and M. Schneider. 2007. “Learning under ambiguity. *Review of Economic Studies* 74: 1275-1303.

Epstein, L. G., and M. Schneider. 2008. Ambiguity, information quality and asset pricing. *Journal of Finance* 63: 197-228.

Epstein, L. G., and M. Schneider. 2010. Ambiguity and asset markets. *Working paper,* Boston University and Stanford University.

Epstein, L. G., and T. Wang. 1994. Intertemporal asset pricing under Knightian uncertainty. *Econometrica* 62: 283-322.

Estrella, A., and G. A. Hardouvelis. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46: 555-576.

Fama, E. F., and K. R. French. 1993. The cross-section of expected returns. *Journal of Financial Economics* 33: 3-56.

Fama, E. F., and K. R. French. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25: 23-49.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81: 607-636.

Fama, E. F., and W. Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5: 115-146.

Ferson, W. E., and C. R. Harvey. 1999. Conditioning variables and the cross-section of stock returns.” *Journal of Finance* 44: 1325-1360.

Ferson, W. E., A. F. Seigel, and P. Xu. 2006, Mimicking portfolios with conditioning information.” *Journal of Financial and Quantitative Analysis* 41: 607-635.

Ferson, W. E., T. Simin, and S. Sarkissian. 2003. Spurious regressions in financial economics? *Journal of Finance* 58: 1393-1414.

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies* 20: 41-81.

Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. West Sussex, UK: John Wiley & Sons.

Goodman, L. A. 1965. On simultaneous confidence intervals for multinomial proportions. *Technometrics* 7: 249-275.

Granger, C. W. J., and T. Teräsvirta. 1993. *Modeling Non-Linear Economic Relationships*. Oxford, UK: Oxford University Press.

Guidolin, M., and F. Rinaldi. 2013. Ambiguity in asset pricing and portfolio choice: A review of the literature." *Theory and Decision* 74: 183-217.

Grünwald, P. D., and A. P. Dawid. 2004. Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory. *The Annals of Statistics* 32: 1367-1433.

Hamilton, J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57: 357-384.

Hamilton, J. D. 1990. Analysis of time series subject to changes in regime. *Journal of Econometrics* 45: 39-70.

Hansen, L. P. 2008. Beliefs, doubts and learning: Valuing macroeconomic risk. *American Economic Review* 97: 1-30.

Hansen, L. P., and T. J. Sargent. 2008. *Robustness*. Princeton, NJ: Princeton University Press.

Hansen, L. P., and T. J. Sargent. 2010. Fragile beliefs and the price of uncertainty. *Quantitative Economics* 1: 129-162.

Hansen, L. P., and T. J. Sargent. 2011. Robustness and ambiguity in continuous time. *Journal of Economic Theory* 146: 1195-1223.

Harremoës, P., and F. Topsøe. 2001. Maximum entropy fundamentals. *Entropy* 3: 191-226.

Jagannathan, R., and Z. Wang. 1998. An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *Journal of Finance* 53: 1285-1309.

Jegadeesh, N. 1990. Evidence of predictable behavior of security returns. *Journal of Finance* 45: 881-898.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48: 65-91.

Ju, N., and J. Miao. 2012. Ambiguity, learning, and asset returns. *Econometrica* 80: 559-591.

Kan, R., C. Robotti, and J. Shanken. 2013. Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance* Forthcoming.

Keynes, J. M. 1921. *A Treatise on Probability*. Cambridge, UK: Macmillan Co.
Knight, F. H. 1921. *Risk, Uncertainty and Profit*. Chicago, IL: University Chicago Press.

Lamont, O. A. 2001. Economic tracking portfolios. *Journal of Econometrics* 105: 161-184.

Leippold, M., F. Trojani, and P. Vanini. 2008. Learning and asset prices under ambiguous information. *Review of Financial Studies* 21: 2565-2597.

Lettau, M., and S. Ludvigson. 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109: 1238-1287.

Lewellen, J., and S. Nagel. 2006. The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics* 82: 289-314.

Lewellen, J., S. Nagel, and J. Shanken. 2010. A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96: 175-194.

Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47: 233-275.

Ludvigson, S.C. 2012. Advances in consumption-based asset pricing: Empirical tests. Forthcoming in volume 2 of the Handbook of the Economics of Finance.

Maenhout, P. J. 2004. Robust portfolio rules and asset pricing. *Review of Financial Studies* 17: 951-983.

Merton, R. C. 1973. An intertemporal capital asset pricing model." *Econometrica* 41: 867-887.

Ozoguz, A. 2009. Good times or bad times? Investors' uncertainty and stock returns. *Review of Financial Studies* 22: 4377-4422.

Pennacchi, G. 2008. *Theory of Asset Pricing*. NJ, USA: Prentice Hall.

Petkova, R. 2006. Do the Fama-French factors proxy for innovations in predictive variables? *Journal of Finance* 61: 581-612.

Petkova, R., and L. Zhang, 2005. "Is value riskier than growth? *Journal of Financial Economics* 78: 187-202.

Roy, A. D. 1952. Safety first and the holding of assets. *Econometrica* 20: 431-449.

Shackle, G. L. S. 1949. *Expectation in Economics*. Cambridge, MA: Cambridge University Press.

Shanken J. 1992. On the estimation of beta pricing models. *Review of Financial Studies* 5: 1-34.

Singleton, K. J. 2006. *Empirical Dynamic Asset Pricing*. Princeton, NJ: Princeton University Press.

Teräsvirta, T. 1998. Modelling economic relationships with smooth transition regressions. In *Handbook of Applied Economic Statistics*, ed. M. Dekker, New York, NY: CRC Press.

Tong, H. 1990. *Non-linear Time Series: A Dynamical Systems Approach*. Oxford, UK: Oxford University Press.

Topsøe, F. 1979. Information theoretical optimization techniques. *Kybernetika* 15: 8-27.

Topsøe, F. 2007. Information theory at the service of science. In *Entropy, Search, Complexity*, ed. I. Csiszár, G. O. H. Katona, G. Tardos, and G. Wiener, G., 179-207. Berlin Heidelberg: Springer.

Uppal, R., and T. Wang. 2003. Model misspecification and underdiversification. *Journal of Finance* 58(6): 2465-2486.

van Dijk, D., T. Teräsvirta, and P. H. Franses. 2002. Smooth transition autoregressive models-A survey of recent developments. *Econometric Reviews* 21: 1-47.

Veronesi, P. 1999. Stock market overreaction to bad news in good times: A rational expectation equilibrium model. *Review of Financial Studies* 12: 975-1007.

Williams, C. D. 2009. "Asymmetric responses to good and bad news: An empirical case for ambiguity." *Working paper*, University of Michigan.

Zhang, Z. 2003. State uncertainty and the cross-sectional returns: Theory and evidence. *Working paper*, Singapore Management University.

**Table 1**
**Summary statistics: Bayesian uncertainty vs. ambiguity**

| **Panel A** | **1927-1961** | | **1962-2007** | |
|---|---|---|---|---|
| ***Ambiguity measures*** | $\boldsymbol{\pi^L}$ | $\boldsymbol{KUNC}$ | $\boldsymbol{\pi^L}$ | $\boldsymbol{KUNC}$ |
| Mean | 0.6243 | 0.2260 | 0.5683 | 0.2104 |
| Std. dev. | 0.3932 | 0.3197 | 0.3500 | 0.2691 |
| Min. | 0.0000 | 0.0238 | 0.0000 | 0.0238 |
| Max. | 0.9762 | 1.0000 | 0.9762 | 1.0000 |
| ***Bayesian uncertainty measures*** | $\boldsymbol{\hat{\pi}}$ | $\boldsymbol{UC}$ | $\boldsymbol{\hat{\pi}}$ | $\boldsymbol{UC}$ |
| Mean | 0.6615 | 0.0668 | 0.6172 | 0.1106 |
| Std. Dev. | 0.3968 | 0.0881 | 0.3549 | 0.0932 |
| Min. | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Max. | 1.0000 | 0.2500 | 1.0000 | 0.2500 |

| **Panel B:** ***Correlation matrix*** | $\boldsymbol{KUNC}$ | $\boldsymbol{UC}$ | $\boldsymbol{\pi^L}$ | $\boldsymbol{\hat{\pi}}$ |
|---|---|---|---|---|
| $KUNC$ | 1 | -0.2100 | -0.8400 | -0.8600 |
| $UC$ | | 1 | -0.2500 | -0.1800 |
| $\pi^L$ | | | 1 | 1.0000 |
| $\hat{\pi}$ | | | | 1 |

The table reports summary statistics for the proxies for investors' beliefs and uncertainty over the hidden state of the market in the standard Bayesian and Ambiguity cases. $\hat{\pi}$ is the Q-MLE of the conditional probability of transitioning to the good economic state using the STAR model of the market return; $\pi^L$ is the distorted (lower bound) conditional probability of transitioning to the good economic state using Goodman's procedure; $UC = \hat{\pi}(1-\hat{\pi})$ denotes the standard Bayesian index of uncertainty constructed using the Q-MLE transition probabilities (Ozoguz 2009); and $KUNC = (\hat{\pi} - \pi^L)/\hat{\pi}$ denotes the ambiguity measure, constructed using the Q-MLE and the distorted probability measures. In Panel A, the sample period is April 1927-December 1961 or January 1962-December 2007. Panel B reports the correlation coefficients among the belief and uncertainty measures over April 1927-December 2007.

## Table 2
## The STAR model

| Panel A: Excess Stock Market Return Equation | | | | |
|---|---|---|---|---|
| | Estimates | Standard Error | *t*-ratio | *p*-value |
| Regime 1 (Good economic times) | | | | |
| Student *t* d.f.^(1/2) | 3.6946 | 0.7294 | -- | -- |
| Log (Skewness) | -0.1939 | 0.0591 | -3.2810 | 0.0010 |
| Intercept $(\hat{\phi}_{1,0})$ | 0.0023 | 0.0083 | 0.3540 | 0.7230 |
| AR1$(\hat{\phi}_1)$ | -0.9524 | 0.0302 | -31.5260 | 0.0000 |
| MA1 $(\hat{\zeta})$ | -0.9638 | 0.0219 | -44.0680 | 0.0000 |
| GARCH Intercept $(\hat{a}_{1,0})$ | 0.0438 | 0.0088 | -- | -- |
| GARCH AR1 $(\hat{b})$ | 0.9499 | 0.0198 | 48.0250 | 0.0000 |
| GARCH MA1 $(\hat{a}_1)$ | 0.8534 | 0.0248 | 34.3830 | 0.0000 |
| Regime 0 (Bad economic times) | | | | |
| Intercept $(\hat{\phi}_{0,0})$ | 0.0176 | 0.0096 | 1.8240 | 0.0680 |
| GARCH Intercept $(\hat{a}_{0,0})$ | 0.0572 | 0.0142 | 4.0180 | 0.0000 |
| Log Likelihood: | 1650.28 | | | |
| **Panel B: Transition Function** | | | | |
| | Estimates | Standard Error | *t*-stat | *p*-value |
| Regime 1 (Good economic times) | | | | |
| $\hat{\delta}_{1,0}$ | -0.6871 | 0.2454 | -2.7990 | 0.0050 |
| $\hat{\delta}_{1,RF(-1)}$ | 0.0022 | 0.0146 | 0.1500 | 0.8810 |
| $\hat{\delta}_{1,DIV(-1)}$ | -0.0164 | 0.0251 | -0.6560 | 0.5120 |
| $\hat{\delta}_{1,TERM(-1)}$ | -0.0477 | 0.0245 | -1.9450 | 0.0520 |
| $\hat{\delta}_{1,DEF(-1)}$ | 0.2252 | 0.0816 | 2.7610 | 0.0060 |
| Regime 0 (Bad economic times) | | | | |
| $\hat{\delta}_{0,0}$ | -0.7142 | 0.1512 | -4.7240 | 0.0000 |
| $\hat{\delta}_{0,RF(-1)}$ | -0.0401 | 0.0177 | -2.2640 | 0.0240 |
| $\hat{\delta}_{0,DIV(-1)}$ | -0.3615 | 0.0923 | -3.9180 | 0.0000 |
| $\hat{\delta}_{0,TERM(-1)}$ | -0.1538 | 0.0382 | -4.0290 | 0.0000 |
| $\delta_{0,DEF(-1)}$ | 0.0231 | 0.0156 | 1.4820 | 0.1390 |
| Smoothness $(\hat{\psi})$ | 36.5380 | 17.6785 | 2.0670 | 0.0390 |

The table reports conditional quasi maximum-likelihood (Q-MLE) estimates of the self-exciting smooth double transition (STAR-GARCH) model for the excess stock market return:

$$R^e_{MKT,t} = \left(\phi_{1,0} + \phi_1 R^e_{MKT.t-1}\right)\left(G(\omega_t;\psi,\boldsymbol{\delta})\right) + \left(\phi_{0,0} + \phi_1 R^e_{MKT.t-1}\right)\left(1 - G(\omega_t;\psi,\boldsymbol{\delta})\right) + (\varepsilon_t - \zeta\varepsilon_{t-1})$$
$$\varepsilon_t = \sqrt{\hbar_t}v_t,\ \hbar_t = \left(G(\omega_t;\psi,\boldsymbol{\delta})\right)\hbar_{1,t} + \left(1 - G(\omega_t;\psi,\boldsymbol{\delta})\right)$$
$$\hbar_{1,t} = a_{1,0} + a_1\varepsilon^2_{t-1} + b\hbar_{1,t-1},\ \hbar_{0,t} = a_{0,0} + a_1\varepsilon^2_{t-1} + b\hbar_{0,t-1}$$
$$G(\omega_t;\psi,\boldsymbol{\delta}) = \left(1 + exp\left\{-\psi\left(R^e_{MKT,t-1} - \delta_{1,0} - \boldsymbol{\delta}'_{\mathbf{1,1}}dz_{t-1}\right)\left(R^e_{MKT,t-1} - \delta_{0,0} - \boldsymbol{\delta}'_{\mathbf{0,1}}dz_{t-1}\right)\right\}\right)^{-1}$$

where the switching "indicator" variable $s_t$ is the one-period lagged dependent variable $r^e_{MKT,t-1}$; $\psi \geq 0$ is the "smoothness" parameter; $dz_{t-1}$ denotes a vector of innovations in a group of exogenous state variables *RF(-1)*, *DIV(-1)*, *TERM(-1)*, and *DEF(-1)* lagged one period; $\phi_{1,0}$, $\phi_{0,0}$ are state dependent intercepts; and $\phi_1$ and $\zeta$ are the autoregressive and moving average parameters of the data generating process (DGP). The errors are assumed to follow a skewed Student *t* distribution with parameters equal to the square of the number of degrees of freedom (d.f.) and Log(Skewness). The GARCH parameters are in ARMA-in-squares form. Robust standard errors are calculated following the Newey-West HAC formula to correct for heteroskedastic and autocorrelation effects, and using a non-parametric Parzen kernel and plug-in bandwidth equal to five in order to account for parametric functional misspecification (see Ferson, Simin, and Sarkissian, 2003). The sample period is 1927-2007.

**Table 3**
**Regression results for the learning under ambiguity model**

| **Panel A: 1927-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{\pi^L}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|
| Estimates: | 0.0090*** | 0.0012 | 0.0020 | 0.0036*** |
| $t_{FM}$ | 6.7836 | 0.5357 | 1.4717 | 4.3529 |
| $t_S$ | 6.6064 | 0.5304 | 1.4659 | 4.3150 |
| $t_{JW}$ | 5.8953 | 0.4949 | 1.4375 | 4.4339 |
| $t_{KRS}$ | 5.1524 | 0.4701 | 1.3039 | 3.7591 |
| $\rho^2$ | 0.1395 | | | |
| $p(\rho^2 = 1)$ | 0.0000 | | | |
| $p(\rho^2 = 0)$ | 0.0045 | | | |
| $Wald$ | 19.7767 | | | |
| $p(Wald)$ | 0.0017 | | | |
| **Panel B: 1962-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{\pi^L}$ | $\hat{\lambda}_{KUNC}$ |
| Estimates | 0.0126*** | -0.0027 | 0.0005 | 0.0046*** |
| $t_{FM}$ | 8.0432 | -1.1130 | 0.3569 | 4.9175 |
| $t_S$ | 7.5256 | -1.0810 | 0.3543 | 4.8338 |
| $t_{JW}$ | 7.1480 | -1.0864 | 0.3519 | 4.9046 |
| $t_{KRS}$ | 6.0201 | -1.0000 | 0.3436 | 4.5038 |
| $\rho^2$ | 0.2281 | | | |
| $p(\rho^2 = 1)$ | 0.0015 | | | |
| $p(\rho^2 = 0)$ | 0.0000 | | | |
| $Wald$ | 31.9242 | | | |
| $p(Wald)$ | 0.0000 | | | |

The table reports cross-sectional regression (CSR) estimates and asset pricing robust test results for the following empirical specification of the learning under ambiguity model:

$$E\left[R^e_{i,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\pi^L}\lambda_{\pi^L} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. The factors are the returns of tracking economic portfolios on "orthogonalized" innovations. $\rho^2$ denotes the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the $p$-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 1$. $p(\rho^2 = 0)$ is the $p$-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 0$. $Wald$ is the result of the joint test of the null hypothesis $H_0: \hat{\lambda} = 0$ with $p$-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-MacBeth (1973) $t$-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted $t$-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted $t$-stats ($t_{JW}$) to correct for Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken's (2013) robust $t$-stats under model misspecification ($t_{KRS}$). (***) denotes a 1% significance level, (**) a 5% significance level, and (*) a 10% significance level.

**Table 4**
**Contribution of the learning factor to the variation of stock returns**

| | 1927-2007 | | | | 1962-2007 | | | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{\gamma}_0$ | $\hat{\gamma}_{MKT}$ | $\hat{\gamma}_{\pi^L}$ | $\hat{\gamma}_{KUNC}$ | $\hat{\gamma}_0$ | $\hat{\gamma}_{MKT}$ | $\hat{\gamma}_{\pi^L}$ | $\hat{\gamma}_{KUNC}$ |
| Estimates: | $0.0090^{***}$ | 0.7882 | $-7.7256^{*}$ | $19.2047^{***}$ | $0.0126^{***}$ | -0.3236 | $-14.5029^{**}$ | $32.3734^{***}$ |
| $t_{FM}$ | 6.7836 | 0.8193 | -2.3685 | 3.3757 | 8.0432 | -0.2229 | -3.7490 | 4.9754 |
| $t_S$ | 6.6064 | 0.7976 | -2.3003 | 3.2693 | 7.5256 | -0.2085 | -3.4693 | 4.5664 |
| $t_{JW}$ | 5.8953 | 0.7113 | -2.0149 | 2.9003 | 7.1480 | -0.2001 | -2.9192 | 3.9296 |
| $t_{KRS}$ | 5.1524 | 0.6262 | -1.5111 | 2.1707 | 6.0201 | -0.1846 | -2.5661 | 3.4372 |

The table reports CSR estimates $\hat{\gamma}$ and $t$-stats obtained by using factor covariances instead of betas in the first-pass time series regressions (see Kan, Robotti, and Shanken 2013). The following empirical specification of the learning under ambiguity model is used:

$$E\left[R^e_{i,t+1}\right] = \lambda_0 + \widehat{COV}_{i,MKT}\gamma_{MKT} + \widehat{COV}_{i,\pi^L}\gamma_{\pi^L} + \widehat{COV}_{i,KUNC}\gamma_{KUNC} + u_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are returns of tracking economic portfolios on "orthogonalized" innovations. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-MacBeth (1973 ) $t$-stats assuming the model is correctly specified ($t_{FM}$); Shanken's (1992) adjusted $t$-stats to correct for the EIV problem ($t_S$); Jagannathan and Wang (1998) adjusted $t$-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors; and Kan, Robotti, and Shanken (2013) robust $t$-stats under model misspecification ($t_{KRS}$). ($^{***}$) denotes 1% significance level; ($^{**}$) 5% significance level; and ($^{*}$) 10% significance level.

**Table 5**
**Incremental explanatory power of *KUNC* vs. the Bayesian model (1927-2007)**

| **Panel A: Bayesian Model** | $\hat{\boldsymbol{\lambda}}_{\mathbf{0}}$ | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{MKT}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}(\mathbf{1}-\hat{\boldsymbol{\pi}})}$ | |
|---|---|---|---|---|---|
| Estimates: | 0.0157*** | -0.0057 | -0.0095* | -0.0012 | |
| $t_{FM}$ | 5.9111 | -1.8511 | -3.0903 | -1.4549 | |
| $t_S$ | 5.4787 | -1.7555 | -2.8841 | -1.3620 | |
| $t_{JW}$ | 4.7633 | -1.5200 | -2.7381 | -1.3044 | |
| $t_{KRS}$ | 3.5296 | -1.1938 | -1.7252 | -0.9396 | |
| $\rho^2$ | 0.1419 | | | | |
| $p(\rho^2 = 1)$ | 0.0050 | | | | |
| $p(\rho^2 = 0)$ | 0.1689 | | | | |
| $Wald$ | 8.1768 | | | | |
| $p(Wald)$ | 0.2423 | | | | |
| **Panel B: Bayesian Model and Ambiguity** | $\hat{\boldsymbol{\lambda}}_{\mathbf{0}}$ | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{MKT}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}(\mathbf{1}-\hat{\boldsymbol{\pi}})}$ | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{KUNC}}$ |
| Estimates: | 0.0167*** | -0.0068 | -0.0046 | -0.0002 | 0.0027*** |
| $t_{FM}$ | 6.2404 | -2.1836 | -1.3795 | -0.2693 | 3.6704 |
| $t_S$ | 6.0010 | -2.1253 | -1.3307 | -0.2602 | 3.6465 |
| $t_{JW}$ | 5.3802 | -1.8936 | -1.2639 | -0.2351 | 3.6158 |
| $t_{KRS}$ | 4.3014 | -1.5742 | -0.9256 | -0.1887 | 3.5917 |
| $\rho^2$ | 0.3020 | | | | |
| $p(\rho^2 = 1)$ | 0.0010 | | | | |
| $p(\rho^2 = 0)$ | 0.0074 | | | | |
| $Wald$ | 21.6935 | | | | |
| $p(Wald)$ | 0.0002 | | | | |

The table reports comparative cross-sectional and GRS asset pricing tests between the three-factor Bayesian model and an augmented version that includes the ambiguity measure *KUNC*. The two models are:

$$\underline{\text{Bayesian model}}: E\left[R^e_{i,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\hat{\pi}}\lambda_{\hat{\pi}} + \hat{\beta}_{i,\hat{\pi}(1-\hat{\pi})}\lambda_{\hat{\pi}(1-\hat{\pi})} + \nu_i.$$

$$\underline{\text{Bayesian model and ambiguity}}: E\left[R^e_{n,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\hat{\pi}}\lambda_{\hat{\pi}} + \hat{\beta}_{i,\hat{\pi}(1-\hat{\pi})}\lambda_{\hat{\pi}(1-\hat{\pi})} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i.$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are returns of tracking economic portfolios on orthogonalized innovations. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the *p*-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 1$; $p(\rho^2 = 0)$ is the *p*-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 0$. *Wald* is the result of the joint test of the null hypothesis $H_0: \hat{\lambda} = 0$ with *p*-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) *t*-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted *t*-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted *t*-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken's (2013) robust *t*-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.

**Table 6**
**Incremental explanatory power of *KUNC* vs. the Bayesian model (1962-2007)**

| **Panel A: Bayesian Model** | $\hat{\boldsymbol{\lambda}}_{\mathbf{0}}$ | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{MKT}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}}$ | $\hat{\boldsymbol{\lambda}}_{\hat{\boldsymbol{\pi}}(\mathbf{1}-\hat{\boldsymbol{\pi}})}$ | |
|---|---|---|---|---|---|
| Estimates: | 0.0158*** | -0.0068** | -0.0094 | 0.0006 | |
| $t_{FM}$ | 7.1166 | -2.3839 | -2.4526 | 0.6842 | |
| $t_S$ | 6.4931 | -2.2565 | -2.2537 | 0.6370 | |
| $t_{JW}$ | 6.5157 | -2.2374 | -2.1306 | 0.6172 | |
| $t_{KRS}$ | 6.0007 | -2.0837 | -1.3682 | 0.4951 | |
| $\rho^2$ | 0.2805 | | | | |
| $p(\rho^2 = 1)$ | 0.0080 | | | | |
| $p(\rho^2 = 0)$ | 0.0286 | | | | |
| $Wald$ | 15.3898 | | | | |
| $p(Wald)$ | 0.0015 | | | | |
| **Panel B: Bayesian Model and Ambiguity** | | | | | |
| Estimates: | 0.0169*** | -0.0078** | -0.0039 | 0.0011 | 0.0025*** |
| $t_{FM}$ | 7.5156 | -2.7088 | -0.9018 | 1.2291 | 2.8225 |
| $t_S$ | 7.1228 | -2.6235 | -0.8576 | 1.1786 | 2.7957 |
| $t_{JW}$ | 7.0906 | -2.5919 | -0.8333 | 1.1891 | 2.7774 |
| $t_{KRS}$ | 6.6118 | -2.3964 | -0.5471 | 0.9326 | 2.7547 |
| $\rho^2$ | 0.4095 | | | | |
| $p(\rho^2 = 1)$ | 0.0138 | | | | |
| $p(\rho^2 = 0)$ | 0.0039 | | | | |
| $Wald$ | 24.9766 | | | | |
| $p(Wald)$ | 0.0000 | | | | |

The table reports comparative cross-sectional and GRS asset pricing tests between the three-factor Bayesian model and an augmented version that includes the proxy for Knightian uncertainty. The two models are:

$$\underline{\text{Bayesian model}}\text{: } E\left[R^e_{i,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\hat{\pi}}\lambda_{\hat{\pi}} + \hat{\beta}_{i,\hat{\pi}(1-\hat{\pi})}\lambda_{\hat{\pi}(1-\hat{\pi})} + \nu_i.$$

$$\underline{\text{Bayesian model and ambiguity}}\text{: } E\left[R^e_{n,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\hat{\pi}}\lambda_{\hat{\pi}} + \hat{\beta}_{i,\hat{\pi}(1-\hat{\pi})}\lambda_{\hat{\pi}(1-\hat{\pi})} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i.$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are returns of tracking economic portfolios on orthogonalized innovations. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the *p*-value of the asset pricing model specification test with null hypothesis $H_0\colon \rho^2 = 1$; $p(\rho^2 = 0)$ is the *p*-value of the asset pricing model specification test with null hypothesis $H_0\colon \rho^2 = 0$. *Wald* is the result of the joint test of the null hypothesis $H_0\colon \hat{\lambda} = 0$ with *p*-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) *t*-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted *t*-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted *t*-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken's (2013) robust *t*-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.

**Table 7**
**Ambiguity and intertemporal risk effects**

| **Panel A: 1927-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|---|---|---|
| Estimates: | 0.0074*** | 0.0029 | -0.0236* | 0.0260 | 0.0096* | 0.0008 | 0.0021*** |
| $t_{FM}$ | 4.9020 | 1.2650 | -2.8865 | 1.1560 | 3.1885 | 0.8365 | 3.3024 |
| $t_S$ | 4.2680 | 1.1848 | -2.5289 | 1.0181 | 2.7927 | 0.7460 | 3.2310 |
| $t_{JW}$ | 4.0923 | 1.1431 | -2.5986 | 1.0558 | 2.8395 | 0.7400 | 3.4017 |
| $t_{KRS}$ | 3.1944 | 0.9939 | -1.6274 | 0.7569 | 1.6478 | 0.5984 | 3.2703 |
| $\rho^2$ | 0.2073 | | | | | | |
| $p(\rho^2 = 1)$ | 0.0028 | | | | | | |
| $p(\rho^2 = 0)$ | 0.0342 | | | | | | |
| $Wald$ | 25.7683 | | | | | | |
| $p(Wald)$ | 0.0046 | | | | | | |
| **Panel B: 1962-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{KUNC}$ |
| Estimates: | 0.0116*** | -0.0019 | -0.0041 | -0.0401 | 0.0094** | 0.0024 | 0.0027*** |
| $t_{FM}$ | 6.7044 | -0.7398 | -0.4364 | -1.7755 | 3.3294 | 2.0468 | 3.1331 |
| $t_S$ | 5.9338 | -0.6959 | -0.3900 | -1.5906 | 2.9866 | 1.8460 | 3.0792 |
| $t_{JW}$ | 6.1394 | -0.7101 | -0.3986 | -1.6500 | 2.9402 | 1.8682 | 3.1287 |
| $t_{KRS}$ | 4.6436 | -0.6040 | -0.2585 | -1.0717 | 2.2464 | 1.4207 | 2.9958 |
| $\rho^2$ | 0.2383 | | | | | | |
| $p(\rho^2 = 1)$ | 0.0013 | | | | | | |
| $p(\rho^2 = 0)$ | 0.0094 | | | | | | |
| $Wald$ | 29.3949 | | | | | | |
| $p(Wald)$ | 0.0002 | | | | | | |

This table reports cross-sectional regression (CSR) estimates and asset pricing robust test results for the following empirical specification of the learning under ambiguity model:

$$E\left[R^e_{i,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,RF}\lambda_{RF} + \hat{\beta}_{i,DIV}\lambda_{DIV} + \hat{\beta}_{i,TERM}\lambda_{TERM} + \hat{\beta}_{i,DEF}\lambda_{DEF} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are returns of tracking economic portfolios on orthogonalized innovations. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the *p*-value of the asset pricing model specification test with null hypothesis $H_0$: $\rho^2 = 1$; $p(\rho^2 = 0)$ is the *p*-value of the asset pricing model specification test with null hypothesis $H_0$: $\rho^2 = 0$. *Wald* is the result of the joint test of the null hypothesis $H_0$: $\hat{\lambda} = 0$ with *p*-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) *t*-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted *t*-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted *t*-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken (2013) robust *t*-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.

**Table 8**
**Ambiguity and intertemporal risk effects using innovations**

| **Panel A: 1927-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|---|---|---|
| Estimates: | 0.0071*** | 0.0003 | -0.0229 | 0.0002 | 0.1159* | 0.0113 | 0.0024*** |
| $t_{FM}$ | 5.7869 | 0.3742 | -2.5869 | 1.1318 | 3.1863 | 0.9554 | 3.2390 |
| $t_S$ | 5.0476 | 0.3313 | -2.2682 | 0.9997 | 2.7956 | 0.8537 | 3.1638 |
| $t_{JW}$ | 5.0226 | 0.3442 | -2.3312 | 1.0267 | 2.8464 | 0.8650 | 3.2970 |
| $t_{KRS}$ | 3.6777 | 0.2038 | -1.3247 | 0.7507 | 1.6395 | 0.6945 | 3.1405 |
| $\rho^2$ | 0.2011 | | | | | | |
| $p(\rho^2 = 1)$ | 0.0023 | | | | | | |
| $p(\rho^2 = 0)$ | 0.0955 | | | | | | |
| $Wald$ | 18.4820 | | | | | | |
| $p(Wald)$ | 0.0003 | | | | | | |
| **Panel B: 1962-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{KUNC}$ |
| Estimates: | 0.0099*** | -0.0001 | -0.0019 | -0.0004 | 0.1109** | 0.0268 | 0.0027*** |
| $t_{FM}$ | 6.8382 | -0.2592 | -0.2014 | -2.0463 | 3.2188 | 1.8713 | 3.1099 |
| $t_S$ | 6.0557 | -0.2337 | -0.1801 | -1.8385 | 2.8879 | 1.6873 | 3.0530 |
| $t_{JW}$ | 6.1904 | -0.2497 | -0.1847 | -1.8991 | 2.8432 | 1.7292 | 3.1801 |
| $t_{KRS}$ | 4.9833 | -0.1825 | -0.1212 | -1.3143 | 2.1580 | 1.2568 | 3.0279 |
| $\rho^2$ | 0.2078 | | | | | | |
| $p(\rho^2 = 1)$ | 0.0003 | | | | | | |
| $p(\rho^2 = 0)$ | 0.0290 | | | | | | |
| $Wald$ | 21.3418 | | | | | | |
| $p(Wald)$ | 0.0001 | | | | | | |

The table reports cross-sectional regression (CSR) estimates and asset pricing robust test results for the following empirical specification of the learning under ambiguity model:

$$E[R^e_{i,t+1}] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,RF}\lambda_{RF} + \hat{\beta}_{i,DIV}\lambda_{DIV} + \hat{\beta}_{i,TERM}\lambda_{TERM} + \hat{\beta}_{i,DEF}\lambda_{DEF} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are orthogonalized innovations from a VAR. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the $p$-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 1$; $p(\rho^2 = 0)$ is the $p$-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 0$. $Wald$ is the result of the joint test of the null hypothesis $H_0: \hat{\lambda} = 0$ with $p$-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) $t$-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted $t$-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted $t$-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken (2013) robust $t$-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.

**Table 9**
**Incremental explanatory power of empirically-motivated factors vs. ambiguity**

| **Panel A: 1927-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{SMB}$ | $\hat{\lambda}_{HML}$ | $\hat{\lambda}_{UMD}$ | $\hat{\lambda}_{\pi^L}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimates: | 0.0068*** | 0.0037 | -0.0259 | 0.0280 | 0.0085 | 0.0009 | 0.0017 | 0.0032** | 0.0067* | 0.0020 | 0.0033*** |
| $t_{FM}$ | 4.1189 | 1.5139 | -2.9032 | 1.2002 | 2.6095 | 0.8886 | 1.2359 | 2.4757 | 2.2463 | 1.5203 | 4.0105 |
| $t_S$ | 3.5237 | 1.3925 | -2.4981 | 1.0391 | 2.2451 | 0.7787 | 1.1497 | 2.3699 | 1.9864 | 1.4794 | 3.7828 |
| $t_{JW}$ | 3.2926 | 1.3570 | -2.6030 | 1.0652 | 2.2170 | 0.8145 | 1.1331 | 2.3993 | 2.0670 | 1.4543 | 4.0796 |
| $t_{KRS}$ | 2.3402 | 1.0604 | -1.5173 | 0.8008 | 1.2687 | 0.6659 | 0.8719 | 2.0502 | 1.5332 | 1.3558 | 3.4746 |
| $\rho^2$ | 0.3055 | | | | | | | | | | |
| $p(\rho^2 = 1)$ | 0.0053 | | | | | | | | | | |
| $p(\rho^2 = 0)$ | 0.0453 | | | | | | | | | | |
| $Wald$ | 31.4597 | | | | | | | | | | |
| $p(Wald)$ | 0.0000 | | | | | | | | | | |

| **Panel B: 1962-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{SMB}$ | $\hat{\lambda}_{HML}$ | $\hat{\lambda}_{UMD}$ | $\hat{\lambda}_{\pi^L}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimates: | 0.0094*** | 0.0005 | -0.0017 | -0.0386 | 0.0095** | 0.0019 | 0.0022 | 0.0045*** | 0.103** | 0.0008 | 0.0042*** |
| $t_{FM}$ | 5.0071 | 0.1790 | -0.1805 | -1.6818 | 3.1028 | 1.5319 | 1.4590 | 3.3921 | 3.4640 | 0.5435 | 4.4841 |
| $t_S$ | 4.2436 | 0.1632 | -0.1549 | -1.4473 | 2.6696 | 1.3264 | 1.3873 | 3.3034 | 3.0772 | 0.5319 | 4.2689 |
| $t_{JW}$ | 4.2446 | 0.1656 | -0.1585 | -1.5534 | 2.7214 | 1.2928 | 1.4364 | 3.2237 | 3.0260 | 0.5211 | 4.4455 |
| $t_{KRS}$ | 3.3343 | 0.1412 | -0.1061 | -1.0478 | 1.9459 | 0.9972 | 1.2599 | 3.0183 | 2.6705 | 0.5154 | 4.0159 |
| $\rho^2$ | 0.3598 | | | | | | | | | | |
| $p(\rho^2 = 1)$ | 0.0155 | | | | | | | | | | |
| $p(\rho^2 = 0)$ | 0.0050 | | | | | | | | | | |
| $Wald$ | 41.1008 | | | | | | | | | | |
| $p(Wald)$ | 0.0000 | | | | | | | | | | |

The table reports cross-sectional regression (CSR) estimates and asset pricing robust test results for the following empirical specification of the learning under ambiguity model:

$$E\left[R^e_{n,t+1}\right] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,RF}\lambda_{RF} + \hat{\beta}_{i,DIV}\lambda_{DIV} + \hat{\beta}_{i,TERM}\lambda_{TERM} + \hat{\beta}_{i,DEF}\lambda_{DEF} + \\ + \hat{\beta}_{i,SMB}\lambda_{SMB} + \hat{\beta}_{i,HML}\lambda_{HML} + \hat{\beta}_{i,UMD}\lambda_{UMD} + \hat{\beta}_{i,\pi^L}\lambda_{\pi^L} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are returns of tracking economic portfolios on orthogonalized innovations. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the *p*-value of the asset pricing model specification test with null hypothesis $H_0\colon \rho^2 = 1$; $p(\rho^2 = 0)$ is the *p*-value of the asset pricing model specification test with null hypothesis $H_0\colon \rho^2 = 0$. $Wald$ is the result of the joint test of the null hypothesis $H_0\colon \hat{\lambda} = 0$ with *p*-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) *t*-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted *t*-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted *t*-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken (2013) robust *t*-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.

**Table 10**
**Incremental explanatory power of empirically-motivated factors vs. ambiguity using innovations**

| **Panel A: 1927-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{SMB}$ | $\hat{\lambda}_{HML}$ | $\hat{\lambda}_{UMD}$ | $\hat{\lambda}_{KUNC}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Estimates: | 0.0078*** | 0.0024 | -0.0267 | -0.0002 | 0.0098 | 0.0128 | 0.0002 | 0.0022 | -0.0018 | 0.0023*** |
| $t_{FM}$ | 5.1767 | 1.0233 | -3.0181 | -0.4137 | 2.5577 | 1.0065 | 0.2141 | 2.7002 | -1.7590 | 3.0662 |
| $t_S$ | 4.3740 | 0.9439 | -2.5661 | -0.3528 | 2.1748 | 0.8720 | 0.1845 | 2.3164 | -1.5152 | 2.9709 |
| $t_{JW}$ | 4.5181 | 0.9755 | -2.6479 | -0.3744 | 2.1853 | 0.9283 | 0.2027 | 2.4229 | -1.7030 | 3.0314 |
| $t_{KRS}$ | 3.4163 | 0.8222 | -1.5674 | -0.2491 | 1.3194 | 0.7331 | 0.1276 | 1.4873 | -1.0032 | 2.9262 |
| $\rho^2$ | 0.2640 | | | | | | | | | |
| $p(\rho^2 = 1)$ | 0.0067 | | | | | | | | | |
| $p(\rho^2 = 0)$ | 0.0789 | | | | | | | | | |
| $Wald$ | 31.7379 | | | | | | | | | |
| $p(Wald)$ | 0.0002 | | | | | | | | | |
| **Panel B: 1962-2007** | $\hat{\lambda}_0$ | $\hat{\lambda}_{MKT}$ | $\hat{\lambda}_{RF}$ | $\hat{\lambda}_{DIV}$ | $\hat{\lambda}_{TERM}$ | $\hat{\lambda}_{DEF}$ | $\hat{\lambda}_{SMB}$ | $\hat{\lambda}_{HML}$ | $\hat{\lambda}_{UMD}$ | $\hat{\lambda}_{KUNC}$ |
| Estimates: | 0.0119*** | -0.0023 | -0.0062 | -0.0004 | 0.1055** | 0.0216 | -0.0008 | 0.0008 | 0.0001 | 0.0030*** |
| $t_{FM}$ | 6.6646 | -0.8945 | -0.6347 | -1.7727 | 3.0062 | 1.4553 | -0.9898 | 0.9089 | 0.0769 | 3.3408 |
| $t_S$ | 5.8884 | -0.8393 | -0.5659 | -1.5882 | 2.6902 | 1.3079 | -0.8905 | 0.8145 | 0.0691 | 3.2710 |
| $t_{JW}$ | 6.0750 | -0.8544 | -0.5929 | -1.7111 | 2.6727 | 1.2866 | -0.9355 | 0.8442 | 0.0708 | 3.3944 |
| $t_{KRS}$ | 4.5734 | -0.7306 | -0.3688 | -1.1437 | 2.0289 | 0.9553 | -0.6911 | 0.5044 | 0.0455 | 3.1962 |
| $\rho^2$ | 0.2593 | | | | | | | | | |
| $p(\rho^2 = 1)$ | 0.0008 | | | | | | | | | |
| $p(\rho^2 = 0)$ | 0.0424 | | | | | | | | | |
| $Wald$ | 33.4840 | | | | | | | | | |
| $p(Wald)$ | 0.0001 | | | | | | | | | |

The table reports cross-sectional regression (CSR) estimates and asset pricing robust test results for the following empirical specification of the learning under ambiguity model:

$$E[R^e_{n,t+1}] = \lambda_0 + \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,RF}\lambda_{RF} + \hat{\beta}_{i,DIV}\lambda_{DIV} + \hat{\beta}_{i,TERM}\lambda_{TERM} + \hat{\beta}_{i,DEF}\lambda_{DEF} + \\ + \hat{\beta}_{i,SMB}\lambda_{SMB} + \hat{\beta}_{i,HML}\lambda_{HML} + \hat{\beta}_{i,UMD}\lambda_{UMD} + \hat{\beta}_{i,\pi^L}\lambda_{\pi^L} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The set of test assets includes 55 portfolios ranked by size, book-to-market, and industry. Factors are orthogonalized innovations from a VAR. $\rho^2$ is the cross-sectional goodness of fit statistic with $p(\rho^2 = 1)$ denoting the *p*-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 1$; $p(\rho^2 = 0)$ is the *p*-value of the asset pricing model specification test with null hypothesis $H_0: \rho^2 = 0$ . *Wald* is the result of the joint test of the null hypothesis $H_0: \hat{\lambda} = 0$ with *p*-value $p(Wald)$. Following the approach in Kan, Robotti, and Shanken (2013), we report standard Fama-Macbeth (1973) *t*-stats assuming the model is correctly specified ($t_{FM}$), Shanken's (1992) adjusted *t*-stats to correct for the EIV problem ($t_S$), Jagannathan and Wang (1998) adjusted *t*-stats ($t_{JW}$) correcting Shanken's test for the distribution of the errors, and Kan, Robotti, and Shanken (2013) robust *t*-stats under model misspecification ($t_{KRS}$). (***) denotes significance at the 1%, (**) at the 5% level, and (*) at the 10% level.

**Table 11**
**Comparative tests of CSR $\rho^2$s: Bayesian uncertainty vs. the three-factor learning under ambiguity model**

| | Learning Under Ambiguity Model | |
|---|---|---|
| | Period: 1927-2007 | Period: 1962-2007 |
| Augmented Model | -0.0123 | -0.0988 |
| p-value(1) | (0.6367) | (0.0846) |
| p-value(2) | (0.6484) | (0.1240) |

The table reports pair-wise tests of equality of the GLS cross-sectional $\rho^2$ between the Augmented Model (AUG) and the Learning Under Ambiguity (LUA) model for two sample periods. The null hypothesis is $H_0: \hat{\rho}^2_{AUG} - \hat{\rho}^2_{LUA} = 0$ with *p*-values in parenthesis computed without [p-value (1)] and with [p-value (2)] the assumption of model misspecification. The Learning Under Ambiguity model includes *MKT*, *DIV*, *TERM*, *DEF*, and *KUNC* as risk factors. The Augmented model includes the same factors as the Learning Under Ambiguity in addition to the Bayesian proxy for uncertainty. The models are estimated using monthly returns on 25 Fama-French portfolios ranked by size and book-to-market. The sample periods are from 1927-2007 (969 observations) and 1962-2007 (552 observations).

**Table 12**
**Comparative tests of CSR $\rho^2$s: the intertemporal CAPM, the Fama-French four-factor model and learning under ambiguity model**

| Panel A: 1927-2007 | | |
|---|---|---|
| | Column Model | |
| **Row Model** | **FF-4** | **Learning Under Ambiguity** |
| **ICAPM** | 0.2242 | -0.0320 |
| p-value(1) | (0.0764) | (0.9943) |
| p-value(2) | (0.0000) | (0.1244) |
| **FF-4** | | -0.2562 |
| p-value(1) | | (0.0683) |
| p-value(2) | | (0.0000) |
| **Panel B: 1962-2007** | | |
| | **FF-4** | **Learning Under Ambiguity** |
| **ICAPM** | 0.0590 | -0.0320 |
| p-value(1) | (0.1406) | (0.2411) |
| p-value(2) | (0.0827) | (0.1846) |
| **FF-4** | | -0.2562 |
| p-value(1) | | (0.4199) |
| p-value(2) | | (0.0001) |

The table reports pair-wise tests of equality of the GLS cross-sectional $\rho^2$ between each of the two "row" models (Model *i*) and each of the two "column" models (Model *j*) with null $H_0\colon \hat{\rho}_i^2 - \hat{\rho}_j^2 = 0$ with p-values in parenthesis computed without [p-value (1)] and with [p-value (2)] the assumption of model misspecification. The FF-4 model includes *MKT*, *SMB*, *HML*, and *UMD* as risk factors,. The ICAPM includes *MKT*, *RF*, *DIV*, *TERM*, and *DEF* as risk factors. The R-DAPM includes *MKT*, *RF*, *DIV*, *TERM*, *DEF*, and *KUNC* as risk factors. The models are estimated using monthly returns on 25 Fama-French portfolios ranked by size and book-to-market. The sample periods are from 1927-2007 (969 observations) and 1962-2007 (552 observations).

**Table 13**
**GLS-CSR robustness checks**

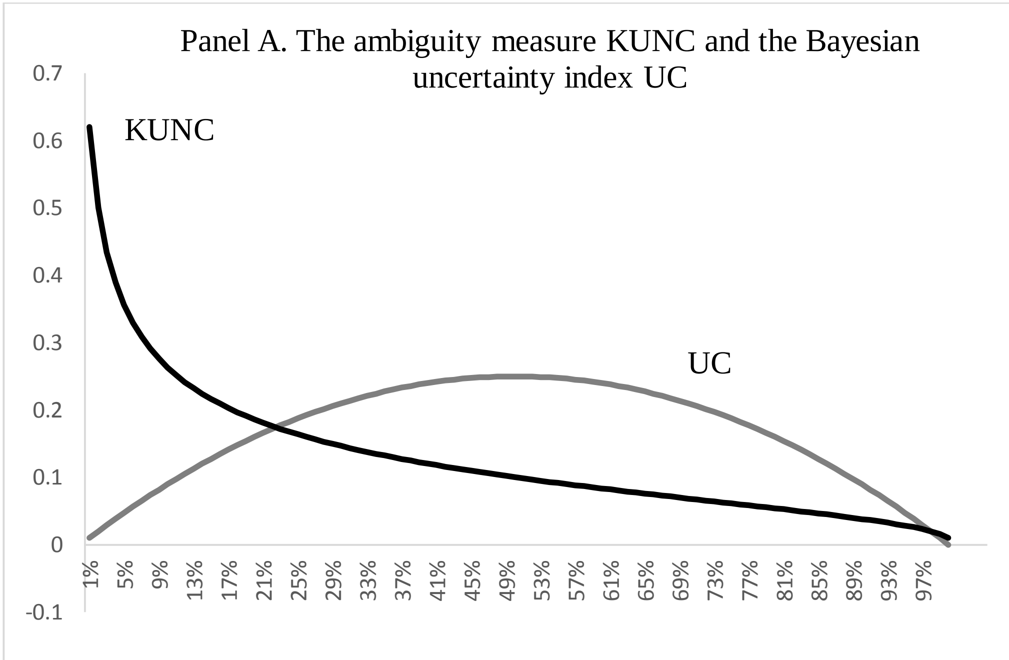
| | | | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{MKT}}$ | | | | | | *t-stat* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| WL/CL | 90 | 80 | 70 | 60 | 50 | WL/CL | 90 | 80 | 70 | 60 | 50 |
| 180 | 0.0072 | 0.0072 | 0.0072 | 0.0072 | 0.0072 | 180 | **4.39** | **4.38** | **4.38** | **4.38** | **4.38** |
| 120 | 0.0077 | 0.0077 | 0.0077 | 0.0077 | 0.0077 | 120 | **3.77** | **3.76** | **3.76** | **3.76** | **3.76** |
| 96 | 0.0070 | 0.0070 | 0.0070 | 0.0070 | 0.0070 | 96 | **3.41** | **3.41** | **3.41** | **3.41** | **3.41** |
| 60 | 0.0070 | 0.0070 | 0.0070 | 0.0070 | 0.0070 | 60 | **3.53** | **3.53** | **3.53** | **3.53** | **3.52** |
| 36 | 0.0071 | 0.0071 | 0.0071 | 0.0071 | 0.0071 | 36 | **3.60** | **3.60** | **3.60** | **3.60** | **3.60** |
| | | | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{\pi^L}}$ | | | | | | *t-stat* | | |
| WL/CL | 90 | 80 | 70 | 60 | 50 | WL/CL | 90 | 80 | 70 | 60 | 50 |
| 180 | 0.0346 | 0.0339 | 0.0336 | 0.0335 | 0.0335 | 180 | 0.66 | 0.64 | 0.64 | 0.64 | 0.64 |
| 120 | 0.0608 | 0.0604 | 0.0601 | 0.0599 | 0.0597 | 120 | 1.19 | 1.18 | 1.18 | 1.18 | 1.17 |
| 96 | 0.0968 | 0.0960 | 0.0951 | 0.0942 | 0.0933 | 96 | **2.11** | **2.09** | **2.08** | **2.06** | **2.04** |
| 60 | 0.0864 | 0.0865 | 0.0867 | 0.0868 | 0.0869 | 60 | **2.30** | **2.31** | **2.31** | **2.32** | **2.32** |
| 36 | 0.0590 | 0.0604 | 0.0613 | 0.0618 | 0.0622 | 36 | **2.19** | **2.24** | **2.27** | **2.29** | **2.30** |
| | | | $\hat{\boldsymbol{\lambda}}_{\boldsymbol{KUNC}}$ | | | | | | *t-stat* | | |
| WL/CL | 90 | 80 | 70 | 60 | 50 | WL/CL | 90 | 80 | 70 | 60 | 50 |
| 180 | 0.0069 | 0.0135 | 0.0195 | 0.0248 | 0.0297 | 180 | **5.20** | **5.18** | **5.13** | **5.03** | **4.89** |
| 120 | 0.0057 | 0.0111 | 0.0160 | 0.0206 | 0.0247 | 120 | **4.16** | **4.15** | **4.10** | **4.00** | **3.83** |
| 96 | 0.0047 | 0.0095 | 0.0143 | 0.0190 | 0.0236 | 96 | **3.57** | **3.72** | **3.80** | **3.81** | **3.75** |
| 60 | 0.0033 | 0.0072 | 0.0115 | 0.0162 | 0.0211 | 60 | **2.52** | **2.82** | **3.04** | **3.14** | **3.16** |
| 36 | 0.0015 | 0.0041 | 0.0075 | 0.0116 | 0.0159 | 36 | 1.19 | 1.69 | **2.10** | **2.35** | **2.46** |

| **CL** | **90** | **80** | **70** | **60** | **50** |
|---|---|---|---|---|---|
| $\hat{Q}$ | 34.41 | 34.36 | 34.32 | 34.29 | 34.25 |
| p-value | 0.10 | 0.10 | 0.10 | 0.11 | 0.11 |
| Adj. $\hat{Q}$ | 48.50 | 48.44 | 48.38 | 48.33 | 48.95 |
| p-value | 0.07 | 0.07 | 0.07 | 0.08 | 0.08 |

The table reports the second-pass Fama-MacBeth (1973) GLS-CSR test results on excess returns on 25 portfolio returns sorted by size and book-to-market:

$$E\left[R_{i,t+1}^{e}\right] = \hat{\beta}_{i,MKT}\lambda_{MKT} + \hat{\beta}_{i,\pi^L}\lambda_{\pi^L} + \hat{\beta}_{i,KUNC}\lambda_{KUNC} + \nu_i$$

The sample period is 1927-2007. The factors are measured in levels. The results are presented across different confidence levels (CL) ranging from 50%-90%, and different lengths in the rolling regression windows (WL) ranging from 36 months to 180 months. The GLS-CSR test has null hypothesis $H_0\colon \hat{Q} = T\bar{\alpha}'(\bar{V}_{22}^{-1})\bar{\alpha} = 0$, with $\widehat{\boldsymbol{\alpha}} = \overline{\mathbf{R} - R_f} - \hat{\beta}\boldsymbol{\lambda}$. Individual $t$-statistics are presented with Shanken's (2009) correction. $\hat{Q}$ are presented with and without Shanken's correction. Highlighted $t$-stats in black denote statistical significance at the 5% level.



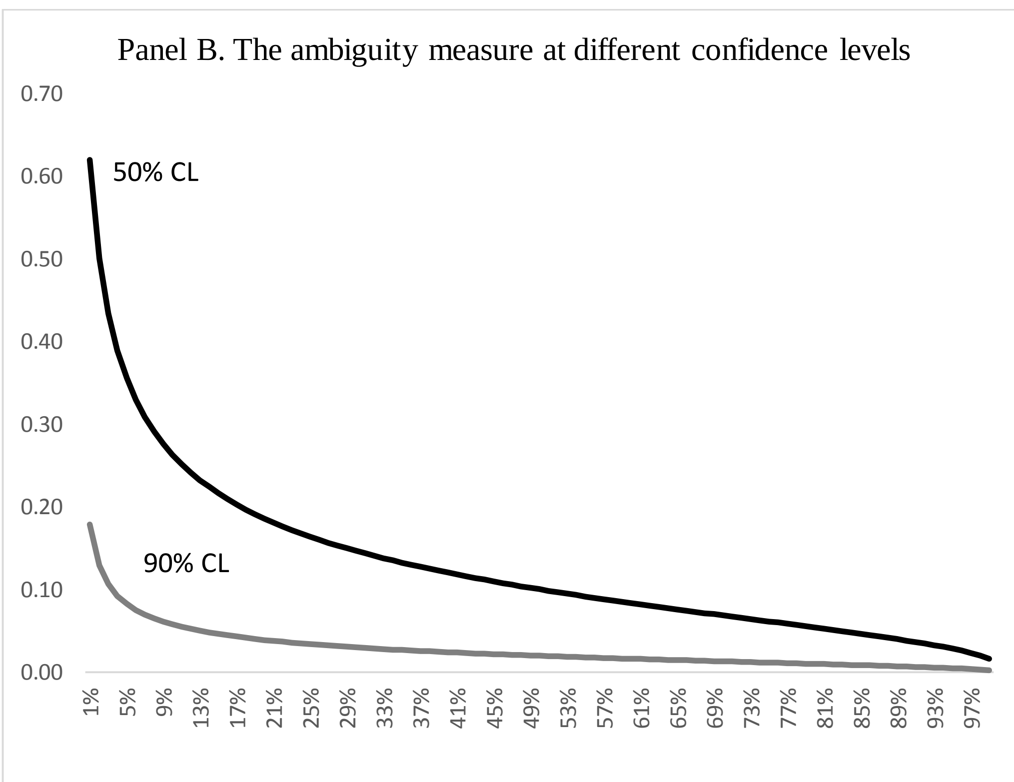




**Figure 1. The ambiguity measure.**

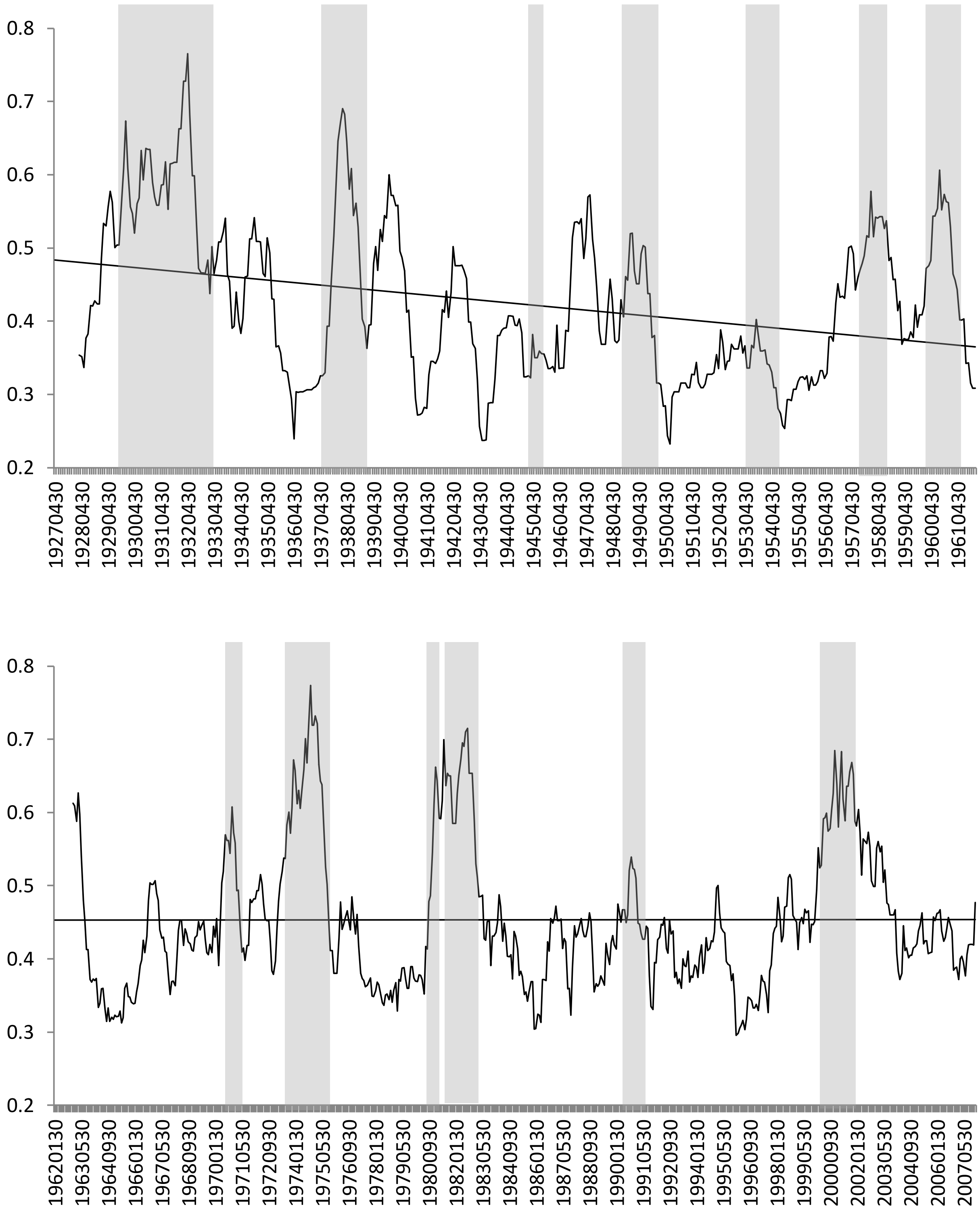



**Figure 2. *KUNC* (CL = 50%, T = 36 months).** The top chart plots the 12-month moving average of the factor *KUNC* from 1927-1961. The bottom chart plots the 12-month moving average of the factor *KUNC* from 1962-2007.