A Simple Robust Single Market Factor Model
Under Statistical Ambiguity

Ariel M. Viale*

This version: January 2017

*Corresponding author: Ariel M. Viale, Marshall E. Rinker School of Business, Palm Beach Atlantic University, 901 South Flagler Dr., West Palm Beach, FL 33416-4708, E-mail: ariel_viale@pba.edu. I would like to thank Minhao Cai and Paulo Maio for their valuable comments. The views expressed are those of the author. All remaining errors are my own responsibility.
A Simple Robust Single Market Factor Model
Under Statistical Ambiguity

Abstract
A well-known result in empirical asset pricing is that the single market factor representation of the CAPM displays idiosyncratic pricing errors (alphas) correlated with firm specific characteristics. Under statistical ambiguity, alphas may reflect investors’ difficulty to approximate a high dimensional state-space with a low dimensional functional. Using a robust econometric estimation method consistent with the multiple-priors approach of the ambiguity literature, I recover from a large cross-section of stock returns the value of the market premium that corresponds to investors’ worst case scenario in terms of indirect utility or wealth. This robust version of the market premium that includes a compensation for statistical ambiguity is statistically and economically significant without the aid of firm-specific factors. The conceptual approach and empirical results suggest that firm-specific factors reflect idiosyncratic ambiguity, which unlike idiosyncratic volatility, cannot be diversified away.

JEL Classification Codes: G12, C58.

Keywords: asset pricing, statistical ambiguity, robust Bayesian analysis, CAPM.
1 Introduction

There is substantial theoretical literature in finance and economics that sets the cornerstones of the economic behavior in equilibrium of a robust representative agent (RA) concerned with model misspecification and ambiguity or uncertainty in the sense of Knight (1921), Keynes (1921), Shackle (1949), and Roy (1952).¹ For an excellent and comprehensive review of the asset pricing literature under ambiguity see Epstein and Schneider (2010).

The hypothesis that statistical ambiguity and robustness should matter in empirical asset pricing tests of the CAPM is motivated by the observation that any empirical asset pricing test of the CAPM represents an example of an ill-posed problem (an under-determined problem in the words of Jaynes, 1984). The intuition that the empirical asset pricing test of the CAPM is an ill-posed problem goes as far back as the early critiques by Roll (1977), Shanken (1987), and Kandel and Stambaugh (1987). These authors argue that true asset pricing factors may never be observable or known and the use of any proxies questions the empirical testability of the CAPM.

Statistical ambiguity arises from the difficulty to estimate a low dimensional functional from observations presumed to be generated from an unknown high dimensional data generating process (DGP). Hansen (2007) coined this problem as statistical ambiguity.²

¹ Ellsberg (1961) and related experiments have provided sufficient evidence of agents' robust behavior or aversion to choices that involve uncertain subjective probabilities (ambiguity) rather than risky objective probabilities (for a survey see Camerer and Weber, 1992). In a static setting, Gilboa and Schmeidler (1989) and Schmeidler (1989), axiomatized the distinct behavior of economic agents concerned with ambiguity by introducing sets of multiple-priors and the MaxMin criteria when solving robust decision problems. Hansen and Sargent (2001, 2008), Chen and Epstein (2002), and Epstein and Schneider (2003, 2007, 2008), among others, extend this work to dynamic settings.

² Further evidence of statistical ambiguity in standard empirical asset pricing tests of the single market factor representation of the CAPM has manifested in the form of instability in regression estimates after some arbitrary change of parameterization. Indeed, as it has been widely documented in the empirical asset pricing literature, the market risk premium can go from being statistically insignificant if the intercept is included, to statistically significant and economically meaningful, with a value comparable to its realized historic average, if the intercept is excluded from the CAPM cross-sectional regression.
With this motivation in mind, I derive a simple robust single market factor representation of the CAPM under statistical ambiguity, and then assess its’ empirical capabilities in a large cross-section of stock returns. Clearly, it seems that in empirical asset pricing applications, the econometrician should use a methodological approach that explicitly accounts for statistical ambiguity. Thus, instead of using generalized least squares (GLS) in the second step of the classic two-pass Fama and MacBeth (1973), and Black, Jensen, and Scholes (1972) estimation procedure, I propose using generalized maximum entropy (GME).³

The proposed approach still falls in the class of simple linear beta pricing models (for a recent survey see e.g., Jagannathan, Schaumburg, and Zhou, 2010). One well known problem with orthodox beta form linear specifications is that within the test of the null hypothesis of zero mispricing (in the cross-section of expected returns) lies an implicit assumption made on the data generating process driving stock returns (in the time-series dimension). Under the proposed approach, the econometrician acknowledges that (misspecified) first-pass time-series regressions constitute only a reference model introducing mispricing errors that GME tackles naturally in the second pass cross-sectional regression independently of the source of the model misspecification.

Notice that this approach is different from other robust econometric methods like standard Bayesian (averaging) model uncertainty and parameter uncertainty. Under the proposed approach, focus is on the lack of confidence in the reference model rather than lack of information. Simply put, the robust econometrician (like the robust investor) gives up in trying to learn all the time-varying features driving stock returns and simply seek to recover the market

³ Kitamura and Stutzer (1997) propose maximum entropy (ME) as an information-theoretic alternative estimation procedure to GMM with better small sample properties in the class of optimal minimum distance (OMD) estimators. The benefits and capabilities of using GME in econometrics has been rigorously covered in Golan, Judge, and Miller (1996) and Mittelhammer, Judge, and Miller (2000).
price of risk and ambiguity that corresponds to the probability law that has maximum divergence or entropy with respect to the reference model.

The choice of GME as a suitable estimation procedure under ambiguity is further motivated by duality between the principle of maximizing entropy in robust Bayesian statistics and investors’ preference for robustness discounting always the worst-case scenario in terms of indirect utility or wealth. Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) show that economic decision problems under ambiguity represented by an entropy-based penalization function, under weak regularity conditions (i.e., dynamic consistency and a Dynkin space), have a dual representation as robust Bayesian statistical decision problems under prior uncertainty. Hence, the proposed approach puts the robust econometrician and robust investor on the same foot when it comes to price risky assets.

Bansal and Lehmann (1997) and Alvarez and Jermann (2005), show that mean stock excess returns place a lower bound on the entropy of any asset pricing kernel or model. Backus, Chernov, and Zin (2014) use this entropy bound to assess the dispersion and horizon dependence of competing asset pricing models. Sims (2003) and Van Nieuwerburgh and Veldkamp (2010), use generalized entropy to assess investors’ capacity to learn the true and hidden DGP driving stock returns. Ghosh, Julliard, and Taylor (2011) and Stutzer (1996) use entropy to assess the divergence between true and risk-neutral probabilities driving asset returns. Here, I use generalized maximum entropy to recover from a large cross-section of stock returns the market premium commensurate with the econometrician-investor’s worst case scenario under statistical ambiguity. In this regard, I provide an empirical response to Al Najjar’s (2009) critique that

---

4 Simply put, the investor chooses without regret the optimal allocation, and stock returns follow a memoryless mechanism defined on a compact state space (e.g., a Markov chain with some transition probability function across states). The Dynkin space that captures the notion of risk is the one where the set of priors collapses to a singleton (the reference model). The Dynkin space that captures the notion of ambiguity is the one with the largest possible set of priors or distorted models.
ambiguity in economics and finance lack empirical motivation, in the sense that the existing literature is unable to connect the multiple-priors’ specification to actual data.

To illustrate, take as example the asset pricing approach under ambiguity developed by Kogan and Wang (2003). They uncover conceptually a pricing equation in beta form that displays a second factor, distinct from market beta, which they refer to as the uncertainty factor. Intrinsic to the problem of assessing empirically the economic and statistical significance of the second factor using observable variables, lies the difficulty in distinguishing the model from an alternative multi-factor representation motivated for example by Merton’s (1973) intertemporal capital asset pricing model (ICAPM). This is exactly the kind of problem that Al Najjar (2009) points at. Under the proposed approach, the robust market premium corresponds one-to-one with the distribution of returns with maximum entropy given the reference model and set of observations.

The results of the asset pricing tests can be summarized as follows. First, a single market factor representation of the CAPM that accounts for robustness towards statistical ambiguity gives a maximum entropy market premium that is economically and statistically meaningful with a value between 0.86% and 1.56% per month (depending on the sample period chosen to run the asset pricing test). This result is consistent with the observation in Backus, Chernov, and Zin (2014) that historical excess returns suggest a one-period compensation for relative entropy about 1% monthly. This result is robust to different time periods, rolling regression windows, error in variables, and sample of test assets. The robust market premium sets the lower bound for stock prices that corresponds to investors’ worst-case scenario in terms of indirect utility or wealth.
I provide numerical evidence that size and value premia are significantly correlated with stocks’ idiosyncratic ambiguity. Consequently, under statistical ambiguity, size and value premia can be interpreted as compensations for stocks’ idiosyncratic ambiguity. The economic intuition lies in the fact that idiosyncratic ambiguity, unlike idiosyncratic volatility, cannot be diversified away and consequently shows up as a statistical characteristic in misspecified asset pricing models.

The rest of the paper proceeds as follows. In section 2, I discuss the conceptual foundation of the simple robust single market factor representation of the CAPM under statistical ambiguity and introduce the methodological approach used in the empirical asset pricing tests. In section 3, I discuss empirical results. Section 4 concludes the study. Proofs and technical details are relegated to Appendices.

2 Conceptual foundation

2.1 The fundamental asset pricing equation under ambiguity

In this section, I follow the asset pricing literature under ambiguity surveyed by Epstein and Schneider (2010). Consider the following discrete Markovian dynamic economy. For each path \(\omega_t\) in the event tree, the robust RA (from now on denoted as the robust investor) must choose the consumption plan \(C_t\) as well as the amount \(\xi_{i,t}\) of total wealth \(W_t\) allocated in \(i\) risky assets with gross returns \(R_i\) (for all \(i = \{1, \ldots, I\}\)) and the risk-free asset with constant gross return \(R_f\). At time \(t + 1\), investor’s conditional wealth is \(W_{t+1}^{s_t}\), where \(s_t\) denotes the state of the economy at time \(t\) given path \(\omega_t\).
A robust (averse to ambiguity) investor casts doubts about the one-step-ahead conditional probabilities for each plausible state.\(^5\) Holds some prior belief \(\hat{\pi}_t\) denoted as the reference model and entertains a set of plausible alternative distorted or perturbed models \(\pi_t^*\). If one of the distorted distributions corresponds to the true data generating process (DGP), as \(T \to \infty\) the expected log likelihood ratio between reference and distorted beliefs converges to the unconditional value of one period relative entropy \(D_t(\pi_t^* \| \hat{\pi}_t)\) (for technical details see the Appendix in Hansen and Sargent, 2008, section 2.3). In the ambiguity literature, the two models are assumed to be statistically close in terms of Kullback-Leibler divergence:\(^6\)

\[
D_t(\pi_t^* \| \hat{\pi}_t) \leq \eta_t, \tag{1}
\]

where \(\eta_t\) restricts the entropy-constrained ball (see Epstein and Schneider, 2010) containing all distorted models \(\pi_t^*\) that are statistically close to the reference model \(\hat{\pi}_t\).\(^7\)

The robust representative investor solves the consumption-investment problem using the following constrained Bellman backward dynamic recursion subject to (1) and the usual budget constraint:

\[
J(W_t, t) = \max_{C_t \in \{\xi_t\}} \{U(C_t) + \min_{\pi_t^*} E_t(\pi_t^* J(W^1_{t+1}, t+1) + (1 - \pi_t^*) J(W^0_{t+1}, t+1))\}, \tag{2}
\]

where \(E_t(\cdot)\) is the conditional expectation operator at time \(t = \{0, 1, \ldots, T-1\}\), \(U(C_t)\) is increasing and concave, and \(J(\cdot, t)\) is the usual indirect utility function. In the ambiguity literature, the conditional expectation on the right-hand side of (2) is interpreted as the ambiguity

---

\(^5\) Epstein and Schneider (2003) pp. 16-17 discuss the technical requirements needed to ensure dynamic consistency or the no regret principle in the entropy framework.

\(^6\) The reason is to avoid infinite pessimism. The Kullback-Leibler divergence is represented in statistics as the expected logarithm of the likelihood ratio between two distributions (see Cover and Thomas, 1991, Chapter 2 for details on the Kullback-Leibler divergence and its various interpretations). It is not a true distance as it is not symmetric and does not satisfy the triangular inequality.

\(^7\) This parameter is generally assumed to be exogenous. As shown below, here the optimal value of the parameter corresponds to the maximum entropy bound, which can be recovered from the cross-section of stock returns. Observe that \(\eta_t = 0\) implies the special case of an investor that is not concerned with ambiguity.
certainty equivalent of the continuation value function under the worst-case scenario in terms of indirect utility or wealth.\(^8\) The constrained MaxMin problem in (2) constitutes a standard specification in the ambiguity literature. It obtains for instance in static multiple-priors settings (see Epstein and Schneider, 2010, p 321) or in a dynamic setting as a specific recursive multiple-priors model (see Epstein and Schneider, 2003, pp. 16-17). It should be interpreted as if the economic agent plays a strategic game against Nature where Nature always chooses the state that corresponds to the (average) worst-case outcome in terms of loss of the economic agent’s indirect utility or wealth. Solving (2) recursively the following proposition results:

**Proposition 1** (See Appendix A for the proof): *The equilibrium fundamental asset pricing equation under ambiguity is:* 

\[
1 = E_t(\tilde{m}_{t,t+1}R_{t,t+1}),
\]

where \(R_{t,t+1}\) is the gross return on risky asset \(i\); and \(\tilde{m}_{t,t+1}\) is the stochastic discount factor or asset pricing kernel under the worst-case scenario probability measure \(\tilde{\pi}_t = \hat{\pi}_t - \sqrt{2\eta_t\hat{\pi}_t(1 - \hat{\pi}_t)}\).

2.2 A simple single market factor representation of the CAPM under statistical ambiguity

Deriving a single-factor linear representation of the pricing kernel is a standard exercise in the asset pricing literature that can be summarized as follows. If the joint distribution of stock returns and investors’ wealth is independent over time (i.e. a constant investment opportunity set), then

---

\(^8\) The investor solves the portfolio problem evaluating each candidate portfolio under the worst-case return distribution for that portfolio. See Epstein and Schneider (2010) Section 2.2 for technical details about the construction of the ambiguity certainty equivalent function.
one can take a Taylor series expansion of marginal indirect utility or wealth and derive an approximate CAPM representation of the pricing kernel $m_{t,t+1}^*$ (see Back, 2010, p. 187).\footnote{If one lifts the assumption of a constant investment opportunity set, but keep the rest of the (weak) assumptions (see Back, 2010, section 10.5), then one obtains the conditional CAPM that holds period by period conditional on the state of the economy.}

In equilibrium, the market portfolio must yield the best reward-to-risk ratio or the maximum Sharpe ratio using variance as a measure of discrepancy or risk. Since the seminal work of Hansen-Jagannathan (1991), a simple way to obtain the single market factor representation of the CAPM is to find the portfolio with minimum-variance bound (HJB). However, under statistical ambiguity variance is no longer the appropriate measure for asset pricing discrepancy.

To illustrate the implications of statistical ambiguity in asset pricing, let’s consider the relation between one-period relative entropy and the conditional distribution of the asset pricing kernel $m_{t+1}$, which is captured by the cumulative generating function and its cumulants.

Following the analysis in Backus et al. (2014), $\log m_{t+1}$ has conditional cumulant-generating function $K(\varepsilon) = \log M(\varepsilon)$, where $M(\varepsilon) = E[\exp(\varepsilon m_{t+1})] = \sum_{r=0}^{\infty} \mu_r \varepsilon^r / r!$ is the moment generating function defined over some range of $\varepsilon$, and $\mu_r$ is the $r$th moment of $\log m_{t+1}$ with density probability function $\pi$. The cumulants $\kappa_r(\varepsilon)$ are the coefficients in the Taylor series expansion of the cumulant generating function around the origin $\varepsilon = 0$:

\begin{equation}
K(\varepsilon) = \log M(\varepsilon) = \sum_{r=0}^{\infty} \frac{\kappa_r \varepsilon^r}{r!},
\end{equation}

where $\kappa_1 = \mu_1$ is the first cumulant that gives the mean return; $\kappa_2$ is the cumulant that gives the variance of returns; $\kappa_3 = E[(\log m_{t+1} - \mu_1)^3]$ is the third cumulant; and so on. In general:

\begin{equation}
D(m_{t+1}^*) = \frac{\kappa_2}{2!} (\log m_{t+1})^2 + \frac{\kappa_3}{3!} (\log m_{t+1})^3 + \frac{\kappa_4}{4!} (\log m_{t+1})^4 + \cdots.
\end{equation}
When using the standard CAPM as reference model, higher order cumulants (above order two) are zero by assumption, so the appropriate metric for asset pricing discrepancy is variance. Under the distorted CAPM, expression (5) involves a complicated combination of higher order cumulants different from zero capturing both negative skewness and excess kurtosis, two stylized facts of stock returns. In this case, variance is no longer the relevant metric for asset pricing discrepancy.

As already mentioned, under the standard CAPM, investors’ optimal allocation maximizes the squared Sharpe ratio, which defined in terms of the cumulant generating function is (Backus et al., 2014):

\[
\frac{\text{var}(m_{t+1})}{\text{E}[m_{t+1}]^2} = \exp[K_t(2) - 2K_t(1)] - 1.
\]  

(6)

The exponent in (6) has Taylor expansion:

\[
K_t(2) - 2K_t(1) = \sum_{j=1}^{\infty} \frac{\kappa_j (2j - 2)}{j!},
\]

(7)

whereas before \( \kappa_1 = \mu_1 \) is the first cumulant that gives the mean return; \( \kappa_2 \) is the second cumulant that gives the variance of returns; \( \kappa_3 = \text{E}[(\log m_{t+1} - \mu_1)^3] \) is the third cumulant; and so on.

Recall that under the standard CAPM, the investor is not concerned about model misspecification and statistical ambiguity, because higher order cumulants (above order two) are zero by assumption. That is, \( K_t(2) - 2K_t(1) = \kappa_{2t} \) and the squared Sharpe ratio is equal to \( \exp[\kappa_{2t}] - 1 \). Assuming a sufficiently small \( \kappa_2 \), the Sharpe ratio is approximately equal to \( \kappa_{2t} \). On the other hand, entropy is equal to \( \kappa_{2t}/2 \). Clearly, under the reference CAPM variance and entropy convey the same information. Under statistical ambiguity, expression (7) involves a complicated combination of cumulants different from zero and both metrics convey different information. Given that the robust optimal allocation will not necessarily correspond to the
minimum variance allocation, entropy is the appropriate metric of asset pricing discrepancy setting the following upper bound on log excess returns (See Backus et al., 2014, section 2.3, pages 5-6 for the proof):

\[ D(m_{t+1}^*) \geq E^{\pi^*}(logR_t) - logR_f. \]  

whereas as before, \( D \) denotes relative entropy. The equality holds under the worst-case scenario distorted probability law \( \pi^*_t \).

Notice that (8) gives a feasible route to recover empirically from the cross-section of stock returns the average worst-case scenario market premium placing an upper bound to the historical market premium. That is, the one that corresponds to the distorted return distribution with maximum entropy with respect to the reference model.

Let’s assume that the reference model is the single market factor representation of the standard CAPM. I follow the ambiguity literature, and think of the source of misspecification as a distortion only in the conditional mean of the vector of innovations in the DGP of stock returns that feedbacks arbitrarily into the history of the state (Hansen and Sargent, 2008 section 2.2). In other words, any plausible alternative (i.e., distorted) asset pricing model will differ from the reference CAPM because the conditional mean of the innovation process in the DGP of stock returns is different from zero (i.e., zero is the value of the conditional mean under the standard CAPM by assumption). The intuition is that because the standard CAPM is static, dynamic misspecification arises naturally.

This approach is computationally convenient because I can still use Sharpe’s ratio to obtain a single market factor representation of the robust pricing kernel even if the variance in returns is no longer the appropriate metric for asset pricing discrepancy. One important caveat though is that imposes an important restriction on the way that misspecification enters the
analysis. That is, the distorted models must be also normally distributed and dynamic misspecification should not affect the variance covariance matrix of stock returns. However, the approach is less restrictive if the reference asset pricing model is the standard CAPM.¹⁰

Let \( \xi^MKT_i \) denote the share of stock \( i \) (for all \( i = \{1, ..., I\} \)) in the market portfolio \( MKT \) such that \( \sum_{i=1}^{n} \xi^MKT_i = 1 \). Under the reference model, the market portfolio expected log return and variance are:

\[
E^\pi[logR_{MKT}] = \sum_{i=1}^{n} (\xi^MKT_i \times E^\pi[logR_i]) , \text{ and}
\]

\[
VAR_{MKT} = \sum_{i=1}^{n} \sum_{j=1}^{m} \xi^MKT_i \xi^MKT_j COVAR_{i,j},
\]

where \( VAR \) and \( COVAR \) are elements of the variance-covariance matrix \( \Sigma \). Under the distorted probability measure \( \pi^* \), the market portfolio expected log return and variance are:

\[
E^{\pi^*}[logR_{MKT}] = \sum_{i=1}^{n} (\xi^MKT_i \times E^{\pi^*}[logR_i]) , \text{ and}
\]

\[
VAR^{\pi^*}_{MKT} = VAR_{MKT}.
\]

In equilibrium, the optimal market portfolio allocations result from solving the following MaxMin robust decision problem:

\[
Max_{\xi_i} \text{Min}_{\pi^*} \left\{ \frac{E^{\pi^*}[logR_{MKT}] - logRF}{\sqrt{VAR^{\pi^*}_{MKT}}} \right\}, \quad (9)
\]

subject to constraint (1). Notice that under statistical ambiguity, the market portfolio is the one that maximizes the reward-to-risk ratio under the probability law \( \pi^*_t \) that simultaneously has maximum entropy with respect to the reference model and corresponds to investors’ worst-case scenario in terms of indirect utility or wealth.

¹⁰ Hansen and Sargent (2008), show that under quadratic preferences and linear transition laws or Normality, the distortion to the variance covariance matrix is insignificant. This result is also true in the more general case of a reference model with Markov transition distribution and misspecification slanting the probabilities of future states. This restriction can be also motivated by the analysis in Merton (1980), where when the transition law follows a standard random walk, then the variance can be estimated precisely by just increasing the frequency of the sample of observations. On the contrary, to estimate the mean will require a longer sample of observations.
Proposition 2 (For the proof see Appendix B):

(i) Under the worst-case scenario distorted probability measure $\pi_t^*$ (true DGP chosen by Nature), the following single market factor linear representation of the robust pricing kernel $m_{t,t+1}^*$ results:

$$
E^{\pi^*}(\log R_{i,t+1}) - \log R_f = \beta_i \lambda_{MKT}^*.
$$

(10)

where $E^{\pi^*}(\log R_{i,t+1}) - \log R_f$ denotes the expected log excess return of asset $i$ under the worst-case scenario probability measure $\pi_t^*$, $\log R_f$ is the one period risk free log rate, $\lambda_{MKT}^* \equiv E^{\pi^*}(\log R_{MKT}) - \log R_f$ is the market premium that accounts for risk and statistical ambiguity and corresponds to investors' worst-case scenario, and $\beta_i$ is the usual factor loading on the market return.

(ii) Under the reference model $\hat{\pi}_t$ (misspecified DGP), the following reference single market factor linear representation results:

$$
E^{\hat{\pi}_t}(\log R_{i,t+1}) - \log R_f = (\beta_i \lambda_{MKT}^* - \hat{\lambda}_i^*) + \beta_i \hat{\lambda}_{MKT}.
$$

(11)

where $E^{\hat{\pi}_t}(\log R_{i,t+1}) - \log R_f$ is the expected log excess return of asset $i$ under the reference probability measure $\hat{\pi}_t$, $\log R_f$ is the one period risk free log rate, $\lambda_i^* = \frac{\xi^{MKT}_{VAR_MKT}^{-1/2}}{\theta(\text{tr}(\Sigma^{-1}))}$, $\lambda_{MKT}^*$ is defined as before, and $\hat{\lambda}_{MKT} = E^{\hat{\pi}_t}(\log R_{MKT}) - \log R_f$ is the usual market premium that accounts for systematic risk.

Remark: Under the reference model, unconditional pricing errors $\alpha_i = (\beta_i \lambda_{MKT}^* - \hat{\lambda}_i^*)$ reflect stock's idiosyncratic ambiguity and the optimal allocation $\xi_{MKT}^*$ is Robust Bayes.
2.3 Econometric methodology: generalized maximum entropy (GME)

Cerreia-Vioglio et al. (2011) discuss the necessary and sufficient technical conditions to obtain duality between robust MaxMin behavioral decision problems that include ambiguity as an entropy-based penalty in indirect utility functions and robust Bayesian statistical problems that involve prior uncertainty or statistical ambiguity (i.e., dynamic consistency and a Dynkin space). Simply put, if one assumes that the investor chooses without regret the optimal allocation and stock returns follow a memoryless mechanism defined on a compact state space (e.g., a Markov chain with given transition function), then the worst-case scenario pricing kernel in terms of indirect utility or wealth admits a dual representation as the maximum entropy distribution of stock returns with respect some reference distribution. Prior uncertainty, is a classical problem of robust Bayesian statistics, with statistical ambiguity represented by a set of priors. Furthermore, Grünwald and Dawid (2004) advocate the use of Jaynes’ (1984) maximum entropy (ME) in robust Bayesian statistics.\(^{11}\)

Maximum entropy (ME) as a formal method of statistical inference first appears in the early works of Jaynes (1957a, 1957b, 1984) who proposed the approach for ill-posed problems in Bayesian statistics. Jaynes’ intuition for the use of ME in Bayesian statistics is based on an extension of LaPlace’s principle of insufficient reasoning. The axiomatic foundation of the ME principle can be found in Levine (1980), Shore and Johnson (1980), and Csiszar (1991). These authors show that the maximum entropy distribution is: 1) unique; 2) and is not affected by a change of specification or parameterization of the reference model. Unlike maximum likelihood (ML), ME does not make any assumptions about the true (hidden) distribution. Jaynes (1984) defines the ME distribution as the least prejudiced (biased) distribution compatible with the

\(^{11}\) As Grünwald and Dawid (2004, page 20) write: “…We see from Theorem 4.1. that when a saddle-point exists, the robust Bayes problem reduces to a maximum entropy problem. This property can thus be regarded as an indirect justification for applying maximum entropy.”
econometrician’s lack of confidence. In this sense, it should be preferred over any other distribution among all available distributions.

In empirical implementations, the main challenge of the econometrician using ME is to construct the multiple priors set and then pick that distribution from the set that corresponds to the worst-case scenario in terms of investors’ indirect utility or wealth. In practical terms, the estimation of the robust CAPM involves a procedure capable of picking the market premium that corresponds to the maximum entropy distribution of returns given the sample of observations. To illustrate the curse of dimensionality of this problem, let’s assume that the number of test assets used by the econometrician is \( N = 25 \), and the number of asset pricing factors included in the reference model is \( K = 3 \) (the true number of assets does not matter as a robust econometrician always discounts model misspecification). The econometrician entertains a bounded convex set of plausible support values for risk premia. The total number of plausible support values and probabilities for risk premia is the astonishing number of \( 2^{(25+3)} = 268,435,456 \). To reduce the dimensionality problem, I follow Golan et al. (1996), and assume that the robust econometrician solves the inverse problem using generalized maximum entropy (GME).

GME tackles the dimensionality of the estimation problem as follows. Assume again that the robust econometrician uses as reference model some multi factor (and misspecified) model in beta form representation (e.g., the Fama-French model encompassing the single factor market model):

\[
\mu - r = \beta \hat{\lambda} + \hat{\alpha},
\]

where \( \mu - r \) is a \((N \times 1)\) vector of test asset excess returns; \( \beta \) is a \((N \times K)\) matrix of risk factor loadings or betas \( \forall k \in [1, K] \) asset pricing factors; \( \hat{\lambda} \) is the \((K \times 1)\) vector of market prices of
risk and statistical ambiguity; and $\hat{\alpha}$ is the $(N\times1)$ vector of mispricing errors i.e., the sample estimates of test asset’s alphas.$^{12}$

Under GME, the robust econometrician re-writes equation (12) so entropy is introduced as the pseudo-metric to discriminate across alternative model parameterizations, each one corresponding to a distorted probability distribution of stock returns in the set of priors. The econometrician parameterizes $\hat{\lambda}_k$ as a discrete random variable with compact support of $M$ plausible outcomes $z_k = [z_{k1}, \ldots, z_{kM}]'$, where $2 \leq M \leq \infty$ and $z_{k1}$ and $z_{kM}$ are plausible upper and lower bounds for $\hat{\lambda}_k$. Thus, $\hat{\lambda}_k$ is now defined by $z_k'p_k$, where $p_k = [p_{k1}, \ldots, p_{kM}]'$ is a $M$-dimensional probability vector satisfying $\mathbf{1}'p_k = 1$ for $k = 1, \ldots, K$, where $\mathbf{1}$ denotes a vector of ones. All possible convex combinations are assembled in matrix form such that $\hat{\lambda} = Zp = \begin{bmatrix} z_1' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z_k' \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_k \end{bmatrix}$, where $Z$ is a $(K\times KM)$ matrix of support values and $P$ is a $KM$-dimensional matrix of probabilities. Notice that although the specification of the support matrix $Z$ influences the finite sample performance of the estimator, Monte Carlo simulations have shown that changing the parameter support has a modest impact on the estimation results (Golan et al., 1996).

Similarly, the vector of misspricing-errors has support matrix $V$ with $2 \leq J \leq \infty$ values, positive weight matrix $W$ such that $\mathbf{1}'w_t = 1$, for $t = 1, \ldots, T$, such that $\hat{\alpha} = VW = \begin{bmatrix} v_1' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_t' \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_t \end{bmatrix}$. Without loss of generality, I assume that the robust econometrician adopts the three-sigma rule of Pukelsheim (1994) to find the upper and lower misspricing error bounds.

$^{12}$ Note that I do not assume independence between the market prices of risk and statistical ambiguity and the mispricing errors (see Appendix B for details).
\( \nu = -3\hat{\sigma}_{\mu-r} \) and \( \bar{v} = 3\hat{\sigma}_{\mu-r} \), where \( \hat{\sigma}_{\mu-r} \) denotes the return standard deviation (Golan et al., 1996). Again, the finite sample performance of the estimator depends on this specification. Intuitively, the larger is the misspricing-error bound the larger is the shrinkage rate of the estimator towards zero guaranteeing an interior solution.

Given \( \mathbf{V}, \mathbf{W}, \mathbf{P} \), equation (12) is re-written as:

\[
\mu - r = \mathbf{\beta}'\mathbf{Z}\hat{\mathbf{P}} + \mathbf{V}\hat{\mathbf{W}},
\]

where \( \hat{\mathbf{\lambda}} = \mathbf{Z}\hat{\mathbf{P}}; \hat{\mathbf{\alpha}} = \mathbf{V}\mathbf{W} \); and \( \hat{\mathbf{P}} \) represents the maximum entropy (ME) probability distribution given the sample of average excess returns.

The estimation problem of the robust econometrician now involves \( 2(N+K) \) unknown plausible solutions. Using the same example as before with \( N = 25 \) and \( K = 3 \), the number of plausible solutions collapses to 56 instead of the original 268 million!

The robust econometrician recovers the optimal values of \( \hat{\mathbf{P}} \) and \( \hat{\mathbf{W}} \) solving an optimization problem by the method of Lagrange multipliers with objective function defined in terms of entropy (for technical details see Appendix C):

\[
F(\mathbf{\beta}) + F(\mathbf{\alpha}) = H(\hat{\mathbf{P}}) - H(\hat{\mathbf{W}}) = -\hat{\mathbf{P}}'ln(\hat{\mathbf{P}}) - \hat{\mathbf{W}}'ln(\hat{\mathbf{W}}),
\]

subject to the moment restriction (given the reference asset pricing model) \( N^{-1}\mathbf{\beta}'(\mu - r) = N^{-1}\mathbf{\beta}'\mathbf{Z}\hat{\mathbf{P}} + N^{-1}\mathbf{\beta}'\mathbf{V}\hat{\mathbf{W}} \).\(^{13}\)

Since the objective function is strictly concave and the feasible set is nonempty, closed, and bounded, the Hessian matrix is negative definite and the solution to GME is globally optimal and unique (Mittelhammer et al., 2000). However, since there is no closed-form solution for the Lagrange multipliers, the global unique solution must be found numerically.

\(^{13}\) The restrictions are usually referred to as consistency constraints. Higher moment restrictions may be introduced to increase the informational content of the estimates. This should be done only if one assumes that the robust econometrician has some prior information about these moments.
The final step in the estimation procedure involves going backwards to recover the market premium that corresponds to the ME return distribution $\hat{\lambda} = E^{\hat{P}}(Z)$ with misspricing errors $\hat{\alpha} = E^{\hat{W}}(V)$.

Conceptually, GME belongs to a broad class of extremum-estimators. For the details of the asymptotic and small sample properties of the estimator we refer the reader to Golan et al. (1996) and Mittelhammer et al. (2000).

3 Asset Pricing Tests

3.1 Data

For the dataset, I follow the standard approach in the empirical asset pricing literature, and use as test assets value-weighted monthly returns on 25 Fama-French portfolios sorted by size (ME) and book-to-market (BE/ME), obtained from Kenneth French’s website. The sample period runs from 1927 to 2007, but I also discuss results for the periods 1927-1961, 1962-2007, and 1927 to 2010 (the latter to assess the effects of the recent financial crisis of 2008). Following Lewellen et al. (2010) critique, I also show results from using an augmented set of test assets that

---

14 For an introduction to extremum-estimators applied in empirical finance see Singleton (2006) Chapter 3. Using a change of measure approach, Kitamura and Stutzer (1997) show that for weakly dependent data generating processes and reasonable regularity conditions, ME estimators have similar large sample properties to GMM i.e., that is the estimator is consistent, asymptotically normal, with the same covariance matrix than GMM, but potentially with better sample properties than GMM in the class of optimal minimum distance (ODM) estimators.

15 The GME estimator is consistent and asymptotically normally distributed. Thus, normal and chi-square based test statistics can be used for hypothesis testing with the usual confidence regions for the unknown parameters. Wald tests, (Lagrange multiplier) LM tests, and pseudo-likelihood ratio (LR) tests can be expressed in the usual way. Golan et al. (1996) report the results of sampling experiments involving 5,000 Monte Carlo trials for many well and ill-conditioned problems to reflect the small sample performance of GME versus LS (i.e., ridge regression) and ML. The minimum square empirical (MSE) risk of GME dominates the rest for both small and large (up to 100) parameter spaces.

includes 30 industry-sorted portfolios of stock returns, also obtained from Kenneth French’s website.

The stock market excess return is calculated as the difference between the nominal return on the CRSP value-weighted stock market index and the (anticipated) one-month Treasury bill yield. Following Fama and French (1992, 1993), and Carhart (1997), I also include the two Fama-French (1992, 1993) firm-specific asset pricing factors related to size \((SMB)\) and book-to-market \((HML)\), as well as Jegadeesh (1990) and Jegadeesh and Titman’s (1993) momentum \((UMD)\) asset pricing factor. Tables 1 and 2 report summary statistics for the test assets and asset pricing factors.

**3.2 Empirical results**

In Table 3, I include the estimation results for the period 1927 – 2007. In Panel A, I report estimates from using a window of 36 months in first pass time series rolling regressions, and GLS in the second pass cross-sectional regression (CSR). These estimates correspond to: 1) a single-factor market model that includes an intercept in the CSR; and 2) a single-factor market model without intercept in the CSR. Then I report results from using GME in the second pass CSR. These estimates correspond to: 1) a misspecified single-factor market model that includes an intercept in the CSR; and 2) (to assess the incremental explanatory power of the Fama-French-Carhart factors) the Fama-French-Carhart four-factor model. In Panel B, I report estimates using instead a window of 240 months in first pass time-series regressions. I also ran the asset pricing tests using the sub-sample periods 1927 - 1961 and 1962 – 2007, but I do not present them to save space.
Under GLS and using a rolling window of 36 months in first pass regressions, the market premium in a single-factor model w/o intercept is statistically significant at the 1% level with monthly values of 0.74% (for the 1927-2007 period), 0.89% (for the 1927-1961 period), and 0.61% (for the Compustat period 1962 – 2007), respectively. If the intercept is included in the CSR, the market premium becomes now statistically insignificant and the incumbent intercept is statistically significant at 1%. This result applies for all three sample periods. If the Fama-French-Carhart factors are included in the asset pricing model, they are also statistically significant. The results are robust to different rolling regression windows and any potential error in variables problem as shown by Shanken’s (1992) corrected t-stats. The results are consistent with previous findings in the empirical asset pricing literature, and confirm the instability of the estimates when using a standard econometric approach.

Using GME and a rolling window of 36 months in first pass regressions, the robust market premium that accounts for both risk and statistical ambiguity is statistically significant at 1% level with monthly values of 1.32% (for the 1927-2007 period), 1.60% (for the 1927-1961), and 1.10% (for the Compustat 1962 – 2007 period), respectively. These results are robust to different rolling regression windows and any potential error in variables problem as shown by Shanken’s (1992) corrected t-stats. The striking result is that GME estimates are quite stable to changes in parameterization of the asset pricing model. The results are robust to the inclusion or not of the intercept and/or additional factors like the Fama-French-Carhart factors.

To illustrate the economic significance of statistical ambiguity in empirical asset pricing tests, I calculate lower and upper bounds for the ambiguity premium in excess of risk. To this purpose, I use minimum and maximum values for the equity risk premium as documented in Mehra and Prescott (2003). Then, I subtract them from the smallest and largest GME estimates.
of the robust market premium. In Table 4, I present calculations that correspond to the period 1927-2007. The monthly value of the premium for statistical ambiguity lies between 0.53% and 0.99%, which is economically meaningful.

3.3 Robustness checks

I assess the robustness of previous results across two dimensions. First, I expand the sample period from 1927 to 2010 to include the impact (if any) of the recent financial crisis of 2008 as shown in Table 5. Then, given Lewellen et al. (2010) critique, I repeat the asset pricing tests using an augmented set of test assets that includes 30 portfolios sorted by industry as shown in Table 6.

Clearly, the original empirical results are robust to the augmented set of test assets and the recent crisis of 2008. Similar results can be obtained if one runs the tests using different sample periods (the later not included to save space).

3.4 Economic interpretation of results

In economics, Lagrange multipliers (LMs) are interpreted as shadow prices. If they are zero, it means that the constraint for that test asset is not binding. Under the proposed robust approach, the LMs reflect each test asset’s idiosyncratic contribution to investors’ lack of confidence in their reference models. In this regard, the LMs can be interpreted as shadow prices of idiosyncratic ambiguity, which unlike idiosyncratic volatility investors cannot diversify away. Thus, they ask for an extra compensation in the market premium.

In the asset pricing literature, market anomalies like size and value premia have been traditionally attributed to exposure to omitted risk factors (like distress risk) not fully captured by
market beta (e.g. Fama and French, 1992), intangible information (Jiang, 2010), and cognitive biases (behavioral literature). The results seem to suggest that size and value are statistical characteristics of a misspecified CAPM.

In Table 7 and Figures 1 and 2, I display the LMs ranked across size (ME) and book-to-market (BE/ME). Conditional on size, the shadow price of idiosyncratic ambiguity across book-to-market increases monotonically from -0.17% (group of growth stocks) to 0.38% (group of value stocks). Conditional on book-to-market, the shadow price of idiosyncratic ambiguity across size is always positive. The average shadow price of ambiguity for small cap stocks is 0.18% and for large cap stocks is 0.08%. The marginal contribution of book-to-market to aggregate market ambiguity is 0.48% while the marginal contribution of size is 0.10%. In Table 8, I include results from running 2-tail t-tests of the null hypothesis that LMs across portfolios ranked by size and book-to-market are equal to zero. The goal is to check if the constraints are binding. The test results at the 1% significance level show that all LMs ranked by book-to-market are statistically different from zero. For LMs ranked by size, I fail to reject the null hypothesis only for the 3 smallest portfolios. In Table 9, I report the robust Bayes portfolio allocations for each test asset that follow the 1/N allocation rule of the ambiguity literature (see Appendix C). Clearly, size and value are closely correlated with idiosyncratic ambiguity.

4. Conclusion

By revisiting the empirical capabilities of one of the most popular asset pricing models in finance under statistical ambiguity: 1) I show that a robust simple single market factor model that accounts for risk and statistical ambiguity can explain the cross-section of stock returns, 2) market anomalies like size and value show up as statistical characteristics of misspecified asset
pricing models given that idiosyncratic ambiguity, unlike idiosyncratic volatility, cannot be diversified away, and 3) I provide an answer to Al Najjar’s (2009) critique that ambiguity or Knightian uncertainty in economics and finance lack empirical motivation.
References


Appendix A

Proof of Proposition 1

The proof follows the analysis in Viale et al. (2014) under their usual measurability conditions. Assume that the RA is endowed with a standard time-additive separable Von Neumann Morgenstern utility of consumption. Let \( R_{i,t+1}^s \equiv \frac{p_{i,t+1}^s}{p_{i,t}} \) denote the gross return of stock \( i \) whereby \( p_{i,t} \) represents the stock price at time \( t \) and \( p_{i,t+1}^s \) at time \( t+1 \). Let \( R_f \) represent the constant gross return of a dollar invested in the risk-free asset. Then, the budget constraint is given by:

\[
W_{t+1}^s = (W_t - C_t) (R_f + \sum_{t=1}^{n} \xi_{i,t} (R_{i,t+1}^s - R_f)). \tag{A.1}
\]

Starting from terminal date \( T \) and using backward recursion one obtains the dynamic Bellman equation (1):

\[
J(W_t, t) = \max_{c_t[\xi_t]} \left\{ u(C_t) + \min_{\pi_t^*} \mathbb{E}_t[\pi_t^*(W_{t+1}^s, t+1) + (1 - \pi_t^*)J(W_{t+1}^0, t+1)] + \frac{1}{2} \theta_t (D_t(\pi_t^*||\pi_t) - \eta_t) \right\}, \tag{A.2}
\]

where \( \theta_t \geq 0 \) is Lagrange multiplier for the ambiguity constraint in equation (1).

Next, insert the approximation \( D(\pi_t^*||\pi_t) \approx \frac{(\pi_t^* - \pi_t)^2}{\pi_t(1-\pi_t)} \) into equation (A.2) to generate the necessary first order condition for optimality of the inner minimization problem:

\[
\pi_t - \pi_t^* = \frac{E_t(f(W_{t+1}^s, t+1) - E_t(f(W_{t+1}^0, t+1)))\pi_t(1-\pi_t)}{\theta_t}. \tag{B.3}
\]

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

From equation (1), write the complementary slackness condition of the entropy constraint as:

\[
\frac{1}{2} \theta_t \left( \frac{(\pi_t - \pi_t^*)^2}{\pi_t(1-\pi_t)} - \eta_t \right) = 0. \tag{A.4}
\]
Substituting equation (A.3) into equation (A.4) gives:

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

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

\[
\pi_t^* = \pi_t - \sqrt{2\eta_t\pi_t(1-\pi_t)}. \tag{A.6}
\]

When Equation (A.6) holds, equation (A.2) turns into the following unconstrained maximization problem:

\[
J(W_t,t) = \max_{C_t} \left\{ u(C_t) + \left( \pi_t^*E_t(J(W_{t+1}^1,t + 1)) + (1 - \pi_t^*)E_t(J(W_{t+1}^0,t + 1)) \right) \right\}. \tag{A.7}
\]

Observe that (A.7) delivers necessary conditions for existence of an interval of attainable minimization solutions with corner equilibria if \(\pi_t^*\) is degenerate. This interval is reminiscent of the so-called “Euler-inequalities” from Epstein and Wang (1994) pp. 300 expression (3.32). Furthermore, since \(0 \leq \pi_t^* \leq 1\), the continuity axiom for Von Neumann and Morgenstern utility functions implies that there exists a unique wealth process \(W_{t+1}^1\) (with \(E_t(W_{t+1}^0) \leq E_t(W_{t+1}^1) \leq E_t(W_{t+1}^1))\) and associated price process \(p_{i,t+1}\) (with \(E_t(p_{i,t+1}^0) \leq E_t(p_{i,t+1}^1) \leq E_t(p_{i,t+1}^1)\) for all \(i = 1, \ldots, I\) such that:

\[
E_t(J(W_{t+1},t + 1)) = \pi_t^*E_t(J(W_{t+1}^1,t + 1)) + (1 - \pi_t^*)E_t(J(W_{t+1}^0,t + 1)), \tag{A.8}
\]

where \(E_t(J(W_{t+1},t + 1))\) is interpreted as the ambiguity-certainty equivalent of the expected continuation values \(E_t(J(W_{t+1}^1,t + 1))\) and \(E_t(J(W_{t+1}^0,t + 1))\) averaged under the worst-case scenario probability \(\pi_t^*\). Consequently, we re-express equation (A.7) as:

\[
J(W_t,t) = \max_{C_t} \left\{ u(C_t) + E_t(J(W_{t+1},t + 1)) \right\}. \tag{A.9}
\]
Expression (A.9) reveals that the original dynamic recursive problem in equation (A.2) reduces to a standard Bellman equation indexed by the distorted measure $\pi_t^*$. Finally, standard first-order conditions (see Back (2008) p.122-123) applied to the outer maximization problem (A.9) delivers Proposition 1. □

Appendix B

Proof of proposition 2

Problem (9) may be expressed as the following unconstrained problem:

\[
\begin{align*}
\text{Max}_{\xi_i} \text{Min}_{\pi^*} \left\{ \frac{E^{\pi^*} \left[ \log R_{MKT} \right] - \log R_f}{\sqrt{\text{VAR}_{MKT}}} - \sum_{t=1}^{T} \theta_t \left[ D_t(\pi_t^* || \hat{\pi}_t) - \eta_t \right] \right\},
\end{align*}
\]  

(B.1)

where $\theta_t \geq 0$ are Lagrange multipliers.

Notice that for the case of multivariate Normal distributions and imposing restriction $\Sigma^* = \Sigma$, the Kullback-Leibler distance can be expressed as:

\[
D_t(\pi_t^* || \hat{\pi}_t) = \frac{1}{2} \text{tr} \left( \Sigma^{-1} (\mu^* - \hat{\mu})(\mu^* - \hat{\mu})' \right),
\]

where $\mu^*$ is a vector of expected returns under the distorted probability law for all $i = \{1, ..., I\}$, $\hat{\mu}$ is a vector of expected returns under the reference probability law (for all $i = \{1, ..., I\}$), and $'$ denotes the transpose operation. Necessary first order conditions for the inner-minimization problem under the worst-case scenario probability law $\pi_t^*$ yield:

\[
E^{\pi^*} \left[ \log R_i \right] = E^\pi \left[ \log R_i \right] + \frac{\text{VAR}_{MKT}^{-1/2} \xi_i}{\theta_t \text{tr}(\Sigma^{-1})},
\]  

(B.2)
for all \( i \) assets. Substituting back into (B.1) and after some algebraic manipulation one can restate (B.1) as the following maximization problem:

\[
\text{Max}_{\xi_i} \left\{ \frac{\sum_{i=1}^{I} \xi_i (E^{\hat{\pi}}[\log R_i] - \log R_f)}{\sum_{i=1}^{I} \sum_{j=1}^{J} \xi_i \xi_j \text{COV}_{ij}} \right\},
\]

Taking the derivative of expression (B.3) at \( \xi_i = \xi_i^{MKT} \) yields for the first term:

\[
\frac{(E^{\hat{\pi}}[\log R_i] - \log R_f) \text{VAR}^{1/2}_{MKT} - (1/2) (\sum_{i=1}^{I} \sum_{j=1}^{J} \xi_i \xi_j \text{COV}_{ij})^{-1/2}}{\sum_{i=1}^{I} \sum_{j=1}^{J} \xi_i \xi_j \text{COV}_{ij}}
\]

where \( \text{COV}_{i,MKT} = \sum_{j=1}^{J} \xi_i^{MKT} \text{COV}_{ij} \). Taking derivative of the second term in expression (B.3) yields:

\[
\frac{\xi_i^{MKT} \text{VAR}^{1/2}_{MKT}}{\theta_i \text{tr}(\Sigma^{-1}) \text{VAR}_{MKT}} - \sum_{i=1}^{I} \left( \frac{\xi_i^{MKT}}{\theta_i \text{tr}(\Sigma^{-1}) \text{VAR}_{MKT}} \right)^2 \text{COV}_{i,MKT}
\]

Adding up (B.4) and (B.5) and setting the sum equal to zero yields after some algebraic manipulation:

\[
E^{\hat{\pi}}[\log R_i] - \log R_f = \left\{ \beta_i \left[ \sum_{i=1}^{I} \left( \frac{\xi_i^{MKT}}{\theta_i \text{tr}(\Sigma^{-1}) \text{VAR}_{MKT}} \right)^2 \text{VAR}_{MKT}^{1/2} \right] - \frac{\xi_i^{MKT}}{\theta_i \text{tr}(\Sigma^{-1}) \text{VAR}_{MKT}^{1/2}} \right\} + \beta_i (E^{\hat{\pi}}[\log R_{MKT}] - \log R_f).
\]

where \( \beta_i = \frac{\text{COV}_{i,MKT}}{\text{VAR}_{MKT}} \). Finally, from the complementary slackness condition:

\[
\theta_i (D_t (\pi^*_i || \hat{\pi}_t) - \eta_i) = 0.
\]

Combining (1) and (B.2) gives:
\[ \lambda_i^* \equiv \frac{\xi_i^{MKT \text{VAR}}_{MKT}^{-1/2}}{\theta_i tr(\Sigma^{-1})} = \sqrt{\frac{2\eta_i}{tr(\Sigma^{-1})}}. \] (B.7)

Furthermore, observe that:

\[ \lambda_{MKT}^* = \sum_{i=1}^I \xi_i^{MKT} \xi_i^{MKT \text{VAR}_{MKT}}^{-1/2} \frac{\theta_i tr(\Sigma^{-1})}{\sqrt{tr(\Sigma^{-1})}} = \sum_{i=1}^I \xi_i^{MKT} \lambda_i^*. \] (B.8)

Substituting (B.7) into (B.2) leads to:

\[ E^{\pi^*}[log R_i] = E^{\hat{\pi}}[log R_i] + \lambda_i^* \] (B.9)

and:

\[ E^{\pi^*}[log R_{MKT}] = \sum_{i=1}^I \xi_i^{MKT} \left( E^{\hat{\pi}}[log R_i] + \lambda_i^* \right) = E^{\hat{\pi}}[log R_{MKT}] + \lambda_{MKT}^* \] (B.10)

Substituting (B.9) and (B.10) into (B.6) yields proposition 2. □

Appendix C

General maximum entropy (GME)

Let:

\[ \mu - r = \beta \hat{\lambda} + \hat{\alpha}, \] (C.1)

where \( \mu - r \) is a \((N\times1)\) vector of stock excess returns; \( \beta \) is a \((N\times K)\) matrix of factor loadings; \( \hat{\lambda} \) is a \((K\times1)\) vector of robust market prices of risk and ambiguity and \( \hat{\alpha} \) is the \((N\times1)\) vector of misspricing errors.

Assumption 1: \( \lambda \in \Lambda \) where \( \Lambda \) is a convex set.

Let \( \beta^* = [\beta, 1] \) and define \( \delta = [\hat{\lambda}, \hat{\alpha}] \). Then,

\[ \mu - r = \beta \hat{\lambda} + \hat{\alpha} = \beta^* \hat{\delta} \] (C.2)
where $\delta$ is in the convex set $\Lambda^* = \Lambda \times V$ and $V$ is the convex hull of $\alpha$. Notice that I do not assume independence between $\lambda$ and $\alpha$. Moreover, I do not assume any distribution on $\alpha$. Thus, cumulative entropy can be defined on the extended convex set $\Lambda^*$ yielding the following estimation problem:

$$
\max_{P,W} F(\delta) = F(\lambda) + F(\alpha) = F(\lambda) + F(\mu - r - \beta \lambda) = H(P) - H(W)
$$

$$
= -P'\ln(P) - W'\ln(W), \quad (C.3)
$$

Subject to $\mu - r = \beta Z \hat{P} + V \hat{W}$, \quad (C.4)

and $(I_K \otimes I_M) p = \iota_K, (I_N \otimes I_j') w = \iota_N$, \quad (C.5)

where $H(P)$ and $H(W)$ are entropy measures for robust market prices of risk and ambiguity $\lambda$ and misspricing errors $\alpha$, respectively, defined over the joint space $\Lambda^* = \Lambda \times V$, such that $\hat{\lambda} = E^P(Z)$ and $\hat{\alpha} = E^W(V)$; $I$ are identity matrices; $\iota_K = \{1,1, \cdots, 1\}$; $\iota_j = \{1,1, \cdots, 1\}$; and $\otimes$ is the Kronecker product. Applying the Lagrange method, the Lagrangian equation takes the form:

$$
L \equiv -P'\ln(P) - W'\ln(W) + \theta'(\mu - r - \beta Z \hat{P} - V \hat{W}) + \phi'(\iota_K - (I_K \otimes I_M)p)
$$

$$
+ \psi'(\iota_N - (I_N \otimes I_j')w). \quad (C.6)
$$

Solving the first order conditions, the inverse probabilities and misspricing error weights are:

$$
(\hat{p}_k^m) = \frac{e^{-(\sum \beta i \hat{x}_i^m \beta_k^i)}}{\sum m e^{-(\sum \beta i \hat{x}_i^m \beta_k^i)}} e^{-(\sum \beta i \hat{x}_i^m \beta_k^i)} / \phi_k(\theta), \quad (C.7)
$$

$$
(\hat{w}_j^i) = \frac{e^{-(\hat{\alpha} v_j)}}{\sum j e^{-(\hat{\alpha} v_j)}} e^{-(\hat{\alpha} v_j)} / \psi_j(\theta). \quad (C.8)
$$

Observe that with the probabilities and weights one can proceed to calculate the point estimates $\hat{\lambda}_k = \sum m z_k^m \hat{p}_k^m$ and misspricing errors $\hat{\alpha}_i = \sum_j v_j \hat{w}_j^i$. Finally, using the cumulative
objective function and the resulting marginal posteriors, one can obtain also the natural weight of each test asset in the market portfolio as a function of the Lagrange multipliers:

\[ x_i(\hat{\theta}) = \frac{e^{-\hat{\theta}_i}}{\sum_i e^{-\hat{\theta}_i}}. \]  

(C.9)

If \( \hat{\theta}_i = 0 \) \( \forall i \), then \( x_i(\hat{\theta}) = \frac{1}{N} \) and the uniform (i.e., Robust Bayes) allocation obtains.
### Table 1
**Summary Statistics: Test Assets (1927-2007)**

#### PANEL A

<table>
<thead>
<tr>
<th>ME/BEME</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>31</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.77</td>
<td>1.14</td>
<td>1.35</td>
<td>1.51</td>
<td>0.88</td>
<td>1.26</td>
<td>1.35</td>
<td>1.42</td>
<td>1.53</td>
<td>0.99</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Std. dev.</td>
<td>12.41</td>
<td>10.71</td>
<td>9.30</td>
<td>8.71</td>
<td>8.03</td>
<td>7.91</td>
<td>7.34</td>
<td>7.62</td>
<td>8.72</td>
<td>7.68</td>
<td>6.58</td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>-49.36</td>
<td>-43.09</td>
<td>-36.58</td>
<td>-35.78</td>
<td>-34.87</td>
<td>-32.50</td>
<td>-30.58</td>
<td>-32.77</td>
<td>-34.64</td>
<td>-29.52</td>
<td>-29.27</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>147.50</td>
<td>139.27</td>
<td>81.04</td>
<td>105.07</td>
<td>105.31</td>
<td>54.13</td>
<td>84.41</td>
<td>78.79</td>
<td>72.57</td>
<td>87.37</td>
<td>60.75</td>
<td>44.32</td>
</tr>
<tr>
<td>Median</td>
<td>0.58</td>
<td>0.94</td>
<td>1.22</td>
<td>1.41</td>
<td>1.52</td>
<td>1.14</td>
<td>1.54</td>
<td>1.50</td>
<td>1.58</td>
<td>1.63</td>
<td>1.39</td>
<td>1.38</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>27.76</td>
<td>57.41</td>
<td>15.75</td>
<td>31.08</td>
<td>31.36</td>
<td>5.06</td>
<td>21.91</td>
<td>23.25</td>
<td>18.91</td>
<td>18.70</td>
<td>10.75</td>
<td>6.97</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.73</td>
<td>4.47</td>
<td>1.81</td>
<td>2.82</td>
<td>3.22</td>
<td>0.37</td>
<td>1.97</td>
<td>2.19</td>
<td>1.80</td>
<td>1.91</td>
<td>1.07</td>
<td>0.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ME/BEME</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>51</th>
<th>52</th>
<th>53</th>
<th>54</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.29</td>
<td>1.30</td>
<td>1.43</td>
<td>0.98</td>
<td>1.06</td>
<td>1.18</td>
<td>1.27</td>
<td>1.40</td>
<td>0.91</td>
<td>0.91</td>
<td>0.99</td>
<td>1.03</td>
<td>1.30</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>6.75</td>
<td>6.80</td>
<td>8.68</td>
<td>6.24</td>
<td>6.27</td>
<td>6.34</td>
<td>7.03</td>
<td>9.02</td>
<td>5.49</td>
<td>5.24</td>
<td>5.74</td>
<td>6.92</td>
<td>7.56</td>
</tr>
<tr>
<td>Max</td>
<td>64.27</td>
<td>56.21</td>
<td>82.06</td>
<td>34.47</td>
<td>57.56</td>
<td>64.91</td>
<td>70.67</td>
<td>86.43</td>
<td>35.52</td>
<td>32.24</td>
<td>48.41</td>
<td>65.04</td>
<td>56.82</td>
</tr>
<tr>
<td>Median</td>
<td>1.56</td>
<td>1.50</td>
<td>1.38</td>
<td>1.20</td>
<td>1.39</td>
<td>1.49</td>
<td>1.51</td>
<td>1.52</td>
<td>1.07</td>
<td>1.05</td>
<td>1.16</td>
<td>1.11</td>
<td>1.34</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>15.05</td>
<td>13.91</td>
<td>20.09</td>
<td>3.60</td>
<td>12.94</td>
<td>15.71</td>
<td>21.27</td>
<td>22.62</td>
<td>5.50</td>
<td>5.35</td>
<td>15.24</td>
<td>24.35</td>
<td>11.81</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.07</td>
<td>1.26</td>
<td>1.97</td>
<td>-0.19</td>
<td>0.92</td>
<td>1.09</td>
<td>1.91</td>
<td>2.11</td>
<td>0.00</td>
<td>-0.07</td>
<td>0.88</td>
<td>1.95</td>
<td>0.74</td>
</tr>
</tbody>
</table>
## Continuation Table 1

### Summary Statistics: Test Assets (1927-2007)

#### PANEL B

<table>
<thead>
<tr>
<th>Food</th>
<th>Beer</th>
<th>Smoke</th>
<th>Games</th>
<th>Books</th>
<th>Hshld</th>
<th>Clths</th>
<th>Hlth</th>
<th>Chems</th>
<th>Txtls</th>
<th>Cnstr</th>
<th>Steel</th>
<th>FabPr</th>
<th>EleEq</th>
<th>Autos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.01</td>
<td>1.23</td>
<td>1.16</td>
<td>1.13</td>
<td>0.96</td>
<td>0.97</td>
<td>0.87</td>
<td>1.09</td>
<td>1.05</td>
<td>0.94</td>
<td>0.97</td>
<td>1.02</td>
<td>1.07</td>
<td>1.24</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>4.87</td>
<td>7.47</td>
<td>5.90</td>
<td>9.00</td>
<td>6.97</td>
<td>6.05</td>
<td>6.03</td>
<td>5.79</td>
<td>6.26</td>
<td>7.56</td>
<td>6.85</td>
<td>8.36</td>
<td>7.22</td>
<td>7.68</td>
</tr>
<tr>
<td>Max</td>
<td>33.48</td>
<td>89.19</td>
<td>33.30</td>
<td>66.81</td>
<td>53.43</td>
<td>59.22</td>
<td>41.17</td>
<td>38.66</td>
<td>46.99</td>
<td>56.59</td>
<td>43.21</td>
<td>80.84</td>
<td>51.94</td>
<td>59.62</td>
</tr>
<tr>
<td>Median</td>
<td>1.23</td>
<td>0.96</td>
<td>1.23</td>
<td>1.16</td>
<td>0.84</td>
<td>1.12</td>
<td>0.83</td>
<td>1.08</td>
<td>1.00</td>
<td>0.86</td>
<td>1.12</td>
<td>1.16</td>
<td>1.43</td>
<td>1.13</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>6.65</td>
<td>22.54</td>
<td>3.48</td>
<td>9.93</td>
<td>7.07</td>
<td>12.95</td>
<td>5.13</td>
<td>7.19</td>
<td>7.33</td>
<td>8.71</td>
<td>6.22</td>
<td>15.00</td>
<td>7.79</td>
<td>9.05</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.06</td>
<td>1.83</td>
<td>0.06</td>
<td>0.66</td>
<td>0.50</td>
<td>0.37</td>
<td>0.35</td>
<td>0.17</td>
<td>0.42</td>
<td>0.87</td>
<td>0.41</td>
<td>1.57</td>
<td>0.57</td>
<td>0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carry</th>
<th>Mines</th>
<th>Coal</th>
<th>Oil</th>
<th>Util</th>
<th>Telecom</th>
<th>Servs</th>
<th>BusEq</th>
<th>Paper</th>
<th>Trans</th>
<th>Whols</th>
<th>Retail</th>
<th>Meals</th>
<th>Fin</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.14</td>
<td>0.98</td>
<td>1.30</td>
<td>1.12</td>
<td>0.91</td>
<td>0.85</td>
<td>1.22</td>
<td>1.13</td>
<td>1.06</td>
<td>0.91</td>
<td>0.86</td>
<td>1.01</td>
<td>1.06</td>
<td>1.08</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>7.75</td>
<td>7.14</td>
<td>8.89</td>
<td>6.09</td>
<td>5.70</td>
<td>4.59</td>
<td>8.68</td>
<td>6.85</td>
<td>5.96</td>
<td>7.28</td>
<td>7.49</td>
<td>6.05</td>
<td>6.75</td>
<td>6.85</td>
</tr>
<tr>
<td>Max</td>
<td>49.60</td>
<td>46.14</td>
<td>77.54</td>
<td>39.16</td>
<td>43.16</td>
<td>28.16</td>
<td>73.32</td>
<td>38.28</td>
<td>43.38</td>
<td>65.35</td>
<td>59.21</td>
<td>37.75</td>
<td>31.49</td>
<td>59.85</td>
</tr>
<tr>
<td>Median</td>
<td>1.21</td>
<td>0.88</td>
<td>0.64</td>
<td>0.92</td>
<td>1.07</td>
<td>0.89</td>
<td>1.09</td>
<td>1.18</td>
<td>0.95</td>
<td>1.06</td>
<td>0.99</td>
<td>1.28</td>
<td>1.18</td>
<td>0.80</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.69</td>
<td>3.69</td>
<td>7.41</td>
<td>4.22</td>
<td>7.67</td>
<td>3.33</td>
<td>16.24</td>
<td>3.27</td>
<td>6.95</td>
<td>13.44</td>
<td>11.70</td>
<td>5.16</td>
<td>2.62</td>
<td>12.27</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.51</td>
<td>0.29</td>
<td>1.07</td>
<td>0.33</td>
<td>0.09</td>
<td>0.07</td>
<td>1.10</td>
<td>-0.21</td>
<td>0.35</td>
<td>1.12</td>
<td>0.67</td>
<td>0.02</td>
<td>-0.34</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Note: This table reports summary statistics for monthly returns on 25 Fama-French portfolios sorted by size (ME) and book-to-market (BE/ME) in Panel A; and 30 stock portfolios sorted by industry in Panel B.
<table>
<thead>
<tr>
<th></th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>UMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.6420</td>
<td>0.2340</td>
<td>0.4120</td>
<td>0.7630</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>5.4300</td>
<td>3.3600</td>
<td>3.5900</td>
<td>4.6500</td>
</tr>
<tr>
<td>Min</td>
<td>-29.0400</td>
<td>-16.7900</td>
<td>-13.4500</td>
<td>-50.6300</td>
</tr>
<tr>
<td>Max</td>
<td>38.2700</td>
<td>39.0400</td>
<td>35.4800</td>
<td>18.3900</td>
</tr>
<tr>
<td>Median</td>
<td>0.9550</td>
<td>0.0600</td>
<td>0.2600</td>
<td>0.9250</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>7.9745</td>
<td>22.3512</td>
<td>16.0341</td>
<td>28.8354</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.2235</td>
<td>2.2104</td>
<td>1.9113</td>
<td>-3.0412</td>
</tr>
</tbody>
</table>

Note: This table reports in summary statistics for the asset pricing factors included in the empirical analysis. MKT denotes the market factor; SMB the size factor; HML the value factor; and UMD the momentum factor.
### Table 3
Second pass cross sectional (CSR) estimates
Single period CAPM - (1927-2007)

<table>
<thead>
<tr>
<th></th>
<th>PANEL A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>36 M Roll Reg</td>
<td>CAPM w/int</td>
<td>CAPM wo/int</td>
<td>FF4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\lambda_0$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_0$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_{SMB}$</td>
<td>$\lambda_{HML}$</td>
<td>$\lambda_{UMD}$</td>
<td></td>
</tr>
<tr>
<td>GLS</td>
<td></td>
<td>0.0080***</td>
<td>0.0002</td>
<td>0.0074***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.3800</td>
<td>0.0800</td>
<td>3.5800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0008)</td>
<td>(0.9345)</td>
<td>(0.0004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td></td>
<td>-0.0028*</td>
<td>0.0132***</td>
<td>-0.0034***</td>
<td>0.0132***</td>
<td>0.0192</td>
<td>0.0017</td>
<td>0.0032</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.8600</td>
<td>16.2000</td>
<td>-2.4600</td>
<td>16.0000</td>
<td>1.3900</td>
<td>0.2400</td>
<td>0.5300</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0635)</td>
<td>(0.0001)</td>
<td>(0.0141)</td>
<td>(0.0001)</td>
<td>(0.1659)</td>
<td>(0.8102)</td>
<td>(0.5984)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.7547</td>
<td>15.4286</td>
<td>-2.3208</td>
<td>15.2381</td>
<td>1.3469</td>
<td>0.2229</td>
<td>0.4812</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0793)</td>
<td>(0.0001)</td>
<td>(0.0203)</td>
<td>(0.0001)</td>
<td>(0.1780)</td>
<td>(0.8236)</td>
<td>(0.6304)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = 890</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PANEL B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>240 M Roll Reg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS</td>
<td></td>
<td>0.0085***</td>
<td>0.0000</td>
<td>0.0074***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.8700</td>
<td>0.0100</td>
<td>4.6300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0042)</td>
<td>(0.9941)</td>
<td>(0.0001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td></td>
<td>-0.0022*</td>
<td>0.0111***</td>
<td>-0.0031***</td>
<td>0.0112***</td>
<td>-0.0136</td>
<td>0.0209</td>
<td>-0.0135</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.8900</td>
<td>18.8900</td>
<td>-2.5300</td>
<td>19.1600</td>
<td>-0.7400</td>
<td>1.7600</td>
<td>-1.4600</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0590)</td>
<td>(0.0001)</td>
<td>(0.0115)</td>
<td>(0.0001)</td>
<td>(0.4606)</td>
<td>(0.0790)</td>
<td>(0.1451)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.7830</td>
<td>17.9905</td>
<td>-2.3868</td>
<td>18.2476</td>
<td>-0.7171</td>
<td>1.6349</td>
<td>-1.3257</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0746)</td>
<td>(0.0001)</td>
<td>(0.0170)</td>
<td>(0.0001)</td>
<td>(0.4733)</td>
<td>(0.1021)</td>
<td>(0.1849)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = 729</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table presents GLS and GME estimates using Fama-MacBeth regressions. The results correspond to the single factor market model (with and without intercepts) and the 4-factor model of Fama-French-Carhart. Excess returns are estimated using monthly returns on 25 portfolios ranked by size (SMB) and book-to-market (HML). Panel A presents results using a window of 36 months in the first pass, and Panel B using a window of 240 months in the first pass. The sample period goes from 1927 to 2007. t-ratios and p-values are reported with (sh) and without (fm) Shanken’s (1992) correction for the error invariables problem (EIV). *** denotes 1% statistical significance level, **5% statistical significance level, and *10% statistical significance level.
Table 4
Robustness check I
Single period CAPM - (1927-2010)

<table>
<thead>
<tr>
<th>PANEL A</th>
<th>36 M Roll Reg</th>
<th>CAPM w/int</th>
<th>CAPM wo/int</th>
<th>FF4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_0$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_0$</td>
</tr>
<tr>
<td>GLS</td>
<td>Estimate</td>
<td>0.0072***</td>
<td>0.0010</td>
<td>0.0074***</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{fm}$</td>
<td>3.0700</td>
<td>0.3500</td>
<td>3.6600</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.0022)</td>
<td>(0.7251)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{sh}$</td>
<td>2.8962</td>
<td>0.3333</td>
<td>3.4857</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.0038)</td>
<td>(0.7389)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>GME</td>
<td>Estimate</td>
<td>-0.0002</td>
<td>0.0122***</td>
<td>-0.0030*</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{fm}$</td>
<td>-0.1500</td>
<td>16.6500</td>
<td>-1.9300</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.8838)</td>
<td>(0.0001)</td>
<td>(0.0544)</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{sh}$</td>
<td>-0.1415</td>
<td>15.8571</td>
<td>-1.8208</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.8838)</td>
<td>(0.0001)</td>
<td>(0.0686)</td>
</tr>
</tbody>
</table>

N = 923

<table>
<thead>
<tr>
<th>PANEL B</th>
<th>240 M Roll Reg</th>
<th>CAPM w/int</th>
<th>CAPM wo/int</th>
<th>FF4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_0$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_0$</td>
</tr>
<tr>
<td>GLS</td>
<td>Estimate</td>
<td>0.0123***</td>
<td>-0.0031</td>
<td>0.0079***</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{fm}$</td>
<td>4.9200</td>
<td>-1.1500</td>
<td>4.7300</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.0001)</td>
<td>(0.2513)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{sh}$</td>
<td>4.6415</td>
<td>-1.0952</td>
<td>4.5048</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.0001)</td>
<td>(0.2513)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>GME</td>
<td>Estimate</td>
<td>-0.0002</td>
<td>0.0123***</td>
<td>-0.0021</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{fm}$</td>
<td>-0.1500</td>
<td>16.6500</td>
<td>-1.5900</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.8838)</td>
<td>(0.0001)</td>
<td>(0.1124)</td>
</tr>
<tr>
<td></td>
<td>$t$-ratio$_{sh}$</td>
<td>-0.1415</td>
<td>15.8571</td>
<td>-1.5000</td>
</tr>
<tr>
<td></td>
<td>(p-value)</td>
<td>(0.8838)</td>
<td>(0.0001)</td>
<td>(0.1336)</td>
</tr>
</tbody>
</table>

N = 765

Note: This table presents GLS and GME estimates using Fama-MacBeth regressions. The results correspond to the single factor market model (with and without intercepts) and the 4-factor model of Fama-French-Carhart. Excess returns are estimated using monthly returns on 25 portfolios ranked by size (SMB) and book-to-market (HML). Panel A presents results using a window of 36 months in the first pass, and Panel B using a window of 240 months in the first pass. The sample period goes from 1927 to 2010. t-ratios and p-values are reported with (sh) and without (fm) Shanken’s (1992) correction for the error invariables problem (EIV). *** denotes 1% statistical significance level, **5% statistical significance level, and *10% statistical significance level.
Table 5  
Robustness check II  
Single period CAPM - (1927-2010)

<table>
<thead>
<tr>
<th>PANEL A</th>
<th>CAPM w/int</th>
<th>CAPM wo/int</th>
<th>FF4</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 M Roll Reg</td>
<td>$\lambda_0$</td>
<td>$\lambda_{MKT}$</td>
<td>$\lambda_{MKT}$</td>
</tr>
<tr>
<td>GLS</td>
<td>0.0065638***</td>
<td>0.0011</td>
<td>0.0070***</td>
</tr>
<tr>
<td>t-ratio/fm</td>
<td>4.4000</td>
<td>0.5300</td>
<td>3.5600</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0001)</td>
<td>(0.5928)</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>t-ratio/sh</td>
<td>4.1509</td>
<td>0.5048</td>
<td>3.3905</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0001)</td>
<td>(0.6137)</td>
<td>(0.0007)</td>
</tr>
</tbody>
</table>

GME

<table>
<thead>
<tr>
<th>PANEL B</th>
<th>240 M Roll Reg</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS</td>
<td>Estimate</td>
</tr>
<tr>
<td>t-ratio/fm</td>
<td>0.7600</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.4488)</td>
</tr>
<tr>
<td>t-ratio/sh</td>
<td>0.7170</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.4734)</td>
</tr>
</tbody>
</table>

GME

<table>
<thead>
<tr>
<th>PANEL B</th>
<th>240 M Roll Reg</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS</td>
<td>Estimate</td>
</tr>
<tr>
<td>t-ratio/fm</td>
<td>-1.2900</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.1970)</td>
</tr>
<tr>
<td>t-ratio/sh</td>
<td>-1.2170</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.2236)</td>
</tr>
</tbody>
</table>

N = 890

N = 729

Note: This table presents GLS and GME estimates using Fama-MacBeth regressions. The results correspond to the single factor market model (with and without intercepts) and the 4-factor model of Fama-French-Carhart. Excess returns are estimated using an augmented set of monthly returns on the 25 portfolios ranked by size (SMB) and book-to-market (HML) plus 30 portfolios ranked by industry. Panel A presents results using a window of 36 months in the first pass, and Panel B using a window of 240 months in the first pass. The sample period goes from 1927 to 2010. t-ratios and p-values are reported with (sh) and without (fm) Shanken’s (1992) correction for the error invariables problem (EIV). *** denotes 1% statistical significance level, **5% statistical significance level, and *10% statistical significance level.
<table>
<thead>
<tr>
<th>Study</th>
<th>Risk premium</th>
<th>Ambiguity premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siegel</td>
<td>4.10%</td>
<td>[9.32%-11.74%]</td>
</tr>
<tr>
<td>Shiller</td>
<td>5.75%</td>
<td>[7.57%-10.09%]</td>
</tr>
<tr>
<td>Mehra-Prescott</td>
<td>6.92%</td>
<td>[6.40%-8.92%]</td>
</tr>
<tr>
<td>Pástor-Stambaugh</td>
<td>[4%-6%]</td>
<td>[7.32%-11.84%]</td>
</tr>
</tbody>
</table>

Note: This table presents plausible values for the equity risk premium and ambiguity premia.
Table 7  
Empirical Lagrange multipliers ranked by size (ME) and book-to-market (BE/ME)

<table>
<thead>
<tr>
<th>Thetas</th>
<th>Low/Growth</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>High/Value</th>
</tr>
</thead>
</table>
| Small  | -0.0046    | 0.0003| 0.0025| 0.0046| 0.0062     | 0.0018  
| 2      | -0.0023    | 0.0022| 0.0034| 0.0040| 0.0042     | 0.0023  
| 3      | -0.0008    | 0.0019| 0.0025| 0.0029| 0.0036     | 0.0020  
| 4      | -0.0002    | 0.0007| 0.0021| 0.0029| 0.0030     | 0.0017  
| Big    | -0.0004    | 0.0000| 0.0012| 0.0010| 0.0020     | 0.0008  

-0.0017    0.0010    0.0023    0.0031    0.0038

Note: This table presents the Lagrange multipliers obtained from GME, interpreted as shadow prices of stocks’ ambiguity ranked by size (ME) and book-to-market (BME).
### Table 8
**Binding Constraints**

<table>
<thead>
<tr>
<th>Confidence Level:</th>
<th>Given Value</th>
<th>Test Value</th>
<th>Critical Value</th>
<th>P-Value</th>
<th>Result</th>
</tr>
</thead>
</table>
| **2-tail t-test results**<br *

**Thetas Averaged and Ranked by BE/ME**

H0: Mean equals 0

|  | | | | | |
|---|---|---|---|---|
| Growth | 0 | -7.0563 | 5.5976 | 0.0021 | Reject |
| 2 | 0 | -10.5024 | 5.5976 | 0.0005 | Reject |
| 3 | 0 | -14.6190 | 5.5976 | 0.0001 | Reject |
| 4 | 0 | -6.5465 | 5.5976 | 0.0028 | Reject |
| Value | 0 | -6.6691 | 5.5976 | 0.0026 | Reject |
| **t-test results**<br *

**Thetas Averaged and Ranked by ME**

|  | | | | | |
|---|---|---|---|---|
| Small | 0 | -2.59154 | 5.5976 | 0.0606 | Fail to Reject |
| 2 | 0 | -3.37095 | 5.5976 | 0.0280 | Fail to Reject |
| 3 | 0 | -5.43695 | 5.5976 | 0.0056 | Fail to Reject |
| 4 | 0 | -7.47512 | 5.5976 | 0.0017 | Reject |
| Big | 0 | -11.7303 | 5.5976 | 0.0003 | Reject |

Note: This table presents the results of a 2-tail t-test of the null hypothesis that the mean of the Lagrange multipliers ranked across book-to-market and size are equal to zero with a 99% confidence level.
### Table 9

Test assets’ natural weights ranked by size (ME) and book-to-market (BE/ME) 1927-2007

<table>
<thead>
<tr>
<th></th>
<th>Low/Growth</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>High/Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.0414</td>
<td>0.0366</td>
<td>0.0405</td>
<td>0.0407</td>
<td>0.0408</td>
<td>0.2000</td>
</tr>
<tr>
<td>2</td>
<td>0.0417</td>
<td>0.0404</td>
<td>0.0406</td>
<td>0.0407</td>
<td>0.0407</td>
<td>0.2041</td>
</tr>
<tr>
<td>3</td>
<td>0.0429</td>
<td>0.0403</td>
<td>0.0405</td>
<td>0.0406</td>
<td>0.0406</td>
<td>0.2050</td>
</tr>
<tr>
<td>4</td>
<td>0.0490</td>
<td>0.0389</td>
<td>0.0404</td>
<td>0.0405</td>
<td>0.0406</td>
<td>0.2094</td>
</tr>
<tr>
<td>Big</td>
<td>0.0451</td>
<td>0.0168</td>
<td>0.0398</td>
<td>0.0395</td>
<td>0.0403</td>
<td>0.1816</td>
</tr>
</tbody>
</table>

|      | 0.2202     | 0.1730| 0.2018| 0.2020| 0.2030     |

Note: This table presents the natural weights of each test asset calculated as a function of their Lagrange multipliers using equation B.9 in Appendix B. The weights are ranked by size (ME) and book-to-market (BE/ME).
Figure 1. Shadow prices of entropy across book-to-market ranked portfolios

Figure 2. Shadow prices of entropy across size ranked portfolios