LECTURE NOTES 11

ARIEL M. VIALE

1. Incomplete Information and Learning in Asset Pricing

1.1. **Filtering theory, Kalman-Bucy filter, and Markov chains.** Unlike engineers, economists use the term "signal" from the standpoint of the "receiver" not the "sender". Assume the complete probability space $(\Omega, \Im, P)$, where the filtering problem of the "receiver" is to estimate a hidden process X from the observations of another process Y (i.e., the signal). The goal is to estimate the conditional expectation $E_t\left[f\left(X_t\right)\middle|\mathcal{F}_t^Y\right]$, where $\left\{\mathcal{F}_t^Y\right\}$ is the filtration generated by Y augmented by the P-null (i.e., negligible) sets in $\Im$, and f is some arbitrary real-valued function satisfying weak regularity conditions. Note that by solving the filtering problem, the "receiver" also obtains the distribution of X_t conditional on the filtration $\mathcal{F}_t^Y$ generated by Y. The filtration $\mathcal{F}_t^Y$ is an increasing sequence (i.e., ordered set) of collection of events (the σ-fields or σ-algebras) where some measure (the probability of the event) is defined. Intuitively, the filtration $\mathcal{F}$ encodes all the information in the data up to time t (without memory). Let W be an n-dimensional Wiener process on its own filtration $\mathcal{F}_t$ generated by $(X_s, W_s; s \leq t)$ and augmented by the P-null sets in $\Im$. We assume that $\mathcal{F}_t$ is independent of the filtration generated by$(W_v - W_u; t \leq u \leq v \leq T)$, which means that future changes in the Wiener process cannot be predicted by X. In a nutshell, we assume that all processes are $\{\mathcal{F}_t\}$-adapted and the Wiener process W creates all the noise that must be "filtered" from the observation process Y. Thus,

$$dY_t = \kappa_t dt + dW_t;\ Y_0 = 0, \tag{1.1}$$

where κ_t is some jointly measurable $\mathbb{R}^n$-valued process satisfying the "Mean-Square-Integrability" condition $E_t \int_0^T \left\|\kappa_t\right\|^2 dt < \infty$. Furthermore, assume that X takes values in some complete separable metric space and,

$$df_t = g_t dt + dM_t, \tag{1.2}$$

where g_t is some jointly measurable process and M is a right-continuous Martingale satisfying the usual technical mean-square integrability condition. The innovation process is defined as,

$$\begin{aligned} dZ_t &= dY_t - \hat{\kappa_t} dt, \\ &= \left(\kappa_t - \hat{\kappa_t}\right) dt + dW_t,\ \ Z_0 = 0. \end{aligned} \tag{1.3}$$

Note that the innovation, shock or surprise in Y consists of two parts: 1) the estimation error in the drift κ_t; and 2) the white noise dW. Following standard filtering theory:

(A.1) The innovation process Z is an $\left\{\mathcal{F}_t^Y\right\}$-Brownian motion. That is, Z is a martingale and the innovations dZ are "unpredictable".

Date: 11/19/13.

(A.2) For any separable L^2-bounded $\{\mathcal{F}_t^Y\}$-martingale K, $\exists$ a jointly measurable $\{\mathcal{F}_t^Y\}$-adapted $\mathbb{R}^n$-valued process ψ satisfying the mean-square-integrability condition and,

$$dK_t = \sum_{i=1}^{n} \psi_t^i dZ_t^i.$$

That is the process Z spans the $\{\mathcal{F}_t^Y\}$-martingales.

(A.3) $\exists$ a jointly measurable adapted process α^i such that $d\left[M, W^i\right]_t = \alpha_t^i dt$ for $i = 1, \ldots, N$. In plain words, the square-bracketed processes are absolutely continuous, although maybe independent, which in that case results in $\alpha^i = 0 \ \forall i$.

(A.4) $\hat{f}$ evolves as,

$$d\hat{f}_t = \hat{g}_t dt + \left(\widehat{f\kappa}_t - \hat{f}_t\hat{\kappa}_t + \hat{\alpha}_t\right)' dZ_t, \tag{1.4}$$

where $\widehat{f\kappa}_t = E_t\left[f_t\kappa_t\ \mathcal{F}_t^Y\right]$; and $\left(\widehat{f\kappa}_t - \hat{f}_t\hat{\kappa}_t\right)$is the covariance between f_t and κ_t. This is the general filtering formula in which the estimate $\hat{f}$ is updated each period because f is expected to change given $\hat{g}_t dt$ and the new information conveyed by dZ. Notice that the general filtering formula generalizes the linear prediction formula $\hat{x}-\bar{x} = \frac{cov(x,y)}{var(y)}\left(y - \bar{y}\right)$.

<u>Kalman-Bucy-filter</u>

Assume X is distributed as a Normal variate with varance σ^2 such that,

$$dX_t = aX_t dt + dB_t,$$

$$dY_t = cX_{tt}dt + dW_t,$$

where B and W are independent real-valued Brownian motions independent of X_0. In this case the distribution of X_t conditional on $\mathcal{F}_t^Y$ is Normal with deterministic variance Σ_t and,

$$d\hat{X}_t = a\hat{X}_t dt + c\Sigma_t dZ_t,$$

$$dZ_t = dY_t dt - c\hat{X}_t dt.$$

The state and observation equations represent the filtering problem in state space form.

<u>Two-state-Markov-Chain</u>

We drop Normality and assume,

$$dX_t = \left(1 - X_{t-}\right) dN_t^0 - X_{t-}dN_t^1,$$

where $X_{t-} \equiv lim_{s\uparrow t} X_s$ and the N^i are independent Poisson processes with parameters λ^i independent of X_0. Intuitively, X stays in each state an exponentially distributed amount of time with λ^i denoting the transition from state i to state j. Thus now,

$$g_t = \left(1 - X_{t-}\right)\lambda^0 - X_{t-}\lambda^1,$$

$$dM_t = (1 - X_{t-})\, dM_t^0 - X_{t-} dM_t^1,$$

where $M_t^i = N_t^i - \lambda_t^i$ is a martingale. Assume,

$$dY_t = \kappa X\,(X_{t-})\, dt + dW_t,$$

where W is a n-dimensional Brownian motion independent of N^i and X_0. Note also that the drift κ is now conditional on the state X_{t-}. Thus, we write π_t for $\hat{X}_t$ denoting the conditional probability that $X_t = 1$. Then the general filtering formula is now,

$$d\pi_t = \left[(1 - \pi_t)\, \lambda^0 - \pi_t \lambda^1\right] dt + \pi_t\,(1 - \pi_t)\,[\kappa\,(0) - \kappa\,(1)]^{'}\, dZ_t,$$

$$dM_t = (1 - X_{t-})\, dM_t^0 - X_{t-} dM_t^1,$$

where the innovation process has dynamics (Wonham, 1965),

$$dZ_t = dY_t - [(1 - \pi_t)\, \kappa\,(0) + \pi_t \kappa\,(1)]\, dt. \tag{1.5}$$

Note that $[\kappa\,(0) - \kappa\,(1)]$ is the "gain" vector c of the Kalman-Bucy filter and $\pi_t\,(1 - \pi_t)$ is the variance of the Kalman-Bucy filter conditional on $\mathcal{F}_t^Y$.

1.2. **Markov chain models of pure exchange economies.** Assume a pure exchange economy a la Lucas in which assets (i.e., trees) are in fixed supply and the behavioral economic problem of the RA is to allocate consumption given assets' dividends (i.e., fruit). Following Veronesi (1999, 2000), we assume that there is a single risky asset with supply normalized to one and paying dividends at the rate D such that,

$$\frac{dD_t}{D_t} = \alpha_D\,(X_{t-})\, dt + \sigma_D dW^1,$$

where X is a two-state Markov chain mimicking the dynamics of the real business cycle. That is, the economy is assumed to switch between a good and bad economic regime at exponentially distributed times; and W^1 is a real-valued Brownian motion independent of X_0. Investors observe the dividend rate D, but do not observe the state vector X_{t-}, which determines the growth rate of dividends. Also, assume that investors observe the process,

$$dK_t = \alpha_K\,(X_{t-})\, dt + \sigma_K dW^2,$$

where W^2 is another real-valued Brownian motion independent of $W^1$and X_0. This process sumarizes any other information investors may have about the state of the economy. Thus, the filtering equations in this model depend on two innovation processes i.e., $Z = \left(Z^1, Z^2\right)$ given that,

$$\frac{dD_t}{D_t} = [(1 - \pi_t)\, \alpha_D\,(0) + \pi_t \alpha_D\,(1)]\, dt + \sigma_D dZ^1, \tag{1.6}$$

and,

$$dK_t = [(1 - \pi_t)\, \alpha_K\,(0) + \pi_t \alpha_K\,(1)]\, dt + \sigma_K dZ^2. \tag{1.7}$$

And the conditional probability of the good (normal) regime corresponding to good economic times evolves as,

(1.8)

$$d\pi_t = \left[(1-\pi_t)\lambda^0 - \pi_t\lambda^1\right]dt + \pi_t(1-\pi_t)\left[\frac{\alpha_D(1)-\alpha_D(0)}{\sigma_D}dZ^1 + \frac{\alpha_K(1)-\alpha_K(0)}{\sigma_K}dZ^2\right]'.$$

The three equations form a Markovian system in which the growth rate of dividends is stochastic.

Assume an infinitely lived RA who maximizes expected discounted utility of consumption with subjective discount rate β. Then, in equilibrium the RA should consume the aggregate dividend and the price of the asset will be determined by her marginal rate of substitution,

$$S_t = E\left[\int_t^\infty \frac{e^{-\beta(s-t)}u'(D_s)}{u'(D_t)}D_s ds\,|\pi_t, D_t\right]. \tag{1.9}$$

In the special case of logarithmic utility (myopic RA) asset returns are given by,

$$\frac{dS_t}{S_t} = \left[(1-\pi_t)\alpha_D(0) + \pi_t\alpha_D(1)\right]dt + \sigma_D dZ^1. \tag{1.10}$$

This special case corresponds to the earlier literature that studied asset pricing under incomplete information. The more interesting case is the one of a RA with a power utility function $u(c) = {}^{c^\gamma}/_{\gamma}$ (constant relative risk aversion). In this case,

$$D_s^\gamma = D_t^\gamma e^{\gamma\int_t^s\left\{\left[\alpha_D(X_{a-})-\sigma_D^2/2\right]da+\sigma_D dW_a^1\right\}}, \tag{1.11}$$

and,

$$S_t = D_t^{1-\gamma}E\left[\int_t^\infty e^{-\beta(s-t)}D_s^\gamma ds\,|\pi_t, D_t\right] \tag{1.12}$$

$$= D_t^{1-\gamma}\left\{(1-\pi_t)E\left[\int_t^\infty e^{-\beta(s-t)}D_s^\gamma ds\,|X_{t-}=0, D_t\right] + \pi_t E\left[\int_t^\infty e^{-\beta(s-t)}D_s^\gamma ds\,|X_{t-}=1, D_t\right]\right\}$$

$$= D_t^{1-\gamma}\left\{(1-\pi_t)E\left[\int_t^\infty e^{-\beta(s-t)}e^{\gamma\int_t^s\left\{\left[\alpha_D(X_{a-})-\sigma_D^2/2\right]da+\sigma_D dW_a^1\right\}}ds\,|X_{t-}=0, D_t\right] + \right.$$

$$\left. +\pi_t E\left[\int_t^\infty e^{-\beta(s-t)}e^{\gamma\int_t^s\left\{\left[\alpha_D(X_{a-})-\sigma_D^2/2\right]da+\sigma_D dW_a^1\right\}}ds\,|X_{t-}=1, D_t\right]\right\}.$$

Note that given the time-homogeneity property of the Markov system [1.6-1.8], conditional expectations are independent of time. We write C_0 and C_1 for the above conditional expectations during bad and good economic times respectively and we have,

$$S_t = D_t\left\{(1-\pi_t)C^0 + \pi_t C^1\right\},$$

with returns,

$$\frac{dS}{S} = \frac{dD}{D} + \frac{\left(C^1-C^0\right)d\pi}{(1-\pi)C^0+\pi C^1} + \frac{\left(C^1-C^0\right)d(D,\pi)}{D\left[(1-\pi)C^0+\pi C^1\right]} \tag{1.13}$$

(1.14)
$$= something\ dt + \sigma_D dZ^1 + \left[\frac{\left(C^1 - C^0\right)\pi\left(1-\pi\right)}{\left(1-\pi\right)C^0 + \pi C^1}\right] \times \left[\frac{\alpha_D\left(1\right) - \alpha_D\left(0\right)}{\sigma_D} dZ^1 + \frac{\alpha_S\left(1\right) - \alpha_S\left(0\right)}{\sigma_S} dZ^2\right].$$

The factor,

$$\left[\frac{\left(C^1 - C^0\right)\pi\left(1-\pi\right)}{\left(1-\pi\right)C^0 + \pi C^1}\right] \tag{1.15}$$

introduces stochastic volatility in the stock return equation. Thus stochastic volatilty can arise in a model with constant dividend volatility. Note that the conditional variance of the state $\pi\left(1-\pi\right)$ is largest when $\pi_t = \frac{1}{2}$ as the RA is uncertain about the state of the economy. Thus, the volatility of the asset is linked to investors' confidence about future economic growth.

Veronesi (1999) assumes that the level of dividends follows an Ornstein-Uhlenbeck process with the RA endowed with a negative exponential utility function (i.e., constant absolute risk aversion). The model is capable of explaining most of the empirical puzzles in the asset pricing literature including the nonlinear relation between the drift and variance of stock returns. One controversial implication of the model though, is the result in Veronesi (2000), where there is no premium for "noisy" signals given the hedging demands of investors. Thus, the quality of the signal (proxied by its precision or the inverse of the standard deviation) does not seem to add to systematic risk. Thus investors "overreact" to bad news during good economic times and "underreact" to good news during bad economic times.

Ozoguz (2009) and Zhang (2003), are two studies that assess the empirical capabilities of the model of Veronesi (1999). They derive the fundamental asset pricing equation of the economy as a conditional scaled CAPM, with the conditional probability of good economic times as the scaling variable. The unconditional asset pricing equation in beta regression form includes additional factors beyond market beta related to: 1) "learning" proxied by changes in the conditional probability π_t; and 2) $UC = \pi\left(1-\pi\right)$ proxying for investors' "uncertainty" about the state of the economy. In a nutshell their results show that the "learning" CCAPM does a good job in explaining the cross-section of average stock returns in the U.S. However, the results show that the "learning" factor seems to be not robust and at sometimes with the wrong sign. Also, the inclusion of additional factor loadings on cross-effects that result from the wedge between the information set of the investor and the econometrician but with no economic content seems controversial. As shown by Viale et. al. (2013), the results seem to obey to the presence of model misspecification. This is relevant, because this model is still rooted in the Rational Expectations Hypothesis (REH), which assumes that investors "know" the data generation process (DGP) driving dividends and stock returns. Their uncertainty is only limited to the state of the economy, not the DGP, which they seek to learn updating their models as "standard" Bayesian agents.

1.3. **Learning under ambiguity in pure exchange economies.** The model in Viale, Garcia-Feijoo, and Giannetti (2013) is an example of a pure exchange economy a la Lucas like the ones discussed bedore but where investors are averse to ambiguity. The motivation for ambiguity comes from the experiments of Ellsberg, but ambiguity may arise by pure statistical arguments given the always present limitations of data under the statistical learning paradigm.

(A.1) Like in the standard Bayesian case, the economy switches between the good economic state (labeled 1) and the bad economic state (labeled 0).

(A.2) For each path ω_t (a random sequence of 0 and 1 moves), the investor chooses his level of consumption C_t as well as the shares $\xi_{i,t}$ of wealth W_t allocated to $i = 1, \ldots, n$ risky assets and the risk-free asset.

(A.3) At time $t+1$, conditional wealth is W^s_{t+1}, where $s_t = \{0, 1\}$ represents the bad (respectively, good) transition of the economy from prior state ω_t.

(A.4) The typical investor is ambiguous (in the sense of Knight, Keynes, Shackle, Roy, and Ellsberg) about the true one-step-ahead conditional probabilities he should attach to the good state of the economy.

(A.5) He holds some prior beliefs $\pi_t(\omega_t)$ possibly derived from historical analysis, which defines his "reference model". However, acknowledging his limitations doubts the reference model and consequently entertains a set of plausible alternative (distorted and close i.e., difficult to distinguish) models $\pi^*_t(\omega_t)$around the reference model i.e., the so-called multiple-priors set that differentiates the robust investor from his standard Bayesian counterpart. This recursive preferences setup with multiple priors closely the approach introduced by Epstein and Wang (1994) and formalized in Epstein and Schneider (2003).

(A.6) The statistical distance between the alternative models and the reference model is restricted by,

$$D_t\left(\pi^*_t \,\|\pi_t\right) \leq \eta_t, \tag{1.16}$$

where $D_t\left(\pi^*_t \,\|\pi_t\right)$ denotes the Kullback-Leibler divergence or "relative entropy"; and η_t is an exogenous state-dependent ambiguity parameter that restricts what Epstein and Schneider call the entropy-constrained ball containing all alternative measures π^*. In plain words, the scalar is related to the typical investor's confidence level on his reference model, and avoids the degenerate case of infinite ambiguity aversion.

The ambiguity averse investor solves the following "robust" Hamilton-Jacobi-Bellman equation with no regrets using backward induction,

,

(1.17)

$$J(W_t, t) = \underset{\{C_t, \xi_{i,t}\}}{Max}\left\{U(C_t) + \underset{\{\pi^*_t \in \prod\}}{Min} E_t\left[\pi^*_t J\left(W^1_{t+1}, t+1\right) + (1-\pi^*_t) J\left(W^0_{t+1}, t+1\right)\right]\right\},$$

subject to the usual budget constraint,

$$W^s_{t+1} = (W_t - C_t)\left[R_{f,t} + \sum_{i=1}^{n} \xi_{i,t}\left(R^s_{i,t+1} - R_f\right)\right], \tag{1.18}$$

where $E_t[\cdot]$ is the conditional expectation operator at time $t = (0, 1, \ldots, T-1)$; $U(C_t)$ is increasing and concave; $J(W_t, t)$ denotes the usual indirect utility function on wealth; $R_{f,t}$ is the gross return of one dollar invested in the risk-free asset; $R^s_{i,t+1}$ is the conditional gross return of risky asset i at period $t+1$; and the rest of the variables are defined as before.

The economic interpretation of the dynamic MaxMin problem is as follows. The ambiguity averse investor first solves the inner constrained minimization problem to identify the worst-case scenario among all alternative models π_t^* given the reference model π_t. The solution is a conditional measure (i.e., what the authors label as the distorted probability of good economic times π_t^L) that allows the investor to calculate the "ambiguity certainty equivalent" of the continuation value function $J\left(W_{t+1}^s, t+1\right)$. In the second (maximization) step, the investor proceeds in the usual way solving a standard expected utility maximization problem, although under the "distorted" probability measure.

Iterating backwards period by period, and accounting for the ambiguity constraint (1) gives the dynamically constrained H-J-B equation,

$$J\left(W_t,t\right) = \underset{\{C_t,\xi_{i,t}\}}{Max}\left\{U\left(C_t\right) + \underset{\{\pi_t^*\in\prod\}}{Min}\,E_t\left[\pi_t^* J\left(W_{t+1}^1,t+1\right) + \left(1-\pi_t^*\right)J\left(W_{t+1}^0,t+1\right) + \frac{1}{2}\theta_t\left(D_t\left(\pi_t^*\,\|\pi_t\right)-\eta_t\right)\right]\right\},$$

where $\theta_t \geq 0$ is the Lagrange multiplier corresponding to the ambiguity constraint (1).

Relative entropy can be defined in terms of log-likelihood ratios,

$$D_t\left(\pi_t^*\,\|\pi_t\right) = E_{\pi_t^*}\left[ln\left(\frac{\pi_t^*}{\pi_t}\right)\right] = \pi_t^* ln\left(\frac{\pi_t^*}{\pi_t}\right) + \left(1-\pi_t^*\right)ln\left(\frac{1-\pi_t^*}{1-\pi_t}\right).$$

Where a first order Taylor series approximation yields,

$$ln\left(\frac{\pi_t^*}{\pi_t}\right) \approx \frac{\pi_t^*-\pi_t}{\pi_t},$$

which is well behaved since $-1 \leq \frac{\pi_t^*-\pi_t}{\pi_t} \leq 1$ as long as π_t is non-degenerate. Substituting the previous equation into the definition of relative entropy results in,

$$ln\left(\frac{\pi_t^*}{\pi_t}\right) \approx \frac{\left(\pi_t-\pi_t^*\right)^2}{\pi_t\left(1-\pi_t\right)},$$

which in the entropy-based econometric literature is known as a "shrinkage estimator". After substituting this approximation into the H-J-B equation one solves the inner minimization problem with F.O.N.C.,

$$\pi_t - \pi_t^* = \frac{\left\{E_t\left[J\left(W_{t+1}^1,t+1\right)\right] - E_t\left[J\left(W_{t+1}^0,t+1\right)\right]\right\}\pi_t\left(1-\pi_t\right)}{\theta_t}, \tag{1.19}$$

and compementary slackness condition,

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

Substituting this into the F.O.N.C. gives,

$$\theta_t = \frac{\left\{E_t\left[J\left(W_{t+1}^1,t+1\right)\right] - E_t\left[J\left(W_{t+1}^0,t+1\right)\right]\right\}\pi_t\left(1-\pi_t\right)}{\sqrt{2\eta_t}}. \tag{1.20}$$

Substituting θ_t back into the F.O.N.C. yields the optimal "lower bound" for the probability of good economic times, which corresponds to the worst case scenario,

$$\pi_t^L \equiv \pi_t - \sqrt{2\eta_t \pi_t (1-\pi_t)}, \tag{1.21}$$

such that $0 \leq \frac{\pi_t - \pi_t^L}{\pi_t} \leq 1$. When equation (21) holds the robust H-J-B equation turns into the usual one,

$$J(W_t, t) = \underset{\{C_t, \xi_{i,t}\}}{Max} \left\{ U(C_t) + E_t^{\pi_t^L} [J(W_{t+1}, t+1)] \right\}, \tag{1.22}$$

where $E_t^{\pi_t^L} [J(W_{t+1}, t+1)] \equiv \pi_t^L J\left(W_{t+1}^1, t+1\right) + \left(1-\pi_t^L\right) J\left(W_{t+1}^0, t+1\right)$ is the ambiguity certainty equivalent of the expected continuation values under the worst-case scenario. Recall that solving this maximization problem leads to the envelope theorem condition and the fundamental asset pricing equation, which under ambiguity is,

$$1 = E_t^{\pi_t^L} [M_{t,t+1} R_{i,t+1}]. \tag{1.23}$$

The final step required for empirical implementations, is to know what is $M_{t,t+1}$. One way to do this, is to follow the conditional asset pricing literature and impose an affine factor structure for the SDF,

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

where the column vector $f_t = \left(R_{MKT,t}^e, R_{d\pi_t^L}^e, R_{dKUNC_t}^e\right)$ includes the excess market return, and excess returns of two tracking portfolios proxying for innovations in the distorted probability π_t^L and the measure of investors' ambiguity$KUNC = \frac{\pi_t - \pi_t^L}{\pi_t}$ as factors. The time-varying coefficients $\tilde{\phi}_{t-1}' = \left(\phi_{t-1}^0, \phi_{t-1}^{f'}\right)$ are also assumed to be an affine function of innovations in a vector of state variables z_{t-1} driving the investment opportunity set,

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

and,

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

This step is similar to the implemenation in Ozoguz (2009) but with some important differences. Under ambiguity, $dz_{t-1} = d\pi_{t-1}^L$ is the scaling variable that should be used to condition down the asset pricing model. Substituting into the SDF gives the unconditional moment condition,

$$1 = E\left[\left(a^0 + b^0 d\pi_{t-1}^L + a^f f_t + b^f d\pi_{t-1}^L f_t\right) R_{i,t}\right] \; \forall i \; risky \; assets. \tag{1.24}$$

By construction, the factors are orthogonal and the scaling variable is white noise. Thus, the additional cross-term factors in Ozoguz (2009) that arise from the wedge between the information set of the typical investor and the econometrician are not present here.

Under ambiguity, the typical investor acts as a robust econometrician that acknowledges his limitations and consequently doubts his reference model. So there is no need for extra factors with no economic content, beyond the ones that span the cross section of stock returns from theory and therefore are included in the SDF.

The resulting empirical specification, which is suitable to be tested empirically, is a three factor model with market beta and loadings on two additional factors proxying for the three dimensions of systematic uncertainty: 1) systematic risk; 2) uncertainty regarding the state of the economy (which could be interpreted as long-run risk); and 3) investors' uncertainty regarding the reference model .

Viale et. al. (2013) test this model and an alternative version of it that includes five factors after subsituting the distorted probability with the vector of state variables in Petkova (2006). This alternative especification can be interpreted as a robust version of the empirical ICAPM of Petkova (2006), and the empirical counterpart of Epstein and Schneider's (2008) learning model when macro news are assumed to be ambiguous. The authors find that ambiguity is priced in the cross-section of average stock returns and that the ambiguity factor is not subsumed by popular firm-specific and macro factors previously discussed in the asset pricing literature. The learning under ambiguity model performs better than the Bayesian learning model of Ozoguz (2009), the ICAPM of Petkova (2006), and the Fama-French-Carhart empirically motivated model.

An important implication of this model is that unlike with noisy signals, ambiguous news seem to carry a premium across stocks and consequently, robust investors always discount more bad news than good news about the continuation probability of the good economic regime, independently of the state of the economy. This hypothesis is tested succesfully in Giannetti and Viale (2013).

References

[Epstein and Schneider (2003) Journal of Economic Theory]
[Epstein and Schneider (2008) Journal of Finance]
[Epstein and Wang (1994) Econometrica]
[Giannetti and Viale (2013) Working Paper]
[Ozozguz (2009) Review of Financial Studies]
[Petkova (2006) Journal of Finance]
[Veronesi (1999) Review of Financial Studies]
[Veronesi (2000) Journal of Finance]
[Viale, Garcia-Feijoo, and Giannetti (2013) Review of Asset Pricing Studies]
[Wang (2003) Working Paper]
[Wonham (1965) SIAM Journal of Control]