Ambiguity and Asset Markets

Larry G. Epstein[1] and Martin Schneider[2]

[1]Department of Economics, Boston University, Boston, Massachusetts 02215; email: lepstein@bu.edu

[2]Department of Economics, Stanford University, Stanford, California 94305; email: schneidr@stanford.edu

Annu. Rev. Financ. Econ. 2010. 2:10.1–10.32

The *Annual Review of Financial Economics* is online at financial.annualreviews.org

This article's doi:
10.1146/annurev-financial-120209-133940

1941-1367/10/1205-0001$20.00

Key Words

portfolio choice, asset pricing

Abstract

The Ellsberg paradox suggests that people's behavior is different in risky situations—when they are given objective probabilities—from their behavior in ambiguous situations—when they are not told the odds (as is typical in financial markets). Such behavior is inconsistent with subjective expected utility (SEU) theory, the standard model of choice under uncertainty in financial economics. This article reviews models of ambiguity aversion. It shows that such models—in particular, the multiple-priors model of Gilboa and Schmeidler—have implications for portfolio choice and asset pricing that are very different from those of SEU and that help to explain otherwise puzzling features of the data.

1. INTRODUCTION

The first part of this article (Section 2) recalls the Ellsberg-based critique of subjective expected utility theory and then outlines some of the models that it has stimulated. Our coverage of preference models is selective—we focus only on models that have been applied to finance or that seem promising for future applications: multiple priors (Gilboa & Schmeidler 1989), "smooth ambiguity" preference (Klibanoff et al. 2005), multiplier utility, and related robust-control-inspired models (Hansen & Sargent 2001, Maccheroni et al. 2006a). We provide a unifying framework for considering the various models. A confusing aspect of the literature is the plethora of seemingly different models, rarely related to one another, and often expressed in drastically different formal languages. Here we put several of these models side-by-side, expressed in a common language, and we examine the properties of each with respect to implications both for one-shot-choice and for sequential choice. In particular, we provide thought experiments to illustrate differences in behavior implied by the various models. We thereby hope to facilitate a more informed choice between models.

The second part of the article (Section 3) derives implications of the models for finance. One common theme shared by all models is that ambiguity-averse agents choose more conservative positions and, in equilibrium, command additional "ambiguity premia" on uncertain assets. Ambiguity aversion can thus help to account for position and price behavior that is quantitatively puzzling in light of subjective expected utility (SEU). A second common theme is that, in dynamic settings, ambiguity-averse agents may adjust their positions to account for future changes in ambiguity, for example, due to learning. This adds a new reason for positions to differ by investment horizon, and, in equilibrium, generates time variation in premia.

Models of ambiguity aversion differ in how ambiguity aversion compares with risk aversion and, thus, in how implications for portfolio choice and asset pricing differ from those of SEU. On the one hand, many of the qualitative implications of multiplier utility and of the smooth ambiguity model are identical to those of SEU. In all three models, with standard specifications, agents are locally risk neutral, portfolios react smoothly to changes in return expectations, and diversification is always beneficial. Consequently, in many settings, the multiplier and smooth models do not expand the range of qualitative behavior that can be explained by SEU. Instead, they offer reinterpretations that may be quantitatively more appealing (for example, ambiguity aversion can substitute for higher risk aversion).

On the other hand, most applications of the multiple-priors model have exploited qualitative differences from SEU. These arise because the multiple-priors model allows uncertainty to have first-order effects on portfolio choice and asset pricing. Thus the model can give rise to selective participation, optimal underdiversification, and portfolio inertia at portfolios that hedge ambiguity. In heterogeneous agent models with multiple priors, portfolio inertia has been used to endogenously generate incompleteness of markets and to account for markets "freezing up" in response to an increase in uncertainty. Finally, uncertainty has a first-order effect on average excess returns, which can be large even if the covariance of payoffs with marginal utility is negligible.

2. MODELS OF PREFERENCE

This section is divided into two major parts. In Section 2.1, we consider static or one-shot-choice settings where all choices are made at a single instant prior to the resolution of uncertainty. Models of preference under uncertainty are typically formulated first for such static settings. However, just as in Epstein & Zin (1989), which studies risk preferences, any such model of static preference can be extended uniquely into a recursive dynamic model of preference. Therefore, the discussion of static models is also revealing about their dynamic extensions, which are outlined in Section 2.2. In addition, a dynamic setting, where choice is sequential, raises new issues—dynamic consistency and updating or learning—and these are the major focus of Section 2.2.

2.1. Static or One-Shot Choice Settings

The first subsection (Section 2.1.1) reviews the Ellsberg paradox and the formal setup. Then, the further subsections introduce the different models of preference.

2.1.1. Ellsberg and the formal set up. Ellsberg's (1961) classic experiments motivate the study of ambiguity. In a variant of one of his experiments, you are told that there are 100 balls in an urn and that each ball is either red or blue. You are not given further information about the urn's composition. Presumably you would be indifferent between bets on drawing either color (take the stakes to be 100 and 0). However, compare these bets with the risky prospect that offers you, regardless of the color drawn, a bet on a fair coin, with the same stakes as above. When you bet on the fair coin, or equivalently on drawing blue from a second risky urn when you are told that there are 50 balls of each color, then you can be completely confident that you have a 50-50 chance of winning. In contrast, in the original "ambiguous" urn, there is no basis for such confidence. This difference motivates a strict preference for betting on the risky urn as opposed to the ambiguous one.

The preference described above is incompatible with expected utility. Indeed, suppose you had in mind a subjective probability about a blue draw from the ambiguous urn. A strict preference for betting on the fair coin over a bet on a blue draw would then reveal that your probability of blue is strictly less than one-half. At the same time, a preference for betting on the fair coin over a bet on a red draw reveals a probability of blue that is strictly greater than one-half, ergo a contradiction. It follows that Ellsberg's choices cannot be rationalized by SEU. Thus SEU cannot afford a distinction between risk and ambiguity. Such a distinction is sometimes alternatively referred to as one between risk and "Knightian uncertainty" or, in terminology introduced by Hansen & Sargent (2001), as the distinction between payoff uncertainty and model uncertainty.

Ellsberg's choices have been confirmed in many laboratory experiments. But this experiment did not need to be run to be convincing—it rings true that confidence, and the amount of information underlying a likelihood assessment, matter. Such a concern is not a mistake or a form of bounded rationality—to the contrary, it would be irrational for an individual who has poor information about her environment to ignore this fact and behave as though she were much better informed. The normative significance of Ellsberg's message distinguishes it from that emanating from the Allais paradox contradicting the vNM model of preference over risky prospects.

We need some formalities to proceed. Following Savage (1954), adopt as primitives a state space Ω, representing the set of relevant contingencies or states of the world $\omega \in \Omega$ and a set of outcomes $C \subset \mathbb{R}^n_+$. (Little is lost by assuming that both Ω and C are finite and have power sets as associated σ-algebras; however, considerable generalization is possible.) Prior to knowing the true state of the world, an individual chooses once and for all a physical action. As in Anscombe & Aumann (1963), suppose that the consequence of an action is a lottery over C, an element of $\Delta(C)$. Then, any physical action can be identified with a (bounded and measurable) mapping $h : \Omega \to \Delta(C)$, which is called an Anscombe-Aumann (AA) act. Thus to model choice between physical actions, we model preference $\succeq$ on the set of AA acts.

To model the Ellsberg experiment above, take $\Omega = \{R, B\}$ as the state space, where a state corresponds to a draw from the ambiguous urn. The relevant bets are expressed as AA acts as shown in **Table 1**.

Bets on a red and a blue draw correspond to acts f_R and f_B, respectively. A bet on the fair coin corresponds to a constant AA act f_C that delivers the same lottery $(100, \frac{1}{2})$ in both states; throughout, we denote by (c, p) the lottery paying c with probability p and 0 with probability $1 - p$.

Two special subsets of acts should be noted. Call h a Savage act if $h(\omega)$ is a degenerate lottery for every ω; in that case, view h as having outcomes in C and write $h : \Omega \to C$. Both f_R and f_B above are Savage acts. For the second subset, we can identify any lottery $\ell \in \Delta(C)$ with the constant act that yields ℓ in every state. An example is the fair-coin lottery above. Consequently, any preference on AA acts includes a ranking of risky prospects. This makes clear the analytical advantage of adopting the large AA domain, because the inclusion of risky prospects makes it straightforward to describe behavior that would reveal that risk is treated differently from other uncertainty. This is a major reason that all the models of preference we discuss have been formulated in the AA framework.

Another analytical advantage of the AA domain is the simple definition it permits for the mixture of two acts. The mixture of two lotteries is well defined and familiar. Given any two AA acts h' and h, and α in [0,1], define the new act $\alpha h' + (1 - \alpha)h$ by mixing their lotteries state by state, that is,

$$\left(\alpha h' + (1-\alpha)\, h\right)(\omega) = \alpha h'(\omega) + (1-\alpha)\, h(\omega), \omega \in \Omega. \tag{2.2}$$

A key property of the Ellsberg urn is that $\frac{1}{2} f_R + \frac{1}{2} f_B = f_C$, and thus a mixture of the bets f_R and f_B gives a lottery that no longer depends on the state. Ellsberg's choices can now be written as

$$\tfrac{1}{2} f_R + \tfrac{1}{2} f_B \succ f_R \sim f_B. \tag{2.3}$$

From this perspective, Ellsberg's example has two important features. First, randomization between indifferent acts can be valuable. This is a violation of the independence axiom and thus a key departure from expected utility. Second, randomization can be valuable because

Table 1 Model of the Ellsberg experiment

Ellsberg's urn: $R + B = 100$		
	R	B
f_R	100	0
f_B	0	100
f_C	$(100, \frac{1}{2})$	$(100, \frac{1}{2})$

it can smooth out, or hedge, ambiguity. The negative comovement in the payoffs of the ambiguous acts f_B and f_R implies that the act $\frac{1}{2}f_R + \frac{1}{2}f_B$ is not ambiguous; it is simply risky. One can be confident in knowing the probabilities of the lottery payoffs, even if one is not confident in those of the underlying bets f_R and f_B.

The literature has identified the first property—a strict preference for randomization between indifferent acts—as the behavioral manifestation of (strict) ambiguity aversion. Accordingly, say that the individual with preference $\succeq$ is (weakly) ambiguity averse if, for all AA acts h' and h,

$$h' \sim h \Rightarrow \alpha h' + (1-\alpha)h \succeq h. \tag{2.4}$$

For a related comparative notion, say that individual 1 is more ambiguity averse than individual 2 if, for all AA acts h and lotteries $\ell \in \Delta(C)$,

$$\ell \succeq_2 h \Rightarrow \ell \succeq_1 h. \tag{2.5}$$

The idea is that if individual 2 rejects the ambiguous act h in favor of the risky prospect ℓ, then so should the more ambiguity-averse individual 1. The uncertainty-aversion axiom (Equation 2.4) is satisfied by all the models reviewed below.

Models of ambiguity aversion differ in why randomization is valuable, in particular, whether it can be valuable even if it does not hedge ambiguity. To see the main point, consider the following extension of the Ellsberg experiment. Let c denote the number of dollars you are willing to pay for the bet f_R. Next, imagine a lottery that delivers either the bet f_R or its certainty equivalent payoff c, each with probability $\frac{1}{2}$. How much would you be willing to pay for such a bet? One reasonable answer is c—randomizing between an asset (here a bet) and its own subjective value cannot be valuable. Intuitively, if you perceive the value of an asset to be low because you are not confident in your probability assessment of its payoff, then your confidence in your assessment should not change just because the asset is part of the lottery. As a result, the asset, its subjective value, and the lottery should all be indifferent.

The above view underlies the MP model of Gilboa & Schmeidler (1989). According to that model, preference for randomization between indifferent acts is valuable only if it hedges ambiguity and thus increases confidence, as in the Ellsberg experiment. When there is no opportunity for hedging—as in the last example where one of the acts (the subjective value of the asset) is constant—then randomization is not valuable. In contrast, "smooth" models of ambiguity aversion, in particular multiplier preferences (Anderson et al. 2003) and the smooth ambiguity model (Klibanoff et al. 2005), assume a pervasive value for randomization. Those models can rationalize Ellsberg's choices only if randomizing between an asset and its subjective value is also valuable.

2.1.2. Multiple-priors utility. When information is scarce and a single probability measure cannot be relied on to guide choice, then it is cognitively intuitive that the decision maker think in terms of a set of probability laws. For example, she might assign the interval $\left[\frac{1}{3}, \frac{2}{3}\right]$ to the probability of drawing a red ball from the ambiguous urn in the Ellsberg experiment. Being cautious, she might then evaluate a bet on red by using the minimum probability in the interval, here $\frac{1}{3}$, which would lead to the strict preference to bet on the risky urn. Similarly for blue. In this way, the intuitive choices Ellsberg highlights can be rationalized.

More formally and generally, the MP model postulates the following utility function on the set of AA acts:

$$U^{MP}(h) = \min_{p \in \mathcal{P}} \int_{\Omega} u(h)dp. \tag{2.6}$$

Here, $u : \Delta(C) \to \mathbb{R}$ is a vNM functional on lotteries that is affine, that is,

$$u(\alpha\ell + (1-\alpha)\ell') = \alpha u(\ell) + (1-\alpha)u(\ell'),$$

for all lotteries ℓ, ℓ' in $\Delta(C)$. In the following, identify c with the degenerate lottery giving c and write $u(c)$. Also, assume that u is strictly increasing for deterministic consumption. The vNM assumption for u excludes risk preferences exhibiting the Allais paradox—ambiguity is the only rationale admitted for deviating from SEU in the MP model as well as in all the other models we discuss. The central component in the functional form is the set $\mathcal{P} \subset \Delta(\Omega)$ of probability measures on Ω—the set of priors. The special case where $\mathcal{P}$ is a singleton gives the AA version of SEU.

Ambiguity aversion, as defined in Equation 2.4, is the central assumption in axiomatization of the MP functional form by Gilboa & Schmeidler (1989). Another important axiom is certainty independence: For all AA acts h' and h, all constant acts c and $\alpha \in (0, 1)$,

$$h' \succ h \Leftrightarrow \alpha h' + (1-\alpha)c \succ \alpha h + (1-\alpha)c. \tag{2.7}$$

In other words, the invariance required by the independence axiom holds so long as mixing involves a constant act. This axiom ensures that Ellsberg-type choices are motivated by hedging. Essentially, moving from expected utility to MP amounts to replacing the independence axiom by uncertainty aversion and certainty independence.

Furthermore, comparative ambiguity aversion is simply characterized: Individual 1 is more ambiguity averse than individual 2 if and only if

$$u_1 = u_2 \text{ and } \mathcal{P}_1 \supset \mathcal{P}_2. \tag{2.8}$$

Thus the model affords a separation between risk attitudes, modeled exclusively by the vNM index u, and ambiguity attitudes, modeled in the comparative sense by the set of priors $\mathcal{P}$. Put another way, expanding $\mathcal{P}$ leaves risk attitudes unaffected and increases ambiguity aversion.

The MP model is very general because the set of priors can take many different forms. Consider briefly two examples that have received considerable attention and that offer scalar parametrizations of ambiguity aversion. Refer to ε-contamination if

$$\mathcal{P} = \{(1-\varepsilon)p^* + \varepsilon p : p \in \mathcal{P}'\}, \tag{2.9}$$

where $\mathcal{P}' \subset \Delta(\Omega)$ is a set of probability measures, $p^* \in \mathcal{P}'$ is a reference measure, and ε is a parameter in the unit interval. This equation is used heavily in robust statistics (see, e.g., Huber 1981). Epstein & Wang (1994) apply it to finance, whereas Kopylov (2009) provides axiomatic foundations. The larger ε is, the more weight is given to alternatives to p^* being relevant and the more ambiguity averse the individual is in the formal sense of Equation 2.5. An act is evaluated by a weighted average of its expected utility according to p^* and its worst-case expected utility:

$$U(h) = (1-\varepsilon)\int_{\Omega} u(h)dp^* + \varepsilon \min_{p \in \mathcal{P}'} \int_{\Omega} u(h)dp. \tag{2.10}$$

In the second example, $\mathcal{P}$ is an entropy-constrained ball. Fix a reference measure $p^* \in \Delta(\Omega)$. For any other $p \in \Delta(\Omega)$, its relative entropy is $R(p \parallel p^*) \in [0,\infty]$, where

$$R(p \parallel p^*) = \int_\Omega \left(\log \frac{dp}{dp^*}\right) dp, \tag{2.11}$$

if p is absolutely continuous with respect to p^*, and ∞ otherwise. Though not a metric, for example, it is not symmetric, $R(p \parallel p^*)$ is a measure of the distance between p and p^*; note that $R(p \parallel p^*) = 0$ if and only if $p = p^*$. Finally, define

$$\mathcal{P} = \{p : R(p \parallel p^*) \leq \eta\}. \tag{2.12}$$

The MP model is sometimes criticized on the grounds that it implies extreme aversion or paranoia. But that interpretation is based on the implicit assumption, not imposed by the model that $\mathcal{P}$ is the set of all logically possible priors. Gajdos et al. (2008) nicely clarify the difference between the subjective set of priors $\mathcal{P}$ and the set of logically possible probability laws. For example, in the Ellsberg example, it is perfectly consistent with the model that the individual use the probability interval $\left[\frac{1}{3}, \frac{2}{3}\right]$, even though any probability in the unit interval is consistent with the information given for the ambiguous urn. Ultimately, the only way to argue that the model is extreme is to demonstrate extreme behavioral implications of the axioms, which has not been done.

2.1.3. The "smooth ambiguity" model of preference. Klibanoff et al. (2005), henceforth KMM, propose the following utility function over AA acts:

$$U^{KMM}(h) = \int_{\Delta(\Omega)} \phi\left(\int_\Omega u(h(\omega))dp(\omega)\right) d\mu(p). \tag{2.13}$$

Here μ is a probability measure on $\Delta(\Omega)$, $u : \Delta(C) \to R$ is a vNM functional as before, and ϕ is continuous and strictly increasing on $u(C) \subset \mathbb{R}$. For simplicity, suppose that u is continuous and strictly increasing on C. Identify a KMM agent with a triple (u,ϕ,μ) satisfying the above conditions. By comparison, the MP functional form is a limiting case: If P is the support of μ, then, up to ordinal equivalence, Equation 2.6 is obtained in the limit as the degree of concavity of ϕ increases without bound.

This functional form admits an appealing interpretation. If the individual were certain of p being the true law, she would simply maximize expected utility using p. However, in general, there is uncertainty about the true law, or "model uncertainty," represented by the prior μ. This uncertainty about the true law matters if ϕ is nonlinear. In particular, if ϕ is concave, then the individual is ambiguity averse in the sense of Equation 2.4, and greater concavity implies greater ambiguity aversion in the sense of Equation 2.5. By contrast, ambiguity (as opposed to the attitude toward it) seems naturally to be captured by μ. Hence, a separation is claimed to be provided between ambiguity and aversion to ambiguity. This separation is highlighted by KMM as an advantage of their model.

Although the above interpretation is intriguing, its validity and, more broadly, the model's appeal depend on the plausibility of the model's axiomatic foundations. Epstein (2010) points out a problematic feature of the foundations provided by KMM, which can be summarized by referring to the Ellsberg paradox we use for intuition. [For other criticisms of the smooth ambiguity model, see Baillon et al. (2009) and Halevy & Ozdenoren (2008).] Given the urn with red and blue balls, and with no information given about their numbers, the smooth model can rationalize a strict preference for betting on the toss of an unbiased coin. However, it has a counterintuitive prediction when we enrich the setting slightly.

Suppose now that there are two urns, I and II, constructed as above, except that for urn II you are told something further about the composition—the specifics are not important except that the information treats red and blue symmetrically and it is not enough to determine the number of red and blue balls, so that you would still strictly prefer to bet on an unbiased coin. Now compare the urns—how much would you be willing to pay for a bet on drawing red from urn I as opposed to a bet on drawing red from urn II? Because there is no information about red versus blue in urn I, whereas some information is provided in urn II, one would expect the willingness to pay to be at least as large for the bet in urn II. However, KMM's axiomatic foundations lead to the prediction that, in many such instances, an individual would be willing to pay strictly more for the bet on drawing red from urn I (and similarly for bets on drawing blue)! The reader is referred to Epstein (2010) for details and also for an explanation of why the preceding calls into question interpretations put forth by KMM, including the noted separation between ambiguity and ambiguity attitude.

Seo (2009) provides alternative foundations for U^{KMM}. In his model, an individual can be ambiguity averse only if she fails to reduce objective (and timeless) two-stage lotteries to their one-stage equivalents. Thus the rational concern with model uncertainty and limited confidence in likelihoods is tied to the failure to multiply objective probabilities, a mistake that does not suggest rational behavior. Such a connection severely limits the scope of ambiguity aversion as modeled by Seo's approach.

Both MP and the smooth model satisfy ambiguity aversion (Equation 2.4) and thus can rationalize Ellsberg-type behavior. However, they represent distinct, indeed, "orthogonal" models of ambiguity aversion—the only point of intersection is SEU. One way to see this, and to highlight their differences, is (following Epstein 2010) to focus on what the models imply about the value of randomization. The MP model satisfies (because of certainty independence) the following: If $f \sim c$, then $\frac{1}{2}f + \frac{1}{2}c \sim f$. Thus mixing with a certainty equivalent is never valuable. By contrast, the smooth model satisfies the following (restricting attention to the special case where ϕ is strictly concave): If $f \sim c \sim \frac{1}{2}f + \frac{1}{2}c$, then for all acts h, $\frac{1}{2}f + \frac{1}{2}h \sim \frac{1}{2}c + \frac{1}{2}h$. In effect, if mixing with a certainty equivalent is not beneficial, then neither is mixing with any other act. To see why this is so, argue as follows, using the functional form of Equation 2.13 and strict concavity and monotonicity of ϕ: If $f \sim c \sim \frac{1}{2}f + \frac{1}{2}c$, then $\int_\Omega u(f)dp = u(c)$ with μ-probability 1, and the expected utility of f is certain in spite of model uncertainty. Thus

$$\begin{aligned}
U^{KMM}\left(\tfrac{1}{2}f + \tfrac{1}{2}h\right) &= \int_{\Delta(\Omega)} \phi\left(\int_{\Omega} \tfrac{1}{2}u(f) + \tfrac{1}{2}u(h)dp\right) d\mu(p) \\
&= \int_{\Delta(\Omega)} \phi\left(\int_{\Omega} \tfrac{1}{2}u(f) + \tfrac{1}{2}u(h)dp\right) d\mu(p) \\
&= \int_{\Delta(\Omega)} \phi\left(\tfrac{1}{2}u(c) + \tfrac{1}{2}\int_{\Omega} u(h)dp\right) d\mu(p) \\
&= U^{KMM}\left(\tfrac{1}{2}c + \tfrac{1}{2}h\right).
\end{aligned}$$

Accordingly, the two properties together imply the independence axiom and hence SEU.

To illustrate the effect of smoothness in applications it is helpful to abstract briefly from risk. Assume that the agent is risk neutral or, equivalently, restrict attention to acts that come with perfect insurance for risk. Formally, take u to be linear and rewrite the utilities as

$$U^{KMM}(h) = E^{\mu}[\phi(E^{p}[h])],$$
$$U^{MP}(h) = \min_{p\in\mathcal{P}} E^{p}[h].$$

For risk-neutral agents, ambiguity matters only if it affects means. Under the smooth model, ambiguity about means is reflected in a nondegenerate distribution of E^p $[h]$ under the prior μ. For a risk-neutral, ambiguity-averse KMM agent, an increase in ambiguity (in means) works like an increase in risk. Under the MP model, ambiguity about means is reflected in a nondegenerate interval for E^p $[h]$. For a risk-neutral MP agent, an increase in ambiguity (in means) can thus work like a change in the mean. The latter is a first-order effect.

2.1.4. Robust control, multiplier utility, and generalizations. Fix a reference measure $p^* \in \Delta(\Omega)$ and define relative entropy $R(p \parallel p^*) \in [0,\infty]$, for any other measure p, by Equation 2.11. Multiplier utility is defined by

$$U^{MU}(h) = \min_{p\in\Delta(\Omega)} \left[\int_{\Omega} u(h)dp + \theta R(p \parallel p^*)\right], \tag{2.14}$$

where $0 < \theta \leq \infty$ is a parameter.

This functional form was introduced into economics by Anderson et al. (2003), who were inspired by robust-control theory, and it was axiomatized by Strzalecki (2010). It suggests the following interpretation: Though p^* is the individual's "best guess" of the true probability law, she is concerned that the true law may differ from p^*. To accommodate this concern with model misspecification, when evaluating any given act h she takes all probability measures into account, weighing more heavily those that are close to her best guess as measured by relative entropy. Reliance on the (weighted) worst-case scenario reflects an aversion to model misspecification or ambiguity. In particular, multiplier utility is ambiguity averse in the sense of Equation 2.4, and ambiguity aversion increases with θ^{-1} in the sense of the comparative notion (Equation 2.5). At the extreme where $\theta = \infty$, the minimum is achieved at $p = p^*$, and $U^{MU}(\cdot) = \int_{\Omega} u(\cdot)dp^*$, reflecting complete confidence in the reference measure.

A key difference between multiplier utility and other models of ambiguity is that for choice among Savage acts—that is, acts that do not involve objective lotteries—it is observationally indistinguishable from SEU. Indeed, utility can be rewritten as

$$U^{MU}(h) = -\theta \log\left(\int_{\Omega} \exp\left(-\frac{1}{\theta}u(h)\right)dp^*\right) \tag{2.15}$$

(see Dupuis & Ellis 1997, proposition 1.4.2; Skiadas 2009b). Thus, on the domain of Savage acts h, for which outcomes are elements of C, U^{MU} conforms to SEU with prior p^* and vNM index

$$u_{\theta}(c) = \exp\left(-\frac{1}{\theta}u(c)\right), c \in C.$$

For Savage acts, introducing robustness ($\theta < \infty$) is thus indistinguishable from increasing risk aversion by moving from u to the more concave u_{θ}. Observational equivalence holds in the strong sense that even if one could observe the entire preference order over Savage acts, and not only a limited set of choices associated with more realistic sets of financial data, one could not distinguish the two models. This observational equivalence matters for applications because most empirically relevant objects of choice in financial markets are Savage acts—objective lotteries are rare. In many settings, multiplier utility may thus help reinterpret behavior that is also consistent with SEU, but it does not expand the range of behavior that can be rationalized. Reinterpretation can be valuable, for example, if there is an a priori bound on the degree of risk aversion. Of course, any exercise along these lines requires taking a stand on u or θ—from choice behavior alone, one can hope to identify, at most, the composite function $\exp\left(-\frac{1}{\theta}u(\cdot)\right)$. Thus, for example, Barillas et al. (2009) and Kleschelski & Vincent (2009) fix $u(c) = \log c$ and then arrive at estimates of the robustness parameter θ.

Multiplier utility has restrictive implications for choice in urn experiments. With one ambiguous urn, it can rationalize the intuitive choices in Ellsberg's experiment surrounding Equation 2.1: Take $p^* = \left(\frac{1}{2}, \frac{1}{2}\right)$ and $\theta < \infty$. However, consider an experiment with two ambiguous urns: In urn I you are told that $R + B = 100$ and $R, B \geq 40$, whereas in urn II you are told only that R and B sum to 100. Because there is more information about the composition of urn I, we would expect a preference to bet on red in urn I to red in urn II, and similarly for black. But this is impossible given multiplier utility. To see this, take the state space $S = \{R_I, B_I\} \times \{R_{II}, B_{II}\}$. The ranking of bets would be determined by how multiplier utility ranks Savage acts over S, but it conforms to subjective expected utility on the Savage domain. Thus, bets would have to be based on a probability measure p on S, which assigns higher probability to R_I than to R_{II}, and similarly for B_I and B_{II}, which is an impossibility.

There is a parallel with constant elasticity of substitution (CES) utility functions in consumer theory that is useful for perspective. The CES utility function is a flexible specification of cross-substitution effects between goods when there are only two goods, because then the elasticity is a free parameter. However, when there are more than two goods, it also imposes the a priori restriction that the noted elasticity is the same for all pairs of goods. Although CES utility remains a useful example, applications may call for more flexible functional forms (translog utility, for example). Analogously, multiplier utility can rationalize intuitive choice with one risky and one ambiguous urn. Once there are two or more ambiguous urns, it imposes additional a priori restrictions that need not be intuitive in applications.

Finally, briefly consider generalizations. Maccheroni et al. (2006a) introduce and axiomatize the following generalization, called variational utility:

$$U^{var}(h) = \min_{p \in \Delta(\Omega)} \left[\int_\Omega u(h)dp + c(p) \right], \tag{2.16}$$

where $c : \Delta(\Omega) \to [0,\infty]$ is a cost or penalty function. Multiplier utility is the special case where $c(p) = \theta R(p \parallel p^*)$. The above model is very general; it even encompasses MP utility, which corresponds to a cost function of the following form: For some set of priors $\mathcal{P} \subset \Delta(\Omega)$,

$$c(p) = \begin{cases} 0 & \text{if } p \in \mathcal{P} \\ \infty & \text{otherwise.} \end{cases}$$

Such a general model has no difficulty accommodating any number of ambiguous urns, and Maccheroni et al. (2006a) describe a number of interesting functional forms for c and hence utility. It remains to be seen whether they are useful in applications.

2.2. Dynamic or Sequential Choice Settings

Here we outline how the preceding models of preference can be extended to recursive, hence dynamically consistent, intertemporal models. Further extensions to accommodate learning are then discussed.

2.2.1. Recursive utility. The formal environment is now enriched as follows. In addition to the (finite) state space Ω, let $\mathcal{T} = \{0,1,\ldots,T\}$ be a time set and $\{\Sigma_t\}_{t=0}^{T}$ a filtration, where $\Sigma_0 = \{\emptyset, \Omega\}$ and $\Sigma_T = 2^{\Omega}$. Each Σ_t can be identified with a partition of Ω; $\Sigma_t(\omega)$ denotes the partition component containing ω. If ω is the true state, then at t the decision maker knows that $\Sigma_t(\omega)$ is true. One can also think of this information structure in terms of an event tree, with nodes corresponding to time-event pairs (t,ω).

For simplicity, assume consumption in any single period lies in the interval $C \subset \mathbb{R}^{+}$. We are interested primarily in C-valued consumption processes and how they are ranked. However, we again enlarge the domain in the AA way and consider the set of all $\Delta(C)$-valued processes. Each such process H is the dynamic counterpart of an AA act; it has the form $H = (H_t)$, where $H_t : \Omega \to \Delta(C)$ is Σ_t -measurable.

The new aspect of the dynamic setting is that choices can be made at all times. To model sequential choice, we assume a preference order at each node in the tree. Formally, let $\succeq_{t,\omega}$ be the preference prevailing at (t,ω), thought of as the ordering conditional on information prevailing then. The primitive is the collection of preferences $\{\succeq_{t,\omega}\} \equiv \{\succeq_{t,\omega} : (t,\omega) \in \mathcal{T} \times \Omega\}$. The corresponding collection of utility functions is $\{V_{t,\omega}\} \equiv \{V_{t,\omega} : (t,\omega) \in \mathcal{T} \times \Omega\}$. They are assumed to satisfy a recursive structure that we now describe [for more detailed formal presentations, see Epstein & Schneider (2003) for the MP-based model and Skiadas (2009a, ch. 6) for the general case, which relaxes the intertemporal additivity that we assume in Equation 2.18 below]. Define $B_t \equiv \left[1 + \beta + \ldots + \beta^{T-t}\right]$. In the infinite-horizon case, these discount terms simplify and each B_t is equal to $(1-\beta)^{-1}$.

To evaluate the act H from the perspective of node (t,ω), observe that it yields the current consumption (lottery) $H_t(\omega)$ and a random future payoff $V_{t+1,\cdot}(H)$. Here $\cdot$ in the subscript indicates that future utility is a function of $\omega' \in \Sigma_t(\omega)$, the realized node in the continuation of the tree from (t,ω). For each such node ω' (and only such nodes matter), let

$$V_{t+1,\omega'}(H) = B_{t+1} u\left(c^{H}_{t+1,\omega'}\right). \tag{2.17}$$

Thus, $c^{H}_{t+1,\omega'}$ is a certainty equivalent in the sense of being the (unique) level of consumption that, if received in every remaining period, would be indifferent from the perspective of $(t+1,\omega')$ to H. Because this certainty equivalent varies with the continuation ω', it defines a "static" act, of the sort discussed above, and whose utility can be computed using one of the static-ambiguity models discussed previously. Finally, the latter utility is aggregated with current felicity in the familiar discounted additive fashion to yield $V_{t,\omega}(H)$.

To be more precise, let * denote any of the models of ambiguity preference discussed above. Let $\{U^{*}_{t,\omega}\}$ be a collection of utility functions conforming to the model *, one for each node in the tree, having fixed risk preferences: $U^{*}_{t,\omega}(\cdot) = u(\cdot)$ on $\Delta(C)$ for every (t,ω). (Some obvious measurability restrictions are also assumed.) Refer to $\{U^{*}_{t,\omega}\}$ as a set of one-step-ahead utility functions. Say that preferences $\{\succeq_{t,\omega}\}$, or the corresponding utilities $\{V_{t,\omega}\}$, are recursive if there exist $u : \Delta(C) \to \mathbb{R}$ affine, a discount factor $0 < \beta < 1$, and a set $\{U^*_{t,\omega}\}$ of one-step-ahead utilities such that for all acts H: (*a*) $V_{T+1,\cdot}\ (H) = 0$ and (*b*) utilities $V_{t,\omega}\ (H)$ are evaluated by backward induction according to, for each (t,ω),

$$V_{t,\omega}(H) = u(H_t(\omega)) + \beta B_{t+1} U^*_{t,\omega}\left(c^H_{t+1,}\cdot\right). \tag{2.18}$$

The primitive components of the recursive model are $u(\cdot)$, modeling attitudes toward current consumption risks (and intertemporal substitutability[1]), a discount factor β, and the set $\{U^*_{t,\omega}\}$. Accordingly, $U^*_{t,\omega}$ represents preference, conditional on (t,ω), over the set of one-step-ahead acts, i.e., acts H for which $H_\tau(\cdot) = H_{t+1}(\cdot)$ for all $\tau > t$, that is, H produces a constant stream (of lotteries) for times $t+1, t+2, \ldots$ and, in particular, all ambiguity (though not risk) is resolved at $t+1$. Thus, $U^*_{t,\omega}$ models preferences over bets on the next step.

There are simple restrictions on preferences, specific to the dynamic setting, that are the main axioms characterizing recursive utility. First, preference at any node depends only on available information. Second, when evaluating H at any node, the individual cares only about what H prescribes in the continuation from that node—unrealized parts of the tree do not matter, an assumption that is commonly called consequentialism. Third, the ranking of risky prospects (lotteries) is the same at every node, i.e., a form of state independence. Finally, the collection of preferences is dynamically consistent; (contingent) plans chosen at any node remain optimal from the perspective of later nodes.

Next we discuss the recursive utility specifications corresponding to each of the static models discussed above. All previous comments remain relevant, as they relate to the ranking of one-step-ahead acts. We add comments that relate specifically to the dynamic setting. As shown from the connections drawn to the applied literature, the recursive model unifies a range of dynamic utility specifications that have been pursued in applications. It excludes specifications adopted in Hansen & Sargent (2007, 2009), Barillas et al. (2009), and in several other papers in the robust-control-inspired literature, which violate either consequentialism or dynamic consistency.

We refer also to continuous-time counterparts of the recursive models. In that case, the recursive construction of utility functions via Equations 2.17 and 2.18 is replaced by backward stochastic differential equations. Duffie & Epstein (1992) introduced these into utility theory in the risk context, and Chen & Epstein (2002) extended them to ambiguity aversion (modeled by MP). See Skiadas (2008) for a nice exposition, original formulations, and references to the technical literature on backward stochastic differential equations.

Recursive SEU. If one-step-ahead acts are evaluated by expected utility, then, from Equations 2.17 and 2.18,

$$V_{t,\omega}(H) = u(H_t(\omega)) + \beta \int_\Omega V_{t+1,\omega'}(H)dp_{t,\omega}(\omega'), \tag{2.19}$$

where $p_{t,\omega} \in \Delta(\Omega, \Sigma_{t+1})$ gives (t,ω) -conditional beliefs about the next step. This is the standard model.

[1]The confounding of risk aversion and substitution in u can be improved upon via a common generalization of Equation 2.18 and Epstein & Zin (1989). The resulting model can (partially) disentangle intertemporal substitution, risk aversion, and ambiguity aversion. Skiadas' (2009, ch. 6) treatment is general enough to admit such a three-way separation. Hayashi (2005) describes such a model where the ranking of one-step-ahead acts conforms to the MP model.

Recursive multiple priors. Let $P_{t,\omega} \subset \Delta(\Omega, \Sigma_{t+1})$ be the *set* of (t,ω)-conditional probability measures describing beliefs about the next step (events in Σ_{t+1}), and let $U^*_{t,\omega}(h) = \min_{p \in P_{t,\omega}} \int u(h)dp$ for any $h : (\Omega, \Sigma_{t+1}) \to \Delta(C)$. Then Equations 2.17 and 2.18 imply

$$V_{t,\omega}(H) = u(H_t(\omega)) + \beta \min_{p \in P_{t,\omega}} \int_\Omega V_{t+1,\omega'}(H)dp\left(\omega'\right). \tag{2.20}$$

Epstein & Wang (1994) first put forth this model; Epstein & Schneider (2003, 2007, 2008) axiomatize and apply it. The special case where each set $P_{t,\omega}$ has the entropy-constrained form in Equation 2.12 was suggested in Epstein & Schneider (2003) and has subsequently been applied by a number of papers in finance (described in Section 3 below). For a continuous-time formulation of recursive MP, see Chen & Epstein (2002).

Recursive smooth ambiguity model. Define $\phi \circ U^*_{t,\omega}$ by Equation 2.13, where μ, but not u or ϕ, varies with (t,ω). One obtains

$$V_{t,\omega}(H) = u(H_t(\omega)) + \beta B_{t+1}\phi^{-1}\left(\int_{\Delta(\Omega)} \phi\left(B^{-1}_{t+1}\int_\Omega V_{t+1,\omega'}(H)dp(\omega')\right)d\mu_{t,\omega}(p)\right). \tag{2.21}$$

This is closely related to the recursive version of the smooth ambiguity model described in Klibanoff et al. (2009) and the specifications in the applied papers by Chen et al. (2009) and Ju & Miao (2009).

Skiadas (2009b) shows that in Brownian and Poisson environments, the continuous-time limit of the recursive smooth ambiguity model is indistinguishable from the one in which the function ϕ is linear, that is, ambiguity aversion vanishes in the limit. He assumes that ϕ is invariant to the length of the time interval. Other ways of taking the continuous-time limit include, for example, allowing the concavity of ϕ to increase suitably as the interval shrinks. However, keeping ϕ fixed seems unavoidable if one sees ambiguity aversion as (separate from ambiguity and as) subject to calibration across settings.

(Recursive) multiplier utility and generalizations. Following Equation 2.15, define

$$\exp\left(-\frac{1}{\theta_{t,\omega}}U^*_{t,\omega}(h)\right) = \left(\int_\Omega \exp\left(-\frac{1}{\theta_{t,\omega}}u(h)\right)dp^*_{t,\omega}\right), \tag{2.22}$$

where $p^*_{t,\omega} \in \Delta(\Omega,\Sigma_{t+1})$ is the reference one-step-ahead measure. For simplicity, and because it is assumed universally, let $\theta_{t,\omega} = \theta$, which is a constant. Then Equations 2.17 and 2.18 imply

$$V_{t,\omega}(H) = u(H_t(\omega)) + \beta B_{t+1}\log\left[-\theta\left(\int_\Omega \exp\left(-\frac{1}{\theta}B^{-1}_{t+1}V_{t+1,\omega'}(H)\right)dp^*_{t,\omega}(\omega')\right)\right].$$

This is a special case of recursive utility as defined by Epstein & Zin (1989), where θ^{-1} parametrizes risk aversion separately from u, which also models intertemporal substitution. In continuous time, one obtains a special case of stochastic differential utility (Duffie & Epstein 1992).

To see the connection to robustness as proposed by Hansen & Sargent (2001), let $p^* \in \Delta(\Omega, \Sigma_T)$ be the reference measure corresponding to $\{p^*_{t,\omega}\}$ and p any other measure on Σ_T, and denote by p_t and p_t^* the restrictions of p and p^* to Σ_t. Define the time-averaged entropy by $\mathcal{R}(p \parallel p^*) = \Sigma_{t \geq 0} \beta^t \, E_p \left[log\left(\frac{dp_t}{dp_t^*}\right)\right]$, if p_t is absolutely continuous with respect to p_t^* for each t, and $\mathcal{R}(p \parallel p^*) = \infty$ otherwise. Then, [see Skiadas (2003) for a general proof for continuous-time], the recursive utility functions above can be written alternatively in the following form paralleling Equation 2.14:

$$V_0^\theta(H) = \min_{p \in \Delta(\Omega)} \left[\int_\Omega \left(\Sigma_{t=0}^T \beta^t u\left(H_\tau(\omega') \right) \right) dp(\omega') + \theta \mathrm{R}\, (p \mathrm{P} p^*) \right]. \tag{2.23}$$

Similar expressions obtain for the conditional utility functions $V_{t,\omega}^\theta(H)$. This reformulation parallels the equivalence of Equations 2.15 and 2.14 in the static context. This permits a reinterpretation of existing risk-based models, [such as the Barillas et al. (2009) reinterpretation of Tallarini (2000) in terms of robustness], but does not add new qualitative predictions.

To accommodate behavior toward several urns, of interest is to extend the model to allow "source dependence," that is, several driving processes and a concern for robustness that is greater for some processes than for others. However, this is hard to square with dynamic consistency and consequentialism. Indeed, let $\Omega = \Pi_{i=1}^n \Omega_i$ and think of n-driving processes. To capture source dependence, extend Equation 2.23 so that for each Ω_i, relative entropy measures distance between Ω_i -marginals with a separate multiplier θ_i for each i. However, unless θ_i are all identical, such a model is not recursive and thus precludes dynamic consistency.

This is in stark contrast to the recursive framework (Equations 2.17 and 2.18) that accommodates a wide range of ambiguity preferences, while having dynamic consistency built in. For example, Skiadas (2008) formulates recursive models that feature source dependence and that are special cases of our general framework (Equations 2.17 and 2.18). Maccheroni et al. (2006b) axiomatize a recursive version of variational utility that is the special case of our recursive model for which one-step-ahead acts are evaluated using variational utility (Equation 2.16).

Skiadas (2009b) derives continuous-time limits for a subclass of recursive variational utility containing the multiplier model (Equation 2.23). He shows that, in a Poisson environment (though not with Brownian uncertainty), these models, with the single exception of multiplier utility, are distinguishable from stochastic differential utility (providing another sense in which multiplier utility is an isolated case). Skiadas also suggests that some of these models have tractability advantages and are promising for pricing, particularly because of the differential pricing of Brownian and Poisson uncertainty.

2.2.2. Updating and learning. The one-step-ahead utility functions $\{U_{t,\omega}^*\}$ are primitives in the recursive model (Equations 2.17 and 2.18) and are unrestricted except for technical regularity conditions. Because they represent the individual's response to data, in the sense of describing his view of the next step as a function of history, they are the natural vehicle for modeling learning. Here, for each of the specific recursive models just described, we consider restrictions on $\{U_{t,\omega}^*\}$. Remaining within the recursive utility framework means the dynamic consistency is necessarily satisfied. The central issue is whether the specification adopted adequately captures intuitive properties of learning under ambiguity.

Learning is sometimes invoked to criticize models of ambiguity aversion. The argument is that, because ambiguity is due to a lack of information and is resolved as agents learn, it is at best a short-run phenomenon. Work on learning under ambiguity has shown that this criticism is misguided. First, ambiguity need not be due only to an initial lack of information. Instead, it may be generated by hard to interpret, ambiguous signals. Second, there are intuitive scenarios where ambiguity does not vanish in the long run. Note, however, the literature has not provided compelling axioms, beyond those underlying recursivity of Equations 2.17 and 2.18, to guide the modeling of learning under ambiguity. Thus, to assess various models, we rely on thought experiments, the following of which is based on Epstein & Schneider (2008).

A thought experiment. You are faced with two sequences of urns. One sequence consists of risky urns and the other of ambiguous urns. Each urn contains black and white balls. At every period, one ball is drawn from each of that period's urns and bets are offered on the draws from next period's urns. The sequence of risky urns is constructed (or perceived) as follows: First, a ball is placed in each urn according to the outcome of a fair coin toss. If the coin toss produces heads, a black "coin ball" is placed in each urn; if tails, then a white coin ball. In addition to a coin ball, each risky urn contains four noncoin balls, two of each color. The sequence of risky urns is thus an example of learning from i.i.d. signals. After a sufficient number of draws, you will become confident about the color of the coin ball from observing the frequency of black draws.

Each urn in the ambiguous sequence also contains a single coin ball with color determined as above (note the coin tosses for the two sequences are independent.) In addition, you are told that each urn contains either $n = 2$ *or* $n = 6$ noncoin balls of which exactly $\frac{n}{2}$ are black and $\frac{n}{2}$ are white. Finally, n varies "independently" across ambiguous urns. The ambiguous urns thus also share a common element (the coin ball), about which you can hope to learn, but they also have idiosyncratic elements (the noncoin balls) that are poorly understood and thus possibly unlearnable.

Ex ante, not knowing the outcome of the coin tosses, would you rather have a bet that pays 100 if black is drawn from the first risky urn (and zero otherwise), or a bet that pays 100 if black is drawn from the first ambiguous urn? The intuition indicated by Ellsberg suggests a strict preference for betting on the risky urn. In the risky urn, drawing black has an objective probability of $\frac{1}{2}$. For the ambiguous urn, the corresponding probability is either in $\left[\frac{4}{7}, \frac{2}{3}\right]$ or in $\left[\frac{1}{3}, \frac{3}{7}\right]$, each with probability $\frac{1}{2}$. Averaging endpoints yields the interval $\left[\frac{19}{42}, \frac{23}{42}\right]$, which has $\frac{1}{2}$ as its midpoint. Thus, ambiguity aversion suggests the preference for the precise $\frac{1}{2}$. However, the unambiguous nature of the bet on the risky urn can be offset by reducing the winning stake there. Let $z < 100$ be such that you are indifferent between a bet that pays z if black is drawn from the risky urn and a bet that pays 100 if black is drawn from the ambiguous urn.

Now sample by drawing one ball from the first urn in each sequence. Suppose that the outcome is black in both cases. With this information, consider versions of the above bets based on the second-period urns. Would you rather have a bet that pays z if black is drawn from the second risky urn or a bet that pays 100 if black is drawn from the second ambiguous urn? Our intuition is that, even with this difference in stakes, betting on the risky urn would be strictly preferable. The reason is that inference about the coin ball is clear for the risky urn—the posterior probability of a black coin ball is $\frac{3}{5}$. As a result, the predictive probability of drawing black is $\frac{3}{5}\left(\frac{3}{5}\right) + \frac{2}{5}\left(\frac{2}{5}\right) = \frac{13}{25}$. In contrast, for the ambiguous urn the signal (a black draw) is harder to interpret, leaving us less confident in our assessment of the composition of that urn. We now elaborate on this point.

Just as for the risky sequence, the only useful inference for the ambiguous sequence is about the coin ball (because noncoin balls are thought to be unrelated across urns in the sequence). But what does a black draw tell us about the coin ball? On the one hand, it could be a strong signal of the color of the coin ball (if $n = 2$ in the sampled urn) and hence also of a black draw from the second urn. On the other hand, it could be a weak indicator (if $n = 6$ in the sampled urn). The posterior probability of the coin ball being black could be anywhere between $\frac{6/2+1}{6+1} = \frac{4}{7}$ and $\frac{2/2+1}{2+1} = \frac{2}{3}$, with a range of predictive probabilities for black ensuing.

The difference in winning stakes, z versus 100, compensates for prior ambiguity, but not for the difficulty in interpreting the realized signal. Thus a preference for betting on the risky urn is to be expected, even given the difference in winning prizes. By analogous reasoning, similar rankings for bets on white are intuitive, both ex ante and ex post conditional on having drawn black balls. Indeed, the lower quality of the signal from the ambiguous urn makes it harder to judge any bet, not just a bet on black.

A multiple-priors model of learning under ambiguity. Epstein & Schneider (2008) propose a model of learning within the recursive MP framework (Equation 2.20) that accommodates the intuitive choices in the thought experiment. It is motivated by the following interpretation of the experiment. The preference to bet on the risky urn ex post is intuitive because the ambiguous signal—the draw from the ambiguous urn—appears to be of lower quality than the noisy signal—the draw from the risky urn. A perception of low information quality arises because the distribution of the ambiguous signal is not objectively given. As a result, the standard Bayesian measure of information quality, i.e., precision, seems insufficient to compare the two signals adequately. The precision of the ambiguous signal is parametrized by the number of noncoin balls n: When there are few noncoin balls that add noise, precision is high. A single number for precision cannot rationalize the intuitive choices because behavior is as if one is using different precisions depending on the bet that is evaluated. When betting on a black draw, the choice between urns is made as if the ambiguous signal is less precise than the noisy one, so that the available evidence of a black draw is a weaker indicator of a black coin ball. In other words, when the new evidence—the drawn black ball—is "good news" for the bet to be evaluated, the signal is viewed as relatively imprecise. In contrast, in the case of bets on white, the choice is made as if the ambiguous signal is more precise than the noisy one, so that the black draw is a stronger indicator of a black coin ball. Now the new evidence is "bad news" for the bet to be evaluated and is viewed as relatively precise. The intuitive choices can thus be traced to an asymmetric response to ambiguous news.

The implied notion of information quality can be captured by combining worst-case evaluation with the description of an ambiguous signal via multiple likelihoods. To see how, think of the decision maker as trying to learn the colors of the two coin balls, believing that is all he needs to learn for the risky sequence. By contrast, for the ambiguous sequence, his perception of noncoin balls as varying independently across urns means there is nothing to be learned from past observations about that component of future urns. For both sequences, his prior over these parameters places probability $\frac{1}{2}$ on the coin ball being black. More generally, the model admits MP over parameters. The intuition given above for the choices indicated in the experiment suggests clearly a translation in terms of multiple likelihoods. Signals for the risky sequence have objective distributions conditional on the color of the coin ball and can be modeled in the usual way by single likelihoods. However, for the ambiguous sequence, the distribution of the signal is unknown, even conditioning on the color of the coin ball, because it varies with n, suggesting multiple likelihoods.

Other models of learning. How do other models perform with respect to the thought experiment? SEU is ruled out by the ex ante ambiguity-averse ranking (the situation is ultimately analogous to Ellsberg's original experiment). The same applies to multiplier utility because it coincides with SEU on Savage acts. Recursive variational utility (Maccheroni et al. 2006b) inherits the generality of variational utility. In particular, it generalizes recursive MP and so can accommodate the thought experiment. The question is whether the added generality that it affords is useful in a learning context. A difficulty is that it is far from clear how to model updating of the cost or penalty function $c(\cdot)$.

The situation is more complicated for the smooth ambiguity model. It can accommodate the ex ante ambiguity-averse choices. To consider also the ex post rankings indicated, it is necessary to specify updating for the recursive smooth model (Equation 2.21). We assume that beliefs $\mu_{t,\omega}$ about the true law are updated by Bayes' rule. Then the recursive smooth model cannot accommodate the intuitive behavior in the thought experiment, at least given natural specifications of the model (outlined next).

Consider the functional form for utility (Equation 2.13). For the risky urns, all relevant probabilities are given; thus, bets on the risky urns amount to lotteries, which are ranked according to u. To model choice between bets on the ambiguous urns, we must first specify the state space Ω. Take $\Omega = \{B, W\}$ so that a state specifies the color of the ball on any single draw.[2] Then a bet on B corresponds to the act f_B, with $f_B(B) = 100$ and $f_B(W) = 0$. The smooth model specifies prior beliefs μ about the true probability of drawing B. Here the latter is determined by the color of the coin ball $\theta = B$ or W and by the number $n = 2$ or 6 of the noncoin balls, according to

$$\ell(B|\theta,n) = \begin{cases} \frac{2}{3} & \theta = B, n = 2 \\ \frac{1}{3} & \theta = W, n = 2 \\ \frac{4}{7} & \theta = B, n = 6 \\ \frac{3}{7} & \theta = W, n = 6. \end{cases}$$

Thus, view μ as a probability measure on pairs (θ,n). Let μ be uniform over the above four possibilities and suppose that ϕ is strictly concave (as in all applications of the modeldiscussed). Then it is a matter of elementary algebra to show that the choices described in the thought experiment cannot be accommodated.

A final comment concerns a theme we have emphasized throughout the discussion of preference models: Appearances can be misleading. The only way to understand a model is through its predictions for behavior, whether through formal axioms or thought experiments. What could be more natural than to use Bayes' rule to update the prior as in the recursive smooth model? Foregoing the issue of how to update sets of priors as in Epstein & Schneider (2007. 2008), one can import results from Bayesian learning theory. The models in Hansen (2007), Chen et al. (2009), and Ju & Miao (2009) share this simplicity. In all cases, updating proceeds exactly as in a Bayesian model, and ambiguity aversion

[2]An alternative is to take the state space to be {2, 6}, corresponding to the possible number of the noncoin balls. However, with this state space, even the (ambiguity-averse) ex ante choices cannot be rationalized.

enters only in the way that posterior beliefs are used to define preference. However, the thought experiment illustrates what is being assumed by adopting such an updating rule: Concern with "signal or information quality" is excluded.

3. AMBIGUITY IN FINANCIAL MARKETS

This section illustrates the role of ambiguity in portfolio choice and asset pricing. We consider simple two- and three-period setups. Those setups are sufficient to illustrate many of the effects that drive more elaborate (and now increasingly quantitative) models studied in the literature. We also focus on the MP model. This is because the range of new effects—relative to models of risk—is arguably larger for that model. Specific differences between MP and smooth models are also highlighted.

3.1. Portfolio Choice

Begin with a two-period problem of savings and portfolio choice. An agent is endowed with wealth W_1 at date 1 and cares for consumption at dates 1 and 2. There is an asset that pays the interest rate r^f with certainty as well as n uncertain assets with log returns collected in a vector r. The returns r could be ambiguous. Let $\mathcal{P}_1$ denote a set of beliefs held at date 1 about returns at date 2. The agent chooses consumption at both dates and a vector of portfolio shares θ for the n uncertain assets to solve

$$\begin{aligned}\max_{C_1,\theta}\ \min_{p\in\mathcal{P}_1}&\{(1-\beta)u(C_1)+\beta E^p[u(C_2)]\}\\ \text{s.t.}\, C_2 &= (W_1-C_1)R_2^w\\ R_2^w &= \left(\exp\left(r^f\right)+\sum_{i=1}^{n}\theta_i\exp(r_i)\right),\end{aligned}$$

where R_2^w is the return on wealth realized at date 2.

Now restrict attention to log utility and lognormally distributed returns. With $u(c) = \log c$, the savings and portfolio-choice problems separate. In particular, the agent always saves a constant fraction $\beta/(1+\beta)$ of wealth, and he chooses his portfolio to maximize the expected log return on wealth. With lognormal returns, a belief in $\mathcal{P}_1$ can be represented by a vector μ^r of expected (log) returns as well as a covariance matrix Σ. Throughout, we use an approximation for the log return on wealth introduced by Campbell & Viceira (1997)

$$\log R_2^w \approx r^f + \theta'\left(r+\tfrac{1}{2}\, diag\,\Sigma - r^f\iota\right) - \frac{1}{2}\theta'\Sigma\theta, \tag{3.1}$$

where $diag\,\Sigma$ is a vector containing the main diagonal of Σ and ι is an n-vector of ones. In continuous time, the formula is exact by Ito's Lemma; in discrete time, it yields simple solutions that illustrate the key effects.

It is convenient to work with excess returns. Define a vector of premia (expected log excess returns, adjusted for Jensen's inequality) by

$$\mu^e = \mu^r + \tfrac{1}{2}\, diag\,\Sigma - r^f\iota.$$

Let Π_1 denote the set of parameters (μ^e, Σ) that correspond to beliefs in $\mathcal{P}_1$. This set can be specified to capture ambiguity about different aspects of the environment. In general, the size of Π_1 reflects the agents' lack of confidence when thinking about returns. For example, worse information about an asset may lead an agent to have a wider interval of possible mean log returns for that asset. In a dynamic setting, the size of the sets Π_1 and $\mathcal{P}_1$ will change over time with new information. Below we discuss the effects of such updating by doing comparative statics with respect features of Π_1.

Using the approximation Equation 3.1, the portfolio-choice problem becomes

$$\max_{\theta} \min_{p \in \mathrm{P}_1} E^p \left[\log R_2^w \right] \approx \max_{\theta} \min_{(\mu^e, \Sigma) \in \Pi_1} \left\{ r^f + \theta' \mu^e - \tfrac{1}{2} \theta' \Sigma \theta \right\}. \tag{3.2}$$

If there is no ambiguity—that is (μ^e, Σ) is known and is therefore the only element of Π_1—then we have a standard mean variance problem, with optimal solution $\theta = \Sigma^{-1} \mu^e$. More generally, the agent evaluates each candidate portfolio under the worst-case return distribution for that portfolio.

3.1.1. One ambiguous asset: nonparticipation and portfolio inertia at certainty. Assume that there is only one uncertain asset. Its log excess return has known variance σ^2 and an ambiguous mean that lies in the interval $[\bar{\mu}^e - \bar{x}, \bar{\mu}^e + \bar{x}]$. Think of $\bar{\mu}^e$ as a benchmark estimate of the premium; $\bar{x}$ then measures the agent's lack of confidence in that estimate. The agent solves

$$\max_{\theta} \min_{\mu^e \in [\bar{\mu}^e - \bar{x}, \bar{\mu}^e + \bar{x}]} \left\{ r^f + \theta \mu^e - \tfrac{1}{2} \sigma^2 \theta^2 \right\}. \tag{3.3}$$

Minimization selects the worst-case scenario depending on the agent's position: $\mu^e = \bar{\mu}^e - \bar{x}$ if $\theta > 0$ and $\mu^e = \bar{\mu}^e + \bar{x}$ if $\theta < 0$. Intuitively, if the agent contemplates going long in the asset, he fears a low excess return, whereas if he contemplates going short, then he fears a high excess return. If $\theta = 0$, the portfolio is not ambiguous and any μ^e in the interval solves the minimization problem.

The optimal portfolio decision anticipates the relevant worst-case scenario. For a given range of premia, the agent evaluates the best nonnegative position as well as the best nonpositive position, and then chooses the better of the two. This leads to three cases. First, if the premium is known to be positive ($\bar{\mu}^e - \bar{x} > 0$), then it is optimal to go long. In this case, any long position is evaluated using the lowest premium and the optimal weight is $\theta = (\bar{\mu}^e - \bar{x})/\sigma^2 > 0$. Similarly, if the premium is known to be negative ($\bar{\mu}^e + \bar{x} < 0$), then the optimal portfolio sells the asset short: $\theta = (\bar{\mu}^e + \bar{x})/\sigma^2 < 0$. Finally, if $\bar{\mu}^e + \bar{x} > 0 > \bar{\mu}^e - \bar{x}$, then it is optimal not to participate in the market ($\theta = 0$). This is because any long position is evaluated using the lowest premium, which is now negative, and any short position is evaluated using the highest premium, which is positive. In both cases, the return on wealth is strictly lower than the riskless rate, so it is better to stay out of the market.

Under ambiguity, nonparticipation in markets is thus optimal for many parameter values. In particular, for any benchmark premium $\bar{\mu}^e$, a sufficiently large increase in uncertainty will lead agents to withdraw from an asset market altogether. This is not true if all uncertainty is risk. Indeed, the participation decision does not depend on the quadratic risk term in Equation 3.3. That term becomes second order as θ goes to zero, that is, agents are "locally risk neutral" at $\theta = 0$. In the absence of ambiguity ($\bar{x} = 0$), agents participate, except in the knife-edge case $\bar{\mu}^e = 0$. Moreover, an increase in the variance σ^2 does not make agents withdraw from the market; all it makes them do is choose smaller positions.

Ambiguity-averse agents exhibit portfolio inertia at $\theta = 0$. Indeed, consider the response to a small change in the benchmark premium $\bar{\mu}^e$. For $\bar{\mu}^e < |\bar{x}|$, an ambiguity-averse agent will not change his position away from zero. This is again in sharp contrast to the risk case, where the derivative of the optimal position with respect to $\bar{\mu}^e$ is everywhere strictly positive. The key point is that an increase in ambiguity can be locally large relative to an increase in risk. Indeed, the portfolio $\theta = 0$ is both riskless and unambiguous. Any move away from it makes the agent bear both risk and ambiguity. However, an increase in ambiguity about means is perceived like a change in the mean and not like an increase in the variance. Ambiguity can thus have a first-order effect on portfolio choice that overwhelms the first-order effect of change in the mean, whereas the effect of risk is second order.

3.1.2. Hedging and portfolio inertia away from certainty. Nonparticipation and portfolio inertia can also arise when the portfolio $\theta = 0$ does not have a certain return, and when the ambiguous asset can help hedge risk. The MP model is sometimes claimed to give rise to inertia only at certainty. The claim is often based on examples with two states of the world, where MP preferences exhibit indifference curves that are kinked at certainty and smooth elsewhere. However, the example here illustrates that in richer settings inertia is a more general phenomenon. To see this, assume that the interest rate is not riskless but instead random with known mean μ_f, variance σ_f^2, and $cov(r^f, r) = \sigma_{fr} < 0$. One interpretation is that r^f is the real return on a nominal bond and r the return on the stock market, which is perceived to be an inflation hedge (stocks pay off more when inflation lowers the real bond return). The agent solves

$$\max_{\theta} \min_{p \in \mathcal{P}_1} E^p \left[\log R_2^w \right] \approx \max_{\theta} \min_{\mu^e \in [\bar{\mu}^e - \bar{x}, \bar{\mu}^e + \bar{x}]} \left\{ \mu^f + \theta\left(\mu^e - \sigma_{fr}\right) - \frac{1}{2}\left(\theta^2 \sigma^2 + \sigma_r^2 \right\}. \right.$$

Investing in stocks is now useful not only to exploit the equity premium μ^e, but also to hedge the risk in a bond position. Moreover, the portfolio $\theta = 0$ (holding all wealth in bonds) is still unambiguous, but it is no longer riskless. Adapting the earlier argument, the agent goes long in stocks if $\bar{\mu}^e - \bar{x} - \sigma_{fr} > 0$, he goes short if $\bar{\mu}^e + \bar{x} - \sigma_{fr} < 0$, and he stays out of the stock market otherwise. For a positive benchmark equity premium $\bar{\mu}^e > 0$, the degree of ambiguity (measured by $\bar{x}$) required to generate nonparticipation is now larger (because of the benefit of hedging), but the basic features of nonparticipation and portfolio inertia remain. The key point is that investing in stocks exposes investors to a source of ambiguity—the unknown equity premium—whereas investing in bonds does not.

Portfolio inertia is a property that is distinct from, and more general than, nonparticipation. This is because even away from certainty there can be portfolios where a small change in a position entails a large change in the worst-case belief. To illustrate, consider an agent who believes in a one-dimensional set of models of excess returns indexed by an ambiguous parameter $x \in [0, \bar{x}]$. In particular, the premium is $\mu^e = \bar{\mu}^e + x$ and the variance is $\sigma^2 = \bar{\sigma}^2 + 2x/\gamma$, where γ is known. Intuitively, the agent believes that risk and expected return go together, but he does not know the precise pair (μ^e, σ^2). Illeditsch (2010) shows that such a family of models can obtain when agents receive bad news of ambiguous precision: More precise bad news lowers both the conditional mean and the conditional variance of returns. The agent solves

$$\max_{\theta} \min_{p \in \mathcal{P}_1} E^p \left[\log R_2^w \right] \approx \max_{\theta} \min_{x \in [0, \bar{x}]} \left\{ \mu^f + \theta(\bar{\mu}^e + x) - \tfrac{1}{2}\theta^2\left(\bar{\sigma}^2 + 2x/\gamma\right) \right\}.$$

There are now two portfolios that are completely unambiguous: $\theta = 0$ and $\theta = \gamma$, and the latter yields the higher return on wealth if $\bar{\mu}^e > \gamma\bar{\sigma}^2/2$. If, moreover, ambiguity is large enough so that $\bar{\mu}^e < \gamma\bar{\sigma}^2 + \bar{x}$, then it is optimal to choose $\theta = \gamma$.

At $\theta = \gamma$, a small increase in θ leads to the worst-case scenario $x = \bar{x}$, whereas a small decrease leads to $x = 0$. Intuitively, risk is taken more seriously relative to expected return at higher positions. Accordingly, the worst-case scenario changes with position size: At high positions, agents fear high risk, whereas at small positions, they fear low expected returns. At $\theta = \gamma$, the two effects offset. The presence of an unambiguous portfolio is a knife-edge case driven by the functional form (or here by the approximation we are using). More generally, even if no portfolio makes the objective function independent of x, there can exist portfolios at which the minimizing choice of x flips discontinuously. Thus, it follows that, at $\theta = \gamma$, any news that slightly changes the benchmark premium $\bar{\mu}^e$ has no effect on portfolio choice. Indeed, changing the portfolio to exploit news about $\bar{\mu}^e$ would require the agent to bear ambiguity. The resulting first-order loss from increased uncertainty overwhelms the gain from a small change in $\bar{\mu}^e$.

3.1.3. Multiple ambiguous assets: selective participation and benefits from diversification.

With multiple assets, ambiguity gives rise to selective participation. To illustrate, consider a set of n uncertain assets such that (a) returns are known to be uncorrelated, i.e., the covariance matrix Σ in Equation 3.2 is diagonal, and (b) the premia μ^e are perceived to be ambiguous but independent, i.e., μ^e lies in the Cartesian product of intervals $\left[\bar{\mu}_i^e - \bar{x}_i,\, \bar{\mu}_i^e + \bar{x}_i\right]$, $i = 1, \ldots, n$. From Equation 3.2, it is optimal to participate in the market for asset i if and only if $0 \notin \left[\bar{\mu}_i^e - \bar{x}_i,\, \bar{\mu}_i^e + \bar{x}_i\right]$, that is, if the premium on asset i is nonzero and not too ambiguous. Agents thus stay away from those markets for which they lack confidence in assessing the distribution of returns. Introducing correlation among returns will change the conditions for participation, but it will not rule out selective participation. The argument is essentially the same as in the previous section.

If ambiguity about premia is independent across assets, then it cannot be diversified away. To see this, specialize further to i.i.d. risk ($\Sigma = \sigma^2 I$) as well as i.i.d. ambiguity about premia. In particular, let all premia lie in the same interval that is centered at $\bar{\mu}_i^e = \hat{\mu}^e$ and has bounds implied by $\bar{x}_i = \hat{x}$. Assume also that it is worthwhile to go long in all markets, or $\hat{\mu}^e - \hat{x} > 0$. Symmetry implies that the optimal portfolio invests the same share, say $\hat{\theta}/n$, to each uncertain asset. Substituting $\theta_i = \hat{\theta}/n$ for all i as well as $\mu^e = (\hat{\mu} - \hat{x})\iota$ in Equation 3.2, the return on wealth is

$$\max_{\theta} \min_{p \in \mathcal{P}_1} E^p \left[\log R_2^w\right] \approx \max_{\hat{\theta}} \left\{ r^f + \hat{\theta}(\hat{\mu}^e - \hat{x}) - \frac{\sigma^2}{2n} \hat{\theta}^2 \right\}.$$

As the number of independent uncertain assets becomes large, the quadratic term becomes small and the effect of risk on the portfolio decreases. At the same time, the effect of ambiguity on portfolio choice remains unchanged. Intuitively, ambiguity reflects confidence in prior information about individual assets that is perceived as a reduction in the mean. Investing in many assets does not raise confidence in that prior information.

Without independence, diversification may be beneficial, because assets hedge ambiguity in other assets. For example, retain the assumption of i.i.d. risk, but suppose now the agent believes premia are $\mu^e = \hat{\mu}^e \iota + x$ with $x' x \leq \kappa^2$, where $\kappa < \hat{\mu}^e$ is fixed. Intuitively, the agent perceives a common factor in mean returns such that if one mean is very far away from the benchmark $\hat{\mu}^e$, then all others must be relatively close. The agent solves

$$\max_{\theta} \min_{p \in \mathcal{P}_1} E^p \left[\log R_2^w\right] \approx \max_{\theta} \min_{x'x \leq \kappa^2} \left\{ r^f + \theta(\hat{\mu}^e \iota + x) - \frac{\sigma^2}{2n} \theta^2 \right\}.$$

Symmetry again implies $\theta = \iota\hat{\theta}/n$ for some $\hat{\theta}$. For $\hat{\theta} > 0$, minimization yields $x = \theta\sqrt{2\kappa/\theta'\theta} = \kappa/\sqrt{n}$ and the portfolio return is thus

$$\max_{\hat{\theta}} \left\{ r^f + \hat{\theta}\left(\hat{\mu}^e - \frac{\kappa}{\sqrt{n}}\right) - \frac{\sigma^2}{2n}\hat{\theta}^2 \right\}.$$

The effect of ambiguity on portfolio choice thus shrinks as n increases, although the speed is slower than for risk.

An extreme case of cross-hedging ambiguity arises when an unambiguous family of portfolios can be constructed. Suppose, for example, that there are only two assets with i.i.d. risk and that $\mu_1^e = \hat{\mu}^e + x$ and $\mu_2^e = \hat{\mu}^e - x$, with $x^2 \leq \kappa^2$. Such a situation may arise when there is pool of assets (e.g., mortgages) with relatively transparent payoff, which has been cut into tranches in a way that makes the payoffs on the individual tranches rather opaque. In this case, holding the entire pool or holding tranches in equal proportions hedges ambiguity. In contrast, an agent holding an individual tranche in isolation bears ambiguity.

3.1.4. Dynamics: entry and exit rules and intertemporal hedging. To illustrate new effects that emerge in an intertemporal context, consider a three-period setup with one uncertain asset. Beliefs can be described by sets of one-step-ahead conditionals. The date 1 one-step-ahead conditionals for date 2 log excess return are normal with variance σ_2^2 and ambiguous mean in the interval $\left[\bar{\mu}_2^e - \bar{x}_2,\, \bar{\mu}_2^e + \bar{x}_2\right]$. As of date 2, the date 3 log excess returns are again viewed as normal, now with variance σ_3^2. Moreover, there is a signal s_2 that induces, via some updating rule, an interval of expected log excess returns $\left[\bar{\mu}_3^e(s_2) - \bar{x}_3(s_2),\, \bar{\mu}_3^e(s_2) + \bar{x}_3(s_2)\right]$. In general, the signal can be correlated with the realized excess return r_2^e. This will be true, for example, if the agent is learning about the true premium, and the realized excess return is itself a signal. Importantly, updating will typically affect both the benchmark mean return $\bar{\mu}_3^e$ and the agents' confidence, as measured by $\bar{x}_3$.

Portfolio choice at date 2 works just like in the one period problem (Equation 3.3) above. The value function from that problem depends on wealth W_2 and the signal s_2. Up to a constant, it takes the form $V_2(W_2,s_2) = \log W_2 + h(s_2)$, where

$$h(s_2) = \frac{1}{2\sigma_3^2}\left(\max\{\bar{\mu}_3^e(s_2) - \bar{x}(s_2),0\}^2 + \min\{\bar{\mu}_3^e(s_2) + \bar{x}(s_2),0\}^2\right).$$

The value function is higher for signals that move the range of equity premia away from zero, so that worst-case expected returns higher than the riskless rate can be obtained. For example, Epstein & Schneider (2007) show in a model of learning about the premium with $s_2 = r_2^e$ that the value function is U-shaped in the signal.

Because the value function V_2 is separable in W_2 and s_2, the portfolio choice problem at date 0 can still be solved separately from the savings problem. The agent solves

$$\max_{\theta} \min_{p \in \mathcal{P}_1} \{E^p[\log R_2^w + h(s_2)]\}.$$

The difference to the one-shot problem (Equation 3.3) is that minimization takes into account the effect on the expected return at the optimal portfolio to be chosen at date 2, captured by h. As a result, it is possible that the choice of p and the choice of the optimal portfolio are different in the two-period problem from those in the one-period problem. In other words, an investor with a two-period horizon does not behave myopically but chooses to hedge future investment opportunities. This hedging is due entirely to ambiguity: With a single prior, log utility guarantees that myopic behavior is optimal. In the expected-utility case, hedging demand is linked to a nonzero cross derivative of the value function V_2. With ambiguity, hedging demand can arise in the log case, even though the cross derivative is zero, because the minimization step creates a link across periods between $E^p\left[\log R_2^w\right]$ and $E^p[h(s_2)]$.

In the intertemporal context, the (recursive) MP model delivers two new effects for portfolio choice. First, the optimal policy involves dynamic exit and entry rules. Indeed, updating shifts the interval of equity premia, and such shifts can make agents move in and out of the market. Second, there is a new source of hedging demand. It emerges if return realizations provide news that shift the interval of equity premia. Portfolio choice optimally takes into account the effects of news on future confidence. The direction of hedging depends on how news affects confidence. For example, Epstein & Schneider (2007) show that learning about premia gives rise to contrarian hedging demand if the empirical mean equity premium is low. Intuitively, agents with a low empirical estimate know that a further low return realization may push them toward nonparticipation and hence a low return on wealth (formally this is captured by a U-shaped value function). To insure against this outcome, they short the asset.

3.1.5. Differences between models of ambiguity. This section illustrates several phenomena that can be traced to first-order effects of uncertainty under the MP model, in particular selective participation, portfolio inertia, and the inability to diversify uncertainty (at least for some sets of beliefs). These effects cannot arise under SEU, which implies local risk neutrality at certainty, smooth dependence of portfolios on the return distribution (at least under the standard assumptions studied here), and benefits of diversification.

The smooth model and multiplier utility resemble SEU in the sense that they also cannot generate the above phenomena. This is immediate for multiplier utility, which is observationally equivalent to SEU on Savage acts, as explained in Section 2.1.4. Moreover, for the smooth model, if u and ϕ are suitably differentiable, then so is U^{KMM}. As a result, selective participation is again a knife-edge property. A theme that is common to smooth models and the MP model is the emergence of hedging demand due to ambiguity.

Some authors have argued that smoothness is important for tractability of portfolio problems. It is true that smoothness permits calculus techniques to be used to solve optimization problems. Moreover, in the expected utility case, closed-form solutions are sometimes available in dynamic problems, and the same may be true for smooth models that are close to expected utility. However, most applied portfolio choice problems considered in the literature today are solved numerically. Even in the expected utility case, they often involve frictions that make closed-form solutions impossible. From a numerical perspective, the additional one-step-ahead minimization step does not appear excessively costly.

3.1.6. Discipline in quantitative applications. In the portfolio choice examples above as well as in those on asset pricing below, the size of the belief set is critical for the magnitude of the new effects. There are two approaches in the literature to discipline the belief set. Anderson et al. (2000) propose the use of detection error probability (for an exposition, see also Barillas et al. 2009). These authors use detection error probabilities in the context of multiplier preference, but the idea has come to be used also to constrain the belief set in MP. The basic idea is to permit only beliefs that are statistically close to some reference belief, in the sense that they are difficult to distinguish from the reference belief based on historical data.

To illustrate, let A denote a reference belief (for example, a return distribution), and let B denote some other belief. We want to describe a sense in which A and B are "statistically close." Let π_A denote the probability, under A, that a likelihood ratio test based on the historical data (of returns, say) would falsely reject A and accept B. Define π_B similarly as the probability under B of falsely rejecting B in favor of A. Finally, define the detection error probability π by $\pi = \frac{1}{2}(\pi_A + \pi_B)$. The set of beliefs is now constrained to include only beliefs with π small enough. One may also choose to make additional functional form assumptions (for example, serial independence of returns).

A second approach to impose discipline involves using a model of learning. For example, the learning model of Epstein & Schneider (2007) allows the modeler to start with a large set of priors in a learning model, resembling a diffuse prior in Bayesian learning, and then shrinking the set of beliefs via updating. A difference between the learning and detection probability approach is that in the former the modeler does not have to assign special status to a reference model. This is helpful in applications where learning agents start with little information, for example, because of a recent structural change. In contrast, the detection probability approach works well for situations where learning has ceased or slowed down, and yet the true model remains unknown.

3.1.7. Literature notes. The nonparticipation result with one uncertain asset is due to Dow & Werlang (1992). More general forms of portfolio inertia appear in Epstein & Wang (1994) and Illeditsch (2010). Mukerji & Tallon (2003) compare portfolio inertia under ambiguity and first-order risk aversion. Garlappi et al. (2007) characterize portfolio choice with multiple ambiguous assets. Bossaerts et al. (2010) and Ahn et al. (2009) provide experimental evidence that supports first-order effects of uncertainty in portfolio choice.

A large empirical literature shows that investors prefer assets that are familiar to them and that the extensive margin matters. One candidate explanation for nonparticipation is that expected-utility investors pay a per-period fixed cost. Vissing-Jorgenson (2002) argues that this approach cannot explain the lack of stock market participation among the wealthy in the United States. Quantitative studies of familiarity bias using the MP model thus seem a promising avenue for future research. Cao et al. (2007) summarize the evidence and discuss ambiguity aversion as a possible interpretation. Most applications of ambiguity to portfolio home bias (Uppal & Wang 2003, Benigno & Nistico 2009) and own-company-stockholdings (Boyle et al. 2003) employ smooth models and do not focus on the extensive margin.

Epstein & Schneider (2007) compute a dynamic portfolio choice model with learning, using the recursive MP approach. They derive dynamic exit and entry rules as well as an intertemporal hedging demand. They also show that, quantitatively, learning about the equity premium can generate a significant trend toward stock market participation and investment, in contrast to results with Bayesian learning. The reason lies in the first-order effect of uncertainty on investment. Roughly, learning about the premium shrinks the interval of possible premia and thus works like an increases in the mean premium, rather than just a reduction in posterior variance, which tends to be second order. Campanale (2010) builds an MP model of learning over the life cycle. He shows that such a model helps explain participation and investment patterns by age in the U.S. Survey of Consumer Finances. Miao (2009) considers portfolio choice with learning and MP in continuous time. Faria et al. (2009) study portfolio choice when volatility is ambiguous.

3.2. Asset pricing

We now use the above results on portfolio choice to derive consumption-based asset pricing formulas. Our formal examples focus on representative agent pricing, because the literature on this issue is more mature and has proceeded to derive quantitative results; notes on new work on heterogenous agent models are provided below.

In equilibrium, a representative agent is endowed with a claim to consumption at date 2 and prices adjust so he is happy to hold on to this claim. Write date 2 consumption as $C_2 = C_1 \exp(\Delta c)$, where Δc is consumption growth. It is useful to distinguish between consumption and dividends. In our two-period economy, we call the payoff to stocks dividends. In a dynamic model, the second-period utility is a value function over wealth, and the payoff on stocks includes the stock price. The basic intuition is the same. Now, assume a share $1-\delta$ of consumption consists of labor income that grows at the constant rate μ_l and a share δ consists of dividends that have a lognormal growth rate Δd with variance σ_d^2 and an ambiguous mean $\mu_d \in [\bar{\mu}_d - \bar{x}, \bar{\mu}_d + \bar{x}]$. Using the same approximation as for the return on wealth above, write consumption growth as

$$\Delta c = (1-\delta)\mu_l + \delta\left(\Delta d + \frac{1}{2}\sigma_d^2\right) - \frac{1}{2}\delta^2\sigma_d^2.$$

The consumption claim trades at date 1 at the price P_c and has log return $r^c = logC_2 - logP_c = \Delta c - log(P_c/C_1)$. The premium on the consumption claim is

$$\mu^e = E[r^c] + \frac{1}{2}\text{var}(r^c) - r^f = (1-\delta)\mu_l + \delta\left(\mu_d + \frac{1}{2}\sigma_d^2\right) - \log(P_c/C_1) - r^f.$$

The representative agent solves a version of problem 3.3, given wealth $W = P_c + C_1$ and a range of premia μ^e generated by ambiguity in dividend growth μ_d. At the equilibrium price and interest rate, he must find it optimal to choose $\theta = 1$ and $C_1 = (P_c + C_1)/(1+\beta)$. The latter condition pins down P_c. With log utility, the price-dividend ratio on a consumption claim depends only on the discount factor.

The condition $\theta = 1$ pins down the interest rate. Because $\theta > 0$, minimization in Equation 3.3 selects the lowest premium, say $\underline{\mu}^e$, by selecting the lowest mean dividend growth rate $\bar{\mu}_d - \bar{x}$. Solving the condition $\theta = 1$ for the interest rate we obtain

$$r^f = -\log\beta + \left\{(1-\delta)\mu_l + \delta\left(\bar{\mu}_d + \frac{1}{2}\sigma_d^2\right) - \frac{1}{2}\delta^2\sigma_d^2\right\} - \frac{1}{2}\delta^2\sigma_d^2 - \delta\bar{x}. \quad (3.4)$$

The interest rate depends on the discount factor, the mean consumption growth rate (in braces), as well as a precautionary savings term. An increase in either risk or ambiguity makes the agent try to save more, which tends to lower the equilibrium interest rate. If $\delta < 1$, an increase in risk also raises the mean growth rate of consumption.

The same price and interest rate would obtain in an economy where the agent is not ambiguity averse but simply pessimistic: He believes that mean consumption growth is $\bar{\mu}_c - \bar{x}$ with certainty. This reflects a general point made first by Epstein & Wang (1994): Asset prices under ambiguity can be computed by first finding the most pessimistic beliefs about the consumption claim and then pricing assets under this pessimistic belief. We emphasize that this does not justify simply modeling a pessimistic Bayesian investor from the onset. For one thing, the worst-case scenario implied by a MP setup may look absurd when interpreted as a dogmatic Bayesian belief. Moreover, a form of the Lucas critique applies: The pessimistic investor is no more than a convenient "reduced form". Thus, focusing on him can give misleading answers to comparative statics (e.g., policy) questions.

Turn now to the stock price and equity premium. If a claim to dividends $D_2 = D_1 e^{\Delta d}$ trades at the price P_d, absence of arbitrage opportunities requires that its premium satisfy $\delta\left(\bar{\mu}_d - \bar{x} + (1/2)\sigma_d^2 - \log(P_d/D_1) - r^r\right) = \underline{\mu}^e$. The price-dividend ratio on equity is thus

$$P_d/D_1 = \exp\left(\bar{\mu}_d + \frac{1}{2}\sigma_d^2\right)\exp\left(-r^f - \left(\bar{x} + \delta\sigma_d^2\right)\right).$$

The price is the expected level of dividends under benchmark growth $\bar{\mu}_d$ (the first term), discounted at an uncertainty-adjusted rate that increases in both risk $\delta\sigma_d^2$ and ambiguity $\bar{x}$. Importantly, the degree of ambiguity $\bar{x}$ affects the discount rate one-for-one, but it affects the interest rate (Equation 3.4) only δ for one. For small δ, changes in ambiguity (for example, due to updating) have a large effect on stock prices but only a small effect on interest rates. This is important for addressing the equity volatility puzzle.

To discuss premia observed in the market, we need to take a stand on the true data-generating process. Suppose that dividend growth is drawn from a distribution with mean μ_d^* and variance σ_d^2. An econometrician who observes many realizations of the economy obtains a sample of excess returns $\Delta d - \log(P_d/D_1) - r^f$. The average premium measured by the econometrician is thus

$$\mu_d^* - \log(P_d/D_1) - r^f + \frac{1}{2}\sigma_d^2 = \delta\sigma_d^2 + \mu_d^* - (\bar{\mu}_d - \bar{x}). \tag{3.5}$$

It consists of a risk premium and an ambiguity premium. The risk premium is the covariance of consumption growth and stock returns. For an asset that represents only a small share of consumption, a large risk premium requires large payoff volatility σ_d^2.

The ambiguity premium consists of the difference between true mean dividend growth and the worst-case mean used by the agent to evaluate the asset. If the belief interval is centered around the truth ($\bar{\mu}_d = \mu_d^*$), then it is equal to $\bar{x}$. In any case, the share of the payoff in consumption does not matter for the ambiguity premium. If a lack of confidence in the asset is reflected in a range of premia, it raises the premium one-for-one. Put together, models of ambiguity aversion hold promise for resolving the equity premium and excess volatility puzzles, especially if the distinction between consumption and dividends is made explicit.

3.2.1. Amplification. With MP preferences, prices may depend very strongly—in fact, discontinuously—on fundamentals that change the representative agent's portfolio. This is the flip side of the portfolio inertia discussed in Section 3.1.2, which says that portfolios may not respond to small changes in prices. To illustrate, consider a family of models for the growth of stock payoffs similar to that used for returns above: Let $\mu_d = \bar{\mu}_d + x(\gamma - 1)/\gamma$

and $\sigma_d^2 = \bar{\sigma}_d^2 + 2x/\gamma$, where $x \in [0, \bar{x}]$ is ambiguous and $\gamma \in (0,1)$ is fixed. Intuitively, the agent believes that high growth goes along with high volatility. While the log-growth rate is decreasing in x, the adjusted-growth rate $\mu_d + \frac{1}{2}\sigma_d^2$ is increasing. Mean consumption growth now depends on x via the term $x\delta(1-\delta/\gamma)$, and the worst case is $x = 0$ if $\delta < \gamma$ and $x = \bar{x}$ if $\delta > \gamma$. If dividends make up a small part of consumption, agents fear low growth. As the share of dividends increases, concern with high risk eventually dominates.

The interest rate takes the form

$$r^f = -\log\beta + \bar{\mu}_c - \frac{1}{2}\delta^2\sigma_d^2 - \delta\bar{x}D,$$

where $D = 1$ if $\delta > \gamma$, $D = 0$ if $\delta < \gamma$ and $D \in [0,1]$ if $\delta = \gamma$. Viewed as a function of δ, the interest rate has a discontinuity at the point $\delta = \gamma$. A small change in fundamentals—here the share of stock payoffs in wealth—can have a large effect on asset prices. Intuitively, a small drop in the share of stock payoffs in wealth redirects agents' concern from high risk to low growth. This results in a jump in interest rates (and a crash in asset prices).

3.2.2. The cross section of returns and idiosyncratic ambiguity. To examine the cross section of stock returns, assume there is no labor income but that the consumption claim consists of n trees of the same size with i.i.d. lognormal dividend growth rates with mean μ_d and variance σ_d^2. Consumption growth is thus

$$\Delta c \approx \frac{1}{n}\sum_{i=1}^{n}\left(\Delta d_i + \frac{1}{2}\sigma_d^2\right) - \frac{1}{2n}\sigma_d^2.$$

Assume that dividend growth rates are perceived as ambiguous but independent: The vector of means μ_d is drawn from a Cartesian product of intervals $[\bar{\mu}_d - \bar{x}_i, \bar{\mu}_d + \bar{x}_i]$. Although the center of the interval is the same for all trees, the agent views some trees as more ambiguous than others.

In equilibrium, tree prices P_i and the interest rate are determined so that the agent is willing to hold all trees. With a true dividend growth rate $\mu_d^* = \bar{\mu}_d$ for all trees, similar algebra as above delivers an interest rate and measured stock premia

$$r^f = -\log\beta + \bar{\mu}_d + \frac{1}{2}\sigma_d^2 - \frac{1}{n}\sum\bar{x}_i - \frac{1}{2n}\sigma_d^2$$

$$\mu_d^* - \log(P_i/D_i) - r^f + \frac{1}{2}\sigma_d^2 = \frac{1}{n}\sigma_d^2 + \bar{x}_i$$

As the number of trees increases, the effects of risk on asset prices vanish. Indeed, both the precautionary savings term in the interest rate and the risk premium on a tree—the covariance of the tree return with consumption growth—go to zero. The effect of ambiguity on the interest rate also vanishes as aggregate consumption becomes less ambiguous. However, the ambiguity premium on an individual tree does not depend on the number of trees; it depends only on the ambiguity perceived about that individual tree.

3.2.3. Literature notes: representative agent pricing. Epstein & Wang (1994, 1995) first studied representative agent asset pricing with MP in discrete time and pointed out the possibility of amplification and price indeterminacy. Chen & Epstein (2002) characterize pricing in continuous time. Sbuelz & Trojani (2008) derive pricing formulas with entropy-constrained priors. Gagliardini et al. (2008) show how to apply detection probabilities in an MP setting. Epstein & Schneider (2008) consider the effect of learning, with a focus on the role of signals with ambiguous precision. They show that such signals induce an asymmetric response to news – bad news is taken more seriously than good news – and contribute to premia for idiosyncratic volatility as well as negative skewness in returns. Williams (2009) provides evidence that in times of greater uncertainty in the stock market the reaction to earnings announcements is more asymmetric. Illeditsch (2010) shows how learning from ambiguous signals can give rise to amplification, particularly in times when bad news arrives.

Another key property of ambiguous signals is that the anticipation of poor signal quality lowers utility. As a result, a shock that lowers the quality of future signals can lower asset prices. In contrast, in a Bayesian setting, the anticipation of less precise future signals does not change utility or prices so long as the distribution of payoffs has not changed. Epstein & Schneider (2008) use a quantitative model to attribute some of the price drop after 9/11 to the discomfort market participants felt because they had to process unfamiliar signals. There is a related literature on "information uncertainty" in accounting. For example, Autore et al. (2009) consider the failure of Arthur Anderson as an increase in (firm-specific) ambiguity about AA clients and document how the price effect of this shock depended on the availability of firm-specific information.

There are now a number of quantitative studies that apply the recursive MP model to different asset markets. Trojani & Vanini (2002. 2004) revisit the equity premium puzzle. Jeong et al. (2009) estimate a model of stock returns with an emphasis on time variation in equity premia. Drechsler (2008) studies the joint behavior of equity returns and option prices. Both Jeong et al. and Drechsler use a general specification of RMP with separate parameters for risk aversion and substitution as in Epstein & Zin (1989) and thus allow for the interaction of ambiguity and "long-run risk." Ilut (2009) addresses the uncovered interest parity puzzle in foreign exchange markets using a model of regime switching under ambiguity. Gagliardini et al. (2008) and Ulrich (2009) consider the term structure of interest rates, focusing on ambiguity about real shocks and monetary policy, respectively. Boyarchenko (2009) studies credit risk in corporate bonds.

There is also a growing literature on quantitative asset pricing with smooth models. Barillas et al. (2009) use multiplier preferences to reinterpret the equity premium results found by Tallarini (2000) in a model with Epstein-Zin utility. Hansen & Sargent (2009) and Chen et al. (2009) consider the behavior of stock returns in models with learning about hidden states, using multiplier and KMM utility, respectively. Kleschinski & Vincent (2009) study the real term structure in a model with robustness. Liu et al. (2005) consider the smirk in option premia.

3.2.4. Literature notes: heterogeneous agent models. Recent work has explored heterogenous agent models where some agents have MP. Epstein & Miao (2003) consider an equilibrium model in which greater ambiguity about foreign as opposed to domestic securities leads to a home bias. Several models center on portfolio inertia as discussed above. Mukerji & Tallon (2001) show that ambiguity can endogenously generate an incomplete market structure. Intuitively, if ambiguity is specific to the payoff on a security, as in 3.1.3 above, then no agent may be willing to take positions in a security with sufficiently ambiguous payoffs. Mukerji & Tallon (2004) build on this idea to explain the scarcity of indexed debt contracts with ambiguity in relative prices. Easley & O'Hara (2009) consider the welfare effects of financial market regulation in models where MP agents choose in what markets to participate.

A shock to the economy that suddenly increases ambiguity perceived by market participants can drive widespread withdrawal from markets, a "freeze." This is why MP has been used to capture the increase in uncertainty during financial crises (Caballero & Krishnamurthy 2008, Caballero & Simsek 2009, Guidolin & Rinaldi 2009, Routledge & Zin 2009). Uhlig (2009) considers the role of ambiguity aversion in generating bank runs.

In heterogenous agent models price generally depend on the entire distribution of preferences. An important point here is that if only some agents become more ambiguity averse, this may not increase premia observed in the market. The reason is that the more ambiguity averse group might leave the market altogether, leaving the less ambiguity averse agents driving prices (Cao et al. 2005, Chapman & Polkovnichenko 2009, Ui 2009). Illeditsch (2010) provides conditions on the distribution of preferences under which amplification effects (as discussed in Section 3.2.3) obtain with heterogeneous agents. Condie (2010) considers conditions under which ambiguity-averse agents affect prices in the long run even if they interact with SEU agents.

A number of papers have recently studied setups with ambiguity-averse traders and asymmetric information. Condie and Ganguli (2009) show that if an ambiguity averse investor has private information, then portfolio inertia (as in Equation 3.1.2) can prevent the revelation of information by prices even there is the same number of uncertain fundamentals and prices. Ozsoylev & Werner (2009) and Caskey (2009) study the response of prices to shocks when ambiguity-averse agents interact with SEU traders and noise traders. Mele & Sangiorgi (2009) focus on the incentives for information acquisition in markets under ambiguity.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

L.G.E. acknowledges financial support from the National Science Foundation (award SES-0917740).

LITERATURE CITED

Ahn D, Choi S, Gale D, Kariv S. 2009. *Estimating ambiguity aversion in a portfolio choice experiment*. Work. Pap., UC Berkeley

Anderson E, Hansen LP, Sargent TJ. 2003. A quartet of semigroups for model specification, robustness, prices of risk and model detection. *J. Eur. Econ. Assoc.* 1:68–123

Anscombe F, Aumann RJ. 1963. A definition of subjective probability. *Ann. Math. Stat.* 34:199–205

Autore D, Billingsley R, Schneller M. 2009. Information uncertainty and auditor reputation. *J. Bank. Financ.* 33:183–92

Boyarchenko N. 2009. *Ambiguity, information quality and credit risk*. Work. Pap., Univ. Chicago

Baillon A, Driesen B, Wakker PP. 2009. *Relative concave utility for risk and ambiguity.* Work. Pap., Maastricht Univ.

Barillas F, Hansen LP, Sargent TJ. 2009. Doubts or variability? *J. Econ. Theory* 144:2388–418

Benigno P, Nistico S. 2009. *International portfolio allocation under model uncertainty.* Work. Pap. 14734, NBER

Bossaerts P, Ghirardato P, Guarnaschelli S, Zame WR. 2010. Ambiguity in asset markets: theory and experiment. *Rev. Finan. Stud.* In press

Caballero R, Krishnamurthy A. 2008. Collective risk management in a flight to quality episode. *J. Financ.* 63(5):2195–230

Caballero R, Simsek A. 2009. *Fire sales in a model of complexity.* Work. Pap., MIT

Campanale C. 2010. Learning, ambiguity and life-cycle portfolio allocation. *Rev. Econ. Dyn.* In press

Cao HH, Wang T, Zhang HH. 2005. Model uncertainty, limited market participation, and asset prices. *Rev. Finan. Stud.* 18:1219–51

Cao HH, Han B, Zhang HH, Hirshleifer D. 2007. *Familiarity and fear of the unknown.* Work. Pap., Univ. North Carolina

Caskey J. 2009. Information in equity markets with ambiguity-averse investors. *Rev. Fin. Stud.* 22(9):3595–627

Chapman D, Polkovnichenko V. 2009. First order risk aversion, heterogeneity and asset market outcomes. *J. Financ.* 64(4):1863–87

Chen Z, Epstein LG. 2002. Ambiguity, risk and asset returns in continuous time. *Econometrica* 70:1403–43

Chen H., Ju N, Miao J. 2009. *Dynamic asset allocation with ambiguous return predictability.* Work. Pap., Boston Univ.

Condie S. 2010. Living with ambiguity: prices and survival when investors have heterogeneous preferences for ambiguity. *Econ. Theory.* In press

Condie S, Ganguli J 2009. *Ambiguity and partially-revealing rational expectations equilibria.* Work. Pap., Univ. Cambridge

Dow J, Werlang SR. 1992. Uncertainty aversion, risk aversion and the optimal choice of portfolio. *Econometrica* 60:197–204

Drechsler I. 2008. *Uncertainty, time-varying fear, and asset prices.* Work. Pap., Stern Sch. Bus., New York Univ.

Duffie D, Epstein LG. 1992. Stochastic differential utility. *Econometrica* 60:353–94; see also Appendix C with C Skiadas

Dupuis P, Ellis RS. 1997. *A Weak Convergence Approach to the Theory of Large Deviations.* New York: Wiley

Easley D, O'Hara M. 2009. Ambiguity and nonparticipation: the role of regulation. *Rev. Finan. Stud.* 22:1817–43

Ellsberg D. 1961. Risk, ambiguity and the Savage axioms. *Q. J. Econ.* 75:643–69

Epstein LG. 2010. *A paradox for the "smooth ambiguity" model of preference. Econometrica.* In press

Epstein LG, Schneider M. 2003. Recursive multiple-priors. *J. Econ. Theory* 113:1–31

Epstein LG, Schneider M. 2007. Learning under ambiguity. *Rev. Econ. Stud.* 74:1275–303

Epstein LG, Schneider M. 2008. Ambiguity, information quality and asset pricing. *J. Financ.* 63:197–228

Epstein LG, Wang T. 1994. Intertemporal asset pricing under Knightian uncertainty. *Econometrica* 62:283–322

Epstein LG, Zin S. 1989. Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica* 57:937–69

Gagliardini P, Porchia P, Trojani F. 2009. Ambiguity aversion and the term structure of interest rates. *Rev. Financ. Stud.* 22:4157–88

Gajdos T, Hayashi T, Talon JM, Vergnaud JC. 2008. Attitude towards imprecise information. *J. Econ. Theory* 140:27–65

Garlappi L, Uppal R, Wang T. 2007. Portfolio selection with parameter and model uncertainty: a multi-prior approach. *Rev. Financ. Stud.* 20:41–81

Gilboa I, Schmeidler D. 1989. Maxmin expected utility with non-unique priors. *J. Math. Econ.* 18:141–53
Guidolin M, Rinaldi F. 2009. *A simple model of trading and pricing risky assets under ambiguity: any lessons for policy makers*. Work. Pap., Fed. Reserve Board, St. Louis
Halevy Y, Ozdenoren E. 2008. *Ambiguity and compound lotteries: calibration*. Work. Pap., Univ. Br. Columbia
Hansen LP. 2007. Beliefs, doubts and learning: valuing macroeconomic risk. *Am. Econ. Rev.* 97:1–30
Hansen LP, Sargent TJ. 2001. Robust control and model uncertainty. *Am. Econ. Rev.* 91:60–66
Hansen LP, Sargent TJ. 2007. Recursive robust estimation and control without commitment. *J. Econ. Theory* 136:1–27
Hansen LP, Sargent TJ. 2009. *Fragile beliefs and the price of uncertainty.* Work. Pap., Univ. Chicago
Hayashi T. 2005. Intertemporal substitution, risk aversion and ambiguity aversion. *Econ. Theory* 25:933–56
Huber PJ. 1981. *Robust Statistics*. New York: Wiley
Illeditsch P. 2010. *Ambiguous information, risk aversion and asset pricing*. Work. Pap., Wharton Sch. Bus.
Ilut CL. 2009. *Ambiguity aversion: implications for the uncovered interest rate parity puzzle*. Work. Pap., Duke Univ.
Jeong D, Kim H, Park Y. 2009. *Does ambiguity matter? Estimating asset pricing models with multiple-priors recursive utility.* Work. Pap., Texas A&M
Ju N, Miao J. 2009. *Ambiguity, learning and asset returns*. Work. Pap., Boston Univ.
Kleschelski I, Vincent N. 2009. *Robust equilibrium yield curves*. Work. Pap., Washington Univ
Klibanoff P, Marinacci M, Mukerji S. 2005. A smooth model of decision making under ambiguity. *Econometrica* 73:1849–92
Klibanoff P, Marinacci M, Mukerji S. 2009. Recursive smooth ambiguity preferences. *J. Econ. Theory* 144:930–76
Kopylov I. 2009. Choice deferral and ambiguity aversion. *Theor. Econ.* 4:199–225
Liu J, Pan J, Wang T. 2005. An equilibrium model of rare-event premia and its implication for option smirks. *Rev. Financ. Stud.* 18:131–64
Maccheroni F, Marinacci M, Rustichini A. 2006a. Ambiguity aversion, malevolent nature, and the variational representation of preferences. *Econometrica* 74:1447–98
Maccheroni F, Marinacci M, Rustichini A. 2006b. Dynamic variational preferences. *J. Econ. Theory* 128:399–424
Mele A, Sangiorgi F. 2009. *Ambiguity, information acquisition and price swings in asset markets*. Work. Pap., Lond. Sch. Econ.
Miao J. 2009. Ambiguity, risk and portfolio choice under incomplete information. *Ann. Econ. Financ.* 10:257–79
Mukerji S, Tallon JM. 2001. Ambiguity aversion and incompleteness of financial markets, *Rev. Econ. Stud.* 68, 883–904
Mukerji S, Tallon JM. 2004. Ambiguity aversion and the absence of indexed debt. *Econ Theory* 24(3):665–85
Ozsoylev H, Werner J. 2009 *Liquidity and asset prices in rational expectations equilibrium with ambiguous information*. Work. Pap., Univ. Minn.
Routledge B, Zin S. 2009. Model uncertainty and liquidity. *Rev. Econ. Dyn.*12(4):543–66
Sbuelz A, Trojani F. 2008. Asset prices with locally-constrained-entropy recursive multiple priors utility. *J. Econ. Dyn. Control* 32:3695–717
Schmeidler D. 1989. Subjective probability and expected utility without additivity. *Econometrica* 57:571–87
Schroder M, Skiadas C. 1999. Optimal consumption and portfolio selection with stochastic differential utility. *J. Econ. Theory* 89:68–126
Seo K. 2009. Ambiguity and second-order belief. *Econometrica* 77:1575–605
Skiadas C. 2003. Robust control and recursive utility. *Financ. Stochast.* 7:475–89

Skiadas C. 2008. Dynamic porfolio choice and risk aversion. In *Financial Engineering*, ed. JR Birge, V Linetsky, 15:789–843. New York: Elsevier

Skiadas C. 2009a. *Asset Pricing Theory.* Princeton, NJ: Princeton Univ. Press

Skiadas C. 2009b. *Smooth ambiguity aversion toward small risks and continuous-time recursive utility.* Work. Pap., Kellogg Sch., Northwestern Univ.

Strzalecki T. 2010. *Axiomatic foundations of multiplier preferences. Econometrica*. In press

Tallarini TD. 2000. Risk-sensitive business cycles. *J. Monet. Econ.* 45:507–32

Trojani F, Vanini P. 2002. A note on robustness in Merton's model of intertemporal consumption and portfolio choice. *J. Econ. Dyn. Control* 26:423–35

Trojani F, Vanini P. 2004. Robustness and ambiguity aversion in general equilibrium. *Rev. Financ.* 8:279–324

Uhlig H. 2009. *A model of a systemic bank run.* Work. Pap., Univ. Chicago

Ui T. 2009. *The ambiguity premium vs. the risk premium under limited market participation.* Work. Pap., Yokohama Natl. Univ.

Ulrich M. 2009. *Inflation ambiguity and the term structure of arbitrage-free U.S. government bonds.* Work. Pap., Columbia Univ.

Uppal R, Wang T. 2003. Model misspecification and under diversification. *J. Financ.* 58:2465–86

Williams C. 2009. *Asymmetric responses to good and bad news: an empirical case for ambiguity.* Work. Pap., Univ. Michigan