Statistical Science
2016, Vol. 31, No. 4, 511–515
DOI: 10.1214/16-STS570
Main article DOI: 10.1214/16-STS592

Ambiguity Aversion and Model Misspecification: An Economic Perspective

Lars Peter Hansen and Massimo Marinacci

1. INTRODUCTION

In their paper, Watson and Holmes (2016) follow the statistical decision approach pioneered by Wald (1950). Under Wald's perspective, the aim of the analysis shifts from the discovery of a statistical "truth" (for instance, as revealed by the correct statistical model), to making decisions that are defensible according to posited objective functions that trade off alternative aims. We also draw on decision theory in our discussion because it provides a formal framework for confronting uncertainty.

Decades ago, Arrow (1951) distinguished two sources of uncertainty: (i) *risk within a model*, where the uncertainty is about the outcomes that emerge in accordance to a (probability) model that specifies fully the outcome probabilities; and (ii) *ambiguity among models*, where the uncertainty is about which alternative model, or convex combination of such models, should be used to assign the probabilities. If the true model is not assumed to be among the original set of models under consideration, a third source of uncertainty emerges, (iii) *model misspecification*, where uncertainty is induced by the approximate nature of the models under consideration to use in assigning probabilities. These different sources of uncertainty are inherent to any analysis that includes decision makers who have (probabilistic) theories about the outcomes and form beliefs over their relevance.

How to accommodate potential model misspecification is a challenging topic. On the one hand, if we have very precise information about the nature of the misspecification, then presumably we would fix or repair the model. On the other hand, if we allow for too large of a set of possible ways for a model to be misspecified, we may find that little can be said of value in confronting the decision problem. The interplay between tractability and conceptual appeal is a central consideration when producing tools that aid in statistical decision making. By using formal decision theory to frame and analyze misspecification, the advances described by Watson and Holmes (2016) help to nurture further connections between statistics and economics, and they allow for the incorporation of additional insights from statistics into economic analyses.

In their essay, Watson and Holmes (2016) draw connections to some research in economics. Our comment will describe other important advances in decision theory within the economics discipline that are designed to confront uncertainty conceived broadly to include an aversion to ambiguity and a concern about model misspecification. We will also delineate some special challenges for applications in economics.

Statisticians and econometricians use decision theory to guide how they estimate parameters and assess the specification of alternative models. But economics and other social sciences also use decision theory for a second purpose. People inside the models that we build must speculate about the future when making investment and other forward-looking decisions. These people inside the models could be individuals from the private sector of the market economy or they could be policy makers who employ econometricians to help them in making better decisions with either self-serving or social objectives. It has long been understood within economics that beliefs about the future are important inputs into model construction, and this opens the door to letting people inside the economic models to use data and statistical methods to help shape their beliefs. Economists debate what degree of statistical sophistication we should ascribe to the people inside our models. Thus, there is a role for depicting decision making under uncertainty to help capture behavior inside the models we build. Doing so can have substantial consequences for outcomes of the economic analysis. Statistical challenges, thus, show up in two places. They are present *inside* the models we build to capture the

Lars Peter Hansen is the David Rockefeller Distinguished Service Professor in Economics, Statistics and the College, University of Chicago, Chicago, Illinois 60637, USA (e-mail: lhansen@uchicago.edu). Massimo Marinacci holds the AXA-Bocconi Chair in Risk, Department of Decision Sciences and IGIER, Università Bocconi, 20136 Milano, Italy (e-mail: massimo.marinacci@unibocconi.it).

behavior of astute decision makers and *outside* these models when econometricians adapt and apply statistical methods to models that interest them. See, for instance, Hansen (2014) for more discussion of this point.

Statistical challenges inside the models we build influence how we shape and apply decision theoretic concerns about model misspecification. Moreover, dynamics are central to many economic analyses. This interest in dynamics has had a substantial impact in the development of decision theory within economics at least since Koopmans (1960) and also within control theory. The interplay between concerns about model misspecification in a dynamic, stochastic environment along with the desire for tractable recursive formulations has led to some recent advances in decision theory within economics that we will briefly describe [see, e.g., Gilboa and Marinacci (2013) and Marinacci (2015) for recent overviews].

2. DECISION THEORY

Decision theory aims to describe how a person should behave in an uncertain environment. We follow Watson and Holmes (2016) in our use of decision theory, except that we extend the notation in ways that will support our discussion of potential model misspecification in dynamic environments.

For a statistician, decision theory typically guides a choice of a model or an estimator of an underlying parameter vector captured by the possibly infinite dimensional parameter vector θ. Econometrics often adopts this same perspective. This is captured by an unknown parameter vector θ that resides in a set Θ. Given θ, a random vector X with realized values x in a set $\mathcal{X}$ is described by a probability model $f(x|\theta)$, a density relative to a measure τ over $\mathcal{X}$ that provides a probabilistic specification of X. A decision maker observes a realization x and takes an action $a \in \mathcal{A}$ that can depend on x. Formally, an *action* (or *decision*) *rule* is a suitably measurable function $A : \mathcal{X} \to \mathcal{A}$.

Represent the decision maker's preferences in terms of a utility function $U(a, x, \theta)$. Integrate over x to construct expected utility conditioned on θ:

$$\overline{U}(A|\theta) = \int_{\mathcal{X}} U\big(A(x), x, \theta\big) f(x|\theta)\tau(dx), \tag{1}$$

which will be an important ingredient of our decision theories. The expected utility, as we have computed it, conditions on the parameter θ that is typically unknown to the decision maker. Statistical decision theory often regards $-U$ as a *loss function* and calls $-\overline{U}$ a *risk function*. Consistent with this label, we view the integration over x conditioned on θ in equation (1) as adjusting for risk.

When it comes to applying decision theories to economics and other fields, the unknown parameter θ may be an intermediate target. For instance, consider a decision maker facing uncertainty captured by a future payoff relevant state. Represent this state as a random vector S with realized values s in a set $\mathcal{S}$. Let $f^*(s|a, x, \theta)$ denote the density relative to a measure τ^* over alternative s's in $\mathcal{S}$ conditioned on the current period action a and observed data x. Consider a next period utility function U^* that depends on (s, a) and integrate over s to construct $U(a, x, \theta) = \int_{\mathcal{S}} U^*(s, a) f^*(s|a, x, \theta)\tau^*(ds)$. Even when not directly payoff relevant, the θ dependence of U is induced by the dependence of f^* on θ. This formulation is amenable to dynamic programming methods where value functions are incorporated into the specification of U and U^*, and S is a future Markov state or a shock that determines this Markov state given current information.

As posed so far, this representation of decision theory is incomplete. Following de Finetti (1937) and Savage (1954), we include a subjective prior probability π, and integrate over the posited θ. With this, we complete the specification:

$$\int_{\Theta} \overline{U}(A|\theta)\pi(d\theta) \tag{2}$$

and can use this integral to rank alternative action rules A. Since the action rule depends on x, given π we may rank alternative actions a conditioned on x by $\int_{\Theta} U(a, x, \theta)\pi^*(d\theta|x)$, where π^* is the familiar Bayesian posterior $\pi^*(d\theta|x) \propto f(x|\theta)\pi(d\theta)$. A Bayesian statistician's job is to construct the posterior π^*.

As stated, however, in this specification there is no obvious scope for the expression of an aversion to model misspecification or to model ambiguity. As noted by Watson and Holmes (2016), both de Finetti and Savage acknowledge the challenge in using subjective probability to address such aversions.

3. MISSPECIFICATION

We share the Watson and Holmes (2016) interest of exploring the impact of model misspecification. Issues that we discuss in this section are already relevant for uncertainty induced by model ambiguity, but concerns about model misspecification magnify their importance potentially by expanding substantially the set of models under consideration. A rich set of extensions of decision theory has emerged to confront uncertainty broadly conceived. These advances include some mentioned by Watson and Holmes (2016), but there are also others. The alternative approaches alter the inputs into a Bayesian decision problem in a variety of ways. Some follow Wald's (1950) approach by relying on the game theoretic analysis of von Neumann and Morgenstern (1944) to shape an approach to uncertainty.

Watson and Holmes (2016) point to a rich literature on robust Bayesian methods that explores prior sensitivity in systematic ways. Motivated by applications in economic dynamics and control theory, decision theorists have found valuable to adopt the following starting point. Consider a convex function C to assess a decision maker's response to ambiguity about the prior π. The decision maker solves the following problem:

$$\max_{A\in\mathcal{A}} \min_{\pi} \int_{\Theta} \overline{U}(A|\theta)\pi(d\theta) + C(\pi). \tag{3}$$

The cost function imposes a penalty on the choice of prior. Penalization methods are well known in both statistics and control theory. The preferences implicit in this decision problem are what Maccheroni, Marinacci and Rustichini (2006b) call *variational preferences*. Such preferences nest the multiple priors specification of Gilboa and Schmeidler (1989), where the cost function C takes on the extreme form of being equal to infinity if π is outside a convex set of priors Π and zero inside. It extends the usual maximin approach to accommodate penalization, thus including formulations with a reference prior and a relative entropy penalty as proposed by Hansen and Sargent (2001). It accommodates robust Bayesian analysis in its systematic exploration of prior sensitivity with the use of a utility or loss function. The choice to minimize represents the aversion to ambiguity over the selection of the prior or a concern about prior misspecification.

We may think of Problem (3) as a zero-sum game. When the order of extremization can be reversed without altering the objective, then we may obtain a so-called worst-case prior under which the decision maker optimizes by taking this prior as given. Bayesians such as Good (1952) argue for assessing the plausibility of this (restrained) worst-case prior. While this forges a link to Bayesian decision theory, notice that this "choice" of prior depends on the utility or loss function rather than on a subjective introspection.

The theory of constrained optimization uses Lagrange or Kuhn–Tucker multipliers as a convenient way to enforce constraints, say on a convex family of priors. This suggests a direct connection between the more special Gilboa and Schmeidler (1989) approach and that of Maccheroni, Marinacci and Rustichini (2006b) in many applications, including those explored by Hansen and Sargent (2001) and Hansen et al. (2006) and mentioned in Watson and Holmes (2016). From the perspective of the axiomatic analysis of Maccheroni, Marinacci and Rustichini (2006b) and of dynamic implementation, there are important conceptual distinctions, however.

There is a seemingly different approach to ambiguity about a prior that also features ambiguity aversion. Instead of penalizing or constraining a family of priors, it introduces aversion to prior ambiguity in a way that is conceptually similar to risk aversion by including a strictly increasing concave function Φ as in the smooth ambiguity model of Klibanoff, Marinacci and Mukerji (2005). The decision maker then solves the following problem:

$$\max_{A\in\mathcal{A}} \Phi^{-1}\left(\int_{\Theta} \Phi[\overline{U}(A|\theta)]\pi(d\theta)\right). \tag{4}$$

While this problem is not posed as one with a concern about prior sensitivity, the informativeness or lack thereof in the prior does play a role in the decision criterion through curvature in the function Φ. As noted by Hansen and Sargent (2007), for the familiar and commonly used relative entropy formulation, there is a simple connection between the penalization approach to prior sensitivity and the smooth ambiguity approach. Let π_o denote a reference prior and g denote a probability density with respect to π_o. If $\mathcal{G}$ denotes the family of such densities, then

$$\begin{aligned}
&\min_{g\in\mathcal{G}} \int_{\Theta} \overline{U}(A|\theta)g(\theta)\pi_o(d\theta) \\
&\qquad + \lambda \int_{\Theta} g(\theta)\log g(\theta)\pi_o(d\theta) \\
&\quad = -\lambda \log \int_{\Theta} \exp\left[-\frac{1}{\lambda}\overline{U}(A|\theta)\right]\pi_o(d\theta).
\end{aligned}$$

This minimization problem is essentially a special case of an optimization problem with a relative entropy penalty that emerges in a variety of areas of applied mathematics and essentially follows from Theorem 4.1 in Watson and Holmes (2016). Thus, a particular form of a smooth ambiguity model emerges from a search over alternative prior densities subject to a penalization.

We close this section by observing that decision theories that extend (2) also violate the sure-thing principle, which has been viewed as a basic normative principle. However, the appeal of this principle becomes questionable under model uncertainty as Ellsberg (1961), pages 653–654, famously illustrated with the three-color paradox.

4. DYNAMICS

Incorporating dynamics raises a host of interesting questions and has nurtured the development of recursive formulations of decision problems. By design, these recursive formulations are amenable to application of dynamic programming methods which render the computation and characterization of solutions to economics models tractable. There are a variety of conceptual changes once we entertain model misspecification.

4.1 Misspecification in Future Dynamics

Connecting the future to the past is central to application of time series statistical methods. So far, we have presumed that the density f^* for the future state, as well as the density f used to represent the observed data, depend on an underlying parameter θ. This gives us the opportunity to learn using Bayesian updating, even though potential model misspecification can impede this process. But if past data are only revealing for some of the components of θ, there may remain uncertainty about how the economic environment will evolve in the future that is not fully captured by random shocks. This type of uncertainty motivated the work of Hansen and Sargent (2001), Chen and Epstein (2002) and Anderson, Hansen and Sargent (2003). It is also a central motivation for the dynamic extension of the class of variational preferences as discussed in Maccheroni, Marinacci and Rustichini (2006a).

4.2 Dynamic Consistency

Dynamic applications of decision making under uncertainty, often, but not always, look for formulations that are dynamically consistent to avoid having decision makers play naive or sophisticated games against future versions of themselves. This aim can shape the formulation of decision problems. Indeed the penalization approach suggested by Hansen and Sargent (2001) and the generalization developed by Maccheroni, Marinacci and Rustichini (2006a) were motivated in part by such concerns.

Dynamic extensions of the multiple priors model of Gilboa and Schmeidler (1989) led Epstein and Schneider (2003) to embed a subjectively specified set of models or priors over models into a potentially larger rectangular set based on considerations of dynamic consistency. Such an approach does not always yield interesting answers, however. Suppose we use dynamic versions of statistical discrepancies, including the ones used by Watson and Holmes (2016), to constrain ways in which models could be misspecified. The resulting rectangular embeddings are so large as to lead to degenerate outcomes [see Hansen and Sargent (2016) for a formal discussion]. This phenomenon has led Hansen et al. (2006) to seek recursive implementations of so-called commitment problems solved from an *ex ante* perspective, much like what often occurs in control theory as reflected say in Petersen, James and Dupuis (2000). The added flexibility of the dynamic version of variational preferences characterized Maccheroni, Marinacci and Rustichini (2006a) gives an alternative attractive way to support the use of statistical discrepancy measures as noted originally by Hansen and Sargent (2001) and developed more fully in Hansen et al. (2006).

4.3 Robustness and Dynamic Learning

In dynamic settings, yesterday's posterior is today's prior. Bayes' rule is a wonderful and often convenient recursion. Even under potential misspecification, guesses about the future are typically tied to past evidence. Robust learning and its impact on decisions provides a fascinating twist to the direct Bayesian approach. There are a variety of approaches that have been suggested, including direct application of Bayes' rule applied to a reference prior over models with an accompanying adjustment to the recursively generated posteriors [see, e.g., Hansen and Sargent (2007) and Klibanoff, Marinacci and Mukerji (2009)]. Alternatively, Epstein and Schneider (2003) suggest a prior-by-prior application of Bayes' rule in a dynamic multiple priors models with a rectangular embedding to enforce consistency. In a model with date zero prior ambiguity, Chamberlain (2000) embraces the prior-by-prior application of Bayes' rule from an *ex ante* perspective without including a rectangular embedding.

There remains scope for further dialog and discussion of this important topic as there are potentially important tradeoffs between conceptual appeal and computational tractability. We very much hope that the essay by Watson and Holmes (2016) and our discussion will encourage further synergistic research linking statistical challenges to economic model building and analysis.

ACKNOWLEDGEMENTS

We thank Pierpaolo Battigalli, Thomas Sargent, Stephen Stigler and an anonymous referee for helpful comments.

REFERENCES

ANDERSON, E., HANSEN, L. P. and SARGENT, T. J. (2003). A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *J. Eur. Econ. Assoc.* **1** 68–123.

ARROW, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situtations. *Econometrica* **19** 404–437. MR0046020

CHAMBERLAIN, G. (2000). Econometric applications of maxmin expected utility. *J. Appl. Econometrics* **15** 625–644.

CHEN, Z. and EPSTEIN, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica* **70** 1403–1443. MR1929974

DE FINETTI, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036

ELLSBERG, D. (1961). Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* **75** 643–669.

EPSTEIN, L. G. and SCHNEIDER, M. (2003). Recursive multiple-priors. *J. Econom. Theory* **113** 1–31. MR2017864

GILBOA, I. and MARINACCI, M. (2013). *Ambiguity and the Bayesian Paradigm*, *in Advances in Economics and Econometrics*: *Theory and Applications* (D. ACEMOGLU, M. ARELLANO and E. DEKEL, eds.). Cambridge Univ. Press, Cambridge.

GILBOA, I. and SCHMEIDLER, D. (1989). Maxmin expected utility with nonunique prior. *J. Math. Econom.* **18** 141–153. MR1000102

GOOD, I. J. (1952). Rational decisions. *J. Roy. Statist. Soc. Ser. B.* **14** 107–114. MR0077033

HANSEN, L. (2014). Nobel lecture: Uncertainty outside and inside economic models. *J. Polit. Econ.* **122** 945–987.

HANSEN, L. P. and SARGENT, T. J. (2001). Robust control and model uncertainty. *Am. Econ. Rev.* **91** 60–66.

HANSEN, L. P. and SARGENT, T. J. (2007). Recursive robust estimation and control without commitment. *J. Econom. Theory* **136** 1–27. MR2888400

HANSEN, L. P. and SARGENT, T. J. (2016). Sets of models and prices of uncertainty. Working Paper No. 22000.

HANSEN, L. P., SARGENT, T. J., TURMUHAMBETOVA, G. and WILLIAMS, N. (2006). Robust control and model misspecification. *J. Econom. Theory* **128** 45–90. MR2229772

KLIBANOFF, P., MARINACCI, M. and MUKERJI, S. (2005). A smooth model of decision making under ambiguity. *Econometrica* **73** 1849–1892. MR2171327

KLIBANOFF, P., MARINACCI, M. and MUKERJI, S. (2009). Recursive smooth ambiguity preferences. *J. Econom. Theory* **144** 930–976. MR2887278

KOOPMANS, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica* **28** 287–309. MR0127422

MACCHERONI, F., MARINACCI, M. and RUSTICHINI, A. (2006a). Dynamic variational preferences. *J. Econom. Theory* **128** 4–44. MR2229771

MACCHERONI, F., MARINACCI, M. and RUSTICHINI, A. (2006b). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* **74** 1447–1498. MR2268407

MARINACCI, M. (2015). Model uncertainty. *J. Eur. Econ. Assoc.* **13** 998–1076.

PETERSEN, I. R., JAMES, M. R. and DUPUIS, P. (2000). Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Trans. Automat. Control* **45** 398–412. MR1762853

SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York. MR0063582

VON NEUMANN, J. and MORGENSTERN, O. (1944). *Theory of Games and Economic Behavior*. Princeton Univ. Press, Princeton, NJ. MR0011937

WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York. MR0036976

WATSON, J. and HOLMES, C. (2016). Approximate models and robust decisions. *Statist. Sci.* **31** 465–589.