How (Not) to Do Decision Theory[1]

Eddie Dekel[2] Barton L. Lipman[3]

Current Draft
October 2009

[1]We thank Daron Acemoglu, Sandeep Baliga, Larry Epstein, Sivan Frankel, Drew Fudenberg, Tzachi Gilboa, Faruk Gul, Edi Karni, Peter Klibanoff, David Kreps, Michael Luca, Mark Machina, Massimo Marinacci, Klaus Nehring, Jawwad Noor, Ady Pauzner, Wolfgang Pesendorfer, Andy Postlewaite, Matt Rabin, Ariel Rubinstein, Aldo Rustichini, Larry Samuelson, Todd Sarver, Uzi Segal, Marciano Siniscalchi, Rani Spiegler, and Asher Wolinsky for comments on this paper and/or valuable discussions on decision theory and related topics. We also thank the National Science Foundation, grants SES–0820333 (Dekel) and SES–0851590 (Lipman), for support for this research.
[2]Economics Dept., Northwestern University, and School of Economics, Tel Aviv University E–mail: dekel@nwu.edu.
[3]Boston University. E–mail: blipman@bu.edu.

Abstract

We discuss the goals and means of positive decision theory and the implications for how to do decision theory. We argue that the goal of positive economic theory generally is to provide predictions and understanding and that representation theorems and other results of decision theory should be seen as ways to achieve these goals. We also argue that the "story" of a model is relevant to whether and how we use the model, that psychological considerations are not necessary for useful decision theory but can be helpful, and that nonchoice data, interpreted properly, can be valuable in predicting choice and so should not be ignored.

Whatever it is, I'm against it.

— Sung by Groucho Marx in *Horsefeathers*.

1 Introduction

In recent years, a basic building block of economics — the theory of individual decision making — has become the renewed focus of an enormous amount of research and reconsideration from many different perspectives. In decision theory, originally the center of such research, the insights of Allais (1953) and Ellsberg (1961, 2001), followed up by the work of Machina (1982), Schmeidler (1989), and Gilboa and Schmeidler (1989) led to a complete reconsideration of the classical notion of expected utility and subjective probability.[1] This research influenced and was influenced by the experimental and theoretical work of Kahneman and Tversky (1979) and many others in psychology who pushed for an even more fundamental reconsideration of how people make economic decisions. The latter research fed into work in economics by Laibson, Rabin, and others, who developed the behavioral approach to economics. In the same period, Rubinstein began the development of formal models which gave voice to the earlier insights of Simon (1982) regarding procedural rationality.

The ideas from outside of decision theory have reverberated back to decision theory again, leading to an explosion of research. Recent work has continued to develop our understanding of issues like ambiguity aversion and has gone further into a reconsideration of more fundamental notions such as optimization itself, new issues such as temptation, and revisiting many older concerns such as incompleteness and regret.[2]

Given this large scale reconsideration, it is perhaps unsurprising that many economists, including but not limited to decision theorists, are rethinking what decision theory is, what it can do, and how we should think about it. For a sampling from the wide variety of critical perspectives, see Caplin and Schotter (2008), Fudenberg (2006), and Gilboa (2009).

Further complicating an assessment of decision theory is the fact that it has never been an entirely unified field. Some work in decision theory has been motivated by essentially

[1]For a detailed bibliography of the literature, see Wakker (2009). For an excellent introduction, see Kreps (1988).

[2]E.g., see, respectively, Maccheroni, Marinacci, and Rustichini (2006) and Siniscalchi (2009); Kalai, Rubinstein, and Spiegler (2002); Gul and Pesendorfer (2001); Bewley (2002) and Dubra, Maccheroni, and Ok (2004); Sarver (2008).

normative questions such as characterizing how a "rational" decision maker should act. Other work takes an entirely positive or predictive approach, seeking convenient modeling tools for summarizing real behavior, rational or not.

In this paper, we discuss our views of positive decision theory and what it can contribute to economic theory. In a nutshell, we argue that the goal of positive economic theory is to provide useful predictions and ways to understand the world, these goals being related but not identical. We distinguish between the "story" of a model and its predictions. While the story need not be literally true for the model to be useful, it plays an important role. Confidence in the story of the model may lead us to trust the models predictions more. Perhaps more importantly, the story affects our intuitions about the model and hence whether and how we use and extend it.

Unfortunately, in seeking to develop a "good" model, one runs into some difficult tradeoffs, the resolution of which will typically depend on the specific goals of the modeler. For example, a theory which helps us understand one particular issue better will often necessarily oversimplify other issues. A more plausible story may come at the cost of reduced tractability.

In Section 2, we discuss the goals of positive economic theory generally and decision theory more specifically. Section 3 turns to a discussion of the kinds of results in decision theory and what they are intended to accomplish. In Section 4, we consider some of the implications of the perspective we offer. Section 5 concludes.

We close this introduction with caveats. In writing this paper and rereading what others have said on these topics, we've learned how difficult it is to make broad methodological statements clearly and succinctly. We apologize if we have unintentionally distorted anyone's views and hope that we have made our views clear enough to avoid misinterpretation. Also, many of the points we make have been made many times before, too many for us to provide anything close to complete citations. We err on the side of providing few or no specific cites on each of these points. Finally, our goal is that this paper says something useful both to decision theorists and to those unfamiliar with decision theory. We hope each audience will be patient with our comments to the other.

2 Positive Decision Theory: Ends

In this section, we first give a broad description of what decision theory does and how this fits into what economic theory, more broadly, does. Next we discuss why we use models for these purposes. Finally, we discuss how we select among alternative models. In our view, since positive decision theory is part of positive economic theory, the goals of the former are a special case of the goals of the latter. Consequently, much of our discussion of decision theory will apply to economic theory more broadly and we sometimes find it convenient to use some examples from outside decision theory.

Briefly, positive economic theory provides a mathematical language to formalize certain informal intuitions about the world to aid us in understanding what we observe and in predicting, two goals we later argue are linked.[3] Decision theory focuses specifically on predictions about *choice* behavior. As such, it provides predictions directly and also provides a tool for use in a broader model to make predictions about other economic issues.

In particular, decision theory provides a formalization of certain intuitive ideas about decision making and relates these to potentially observable choice behavior. Typically, decision theory develops a "model" in the form of a class of utility functions (e.g., expected utility) which is used to predict choice behavior.[4]

For example, suppose we want to model "temptation." The first step is to identify the behavior that we think corresponds to our intuitive notion of temptation. This step, giving "temptation" a behavioral definition, is in many ways the crux of what decision theory does — see Section 4.3 for more on this point. Since it would be too difficult to work with a model that assumes only this behavioral definition, we add some assumptions on behavior that simplify and either seem plausibly true or at least plausibly unrelated to the issue of temptation. We then relate these assumptions to a particular functional form for the utility function. This result thus identifies a functional form that we can think of as a useful formal embodiment of our intuitive notion of temptation.

Decision theory is used for a range of purposes, from predictions about individual choice (as in demand theory) to being an ingredient in a much larger model. For an example focused on individual choice, we might use the model of temptation to analyze how demand functions are affected by temptation. For an example of using decision theory as an ingredient, suppose we observe that the addition of dessert items by a

[3] Much work in theory is to achieve a better understanding of the models themselves with an indirect, long run goal of using this to understand the world better. While the boundary between understanding the models and understanding the world is not always clear, we generally focus here on the latter. This should not be taken to imply a view that the latter is superior, only that we think its role is more central in the debate over decision theory.

[4] For simplicity, throughout this paper, we will describe a model in decision theory in terms of utility functions. However, classic models in decision theory also include other objects such as subjective probabilities. Recent models have introduced many other approaches to representing behavior. For example, the agent is represented with more than one utility function in the literature on temptation (Gul and Pesendorfer (2001)) and incomplete preferences (Dubra, Maccheroni, and Ok (2004)) and with a set of probabilities in the literature on ambiguity (Bewley (2002), Gilboa and Schmeidler (1989)). Another type of model involves criteria which eliminate certain options before applying utility maximization (Eliaz and Spiegler (2008), Manzini and Mariotti (2007)).

restaurant leads to a short–run increase in sales, followed by a severe downturn. Suppose we hypothesize that people are surprised by the tempting items at first and give in to temptation, but afterward avoid these restaurants. Then we can use our model of temptation as an ingredient in analyzing this question.

To discuss how to select among models, we find it necessary to first explain the purpose of models. Suppose empirical observations or our intuition about the world suggests the hypothesis that A leads to X where A and X are statements about observable variables.[5] We claim that having a model explaining why A leads to X would be useful above and beyond our conjecture or whatever empirical observations led to it. Constructing such a model forces us to define a mathematical analog to A and X as well as to make various auxiliary hypotheses relating them. What do we learn from this process?

First, we may learn whether our intuition is flawed or not. For example, we may find that our seemingly sensible intuition of why A leads to X actually requires an additional, unpalatable assumption of B, leading us to question our initial conjecture. Alternatively, perhaps the model also implies an unexpected and unpalatable conclusion Y or implies that there is an internal inconsistency in the logic that A would lead to X.[6] On the other hand, we may find that only relatively weak and plausible additional assumptions are required, reinforcing our initial intuition.

Second, we may flesh out our initial intuition in ways that enable additional or better predictions. For example, if we see that A also leads to an unexpected implication of Y, then we have a new prediction. Also, if we see that A only leads to X when some other condition B holds and that A has a different implication otherwise, then we improve our ability to predict. If B is observable, the recognition of its role gives us a different prediction for those out of sample forecasts when A holds and B does not. For example, the initial discovery of revenue equivalence in independent private value auctions indicated that some other feature (the B) such as a common–value component is needed to explain the ubiquity of English auctions (the X).

Finally, even if a model doesn't immediately change or enlarge our set of predictions, it may yield a clearer understanding of why A might cause X. Why would such an understanding be useful? We see the primary value of such understanding in the way it may lead in the long run to more or better predictions. Lest this comment be misinterpreted, we emphasize that understanding may involve concepts whose translation into observables is not direct or immediate. Consequently, in the short run it may be

[5] Our arguments also naturally apply to other kinds of hypotheses. For example, we may instead hypothesize that A and X are correlated because of some (possibly unobservable) common cause Z.

[6] Paul Krugman gives a good illustration of this in his first Lionel Robbins lecture in June 2009: "What I did at the time, a very economist thing to do, was to build myself a little model to prove the point that I believed. So I built a little intertemporal optimizing whatever and to my shock — and this is the point, of course, of doing models — it actually gave me the opposite answer."

difficult to know exactly what these predictions might be. Finally, we emphasize that "prediction" should be broadly understood to mean a perhaps imprecise statement about a relationship between observable variables. An imperfect example is Spence's (1973) paper on job market signaling. While he did provide the new prediction that, controlling for human capital, workers with higher education would receive higher wages, the greater contribution of the paper seems to be in generating an understanding of signaling that ultimately led to many other predictions in other environments.

Given that these are the ways models are useful to us, how should we select among models? Of course, in light of the importance of a model's predictions, all else equal, a more accurate model with a wider range of predictions is a better model. For example, if A is the description of an agent's choice problem and X is his purchase of insurance, we could trivially explain the choice by saying that he just likes to buy insurance policies. However, a fuller and therefore more appealing explanation is that insurance reduces risk and the agent values it for this reason. One reason this explanation would be more appealing is that it would lead us to make other predictions about his behavior — e.g., investment decisions. Hence a decision–theoretic model which provides a formal notion of risk and risk aversion provides a broader range of other predictions.

Some discussions of model selection seem to stop at this point, taking the range and accuracy of the set of predictions of a model to be the sole criterion for judging its usefulness. It is common to suggest that we reject a model if and only if its predictions are refuted by the data, retaining it as "unrejected" in the meantime. We think such an exclusive focus on refutation misses some important considerations. Indeed, in choosing between two models, there may be good reasons in some cases for preferring the less accurate one. Many of these reasons have to do with the "story" of the model — the informal interpretation of the mathematics.

Consider, for example, the theory of subjective expected utility. The story of this model is that the agent forms subjective probability beliefs about the uncertainty he faces and chooses that action which maximizes expected utility as computed with these beliefs. To us, this story seems very idealized but intuitive at a basic level. While we don't believe the story is literally true, the idea that agents have beliefs of some kind which guide their choices, choose actions they consider more likely to work out well, etc., seems natural. (Of course, the degree of idealization involved may lead some readers to conclude that the story of subjective expected utility is not plausible at all.)

There are at least two reasons why the story of subjective expected utility matters. First, if we accept the story as plausibly true, this should make us more comfortable in relying on the models predictions. As Kreps (1990) argues, this consistency with intuition is just another kind of consistency with data. Thus in making out of sample predictions, we may be more inclined to trust an intuitive model with slightly worse predictions in sample than a less intuitive model that is more consistent with sample data. Conversely, if we find the story implausible, this may make us less willing to accept the predictions.[7]

Second, even if we dismiss the story as descriptively implausible, it is still very handy for the modeler's reasoning process. That is, it is useful to organize our thinking around ideas like beliefs, information, expectations, and the other concepts suggested by this model. Moreover, since models don't simply sit on a shelf but are to be used, tractability is valuable. For example, a model may be taken to new domains to generate new predictions/understanding. While the story isn't the sole determinant of tractability, having a nice intuition about the story facilitates the use of the model. Similarly, a more intuitive model is likely to be more valuable in helping us understand the world. To see the point, suppose we had a black box which, when input a choice set, etc., would immediately tell us the agent's choice from that set. E.g., in a game, we could input the behavior of the opponents and it would give the response. Even if this predicts perfectly, it's not clear how useful it is. For example, we couldn't use this to compute an optimal mechanism since without understanding the structure of the function describing this agent, we'd have to search over every possible mechanism. It is also unclear how we would "use" this black box in an applied model.[8] Hence even a flawed intuition might be more useful. In short, general principles, even ones which are not entirely accurate, may be more comprehensible and hence more useful than very detailed and accurate specifications.

A final problem with refutation as the exclusive criterion is the obvious but important point that a model is *never* a completely accurate description of the world. Having a model which is exactly correct would require the model to be as complex as the world, something about as useful as a map with a scale of "1 inch = 1 inch."[9] Less facetiously, all understanding is generated by deciding which aspects of reality to ignore. The world never repeats exactly the same situation twice, so if all aspects of reality are relevant, it is impossible to learn from one situation how to predict what will happen in another.

Thus almost every model is refuted in the strictest sense of the term. To see the implication, consider again our discussion of a model of above regarding how A might imply X. We noted that the model might tell us that this intuitive story, once formalized, has the additional, unexpected, prediction of Y. Suppose that Y is something absurd,

[7]Gul and Pesendorfer (2008) argue forcefully that the implausibility of the story of a model cannot refute the model. We entirely agree. However, the implausibility may make us less confident in the predictions of the model.

[8]Of course such a black box would change the way we study models. For example, we would compare the models predictions to those of the black box. But it seems to us that the need for simple and extendible models would remain.

[9]According to Wikipedia, this idea originated in Lewis Carroll's 1893 *Sylvie and Bruno Concluded*, though other readers may recognize it from Jorge Luis Borges' 1946 short story "Del Rigor en la Ciencia" ("The Rigors of Science"), Umberto Eco's essay "On the Impossibility of Drawing a Map of the Empire on a Scale of 1 to 1" in his 1995 book *How to Travel with a Salmon and Other Essays*, or a joke by Steven Wright.

either rejected by data or intuitively implausible. Does this mean we reject the model?

Not necessarily. We know that our model is unrealistically simple. Hence we should not be surprised if it has some odd implications, either about other issues such as Y or in the form of an excessive magnitude of the effect of A on X. The key question is whether we believe that the simplifications which led to the odd predictions are crucial to the model's explanation of why A leads to X. If so, we should conclude the model is inadequate in providing such an explanation, undermining our confidence in its predictions more broadly. Depending on whether we continue to believe our original hypothesis about the relationship between A and X, we would reconsider either the formalization of our idea or the idea itself. On the other hand, if the simplified aspects of the model which lead to implausible predictions are not important to our explanation for the relationship between A and X, then it seems reasonable to continue to use the model as a working hypothesis for formulating predictions about A and X, at least qualitatively.

To explain the point, return to the insurance example above. We suggested that an intuitive explanation of the observation of an agent purchasing insurance is that he is averse to risk and buys insurance to reduce his risk. We might then adopt the simplest approach to formalizing this explanation, assuming that we can write utility as a function of money only, assuming the agent maximizes expected utility, defining risk and risk aversion in this context, and assuming that the risk aversion property holds globally. This yields a simple, intuitive model with many predictions. Suppose that our goal with this model is to make predictions about the agent's other insurance purchases. For example, perhaps our initial observation concerned the purchase of house insurance and we wish to predict choices regarding car and boat insurance. If our predictions of these decisions are reasonably accurate, then we should not be too troubled by the observation that the agent occasionally plays poker with his friends, even though this contradicts the joint hypothesis that utility depends only on money and that the agent is globally risk averse. We would argue this is not troubling for two reasons. First, we know that our hypothesis that the agent cares only about the amount of money he has is a simplification and that we could easily change that assumption to incorporate the utility of games with friends to reconcile the model with this observation. Second, the poker games were not our focus in constructing the model, so inaccuracy in that dimension is not costly. Thus we would not bother to address this observation.[10]

Similarly, if we have data which refutes an auxiliary assumption of the model, this need not be a reason to abandon the model. For example, suppose we have evidence that the agents we are studying do not maximize expected utility, but the model still does a good job of predicting insurance decisions. Since this data does not refute the predictions of the model, we see no reason to abandon it, though this may reduce confidence in its out

[10] Of course, if the original model is consistent with the agent's poker playing, this is only mild support for the model for the same reasons.

of sample predictions. Of course, our confidence in the predictions would be enhanced if we knew it were possible to modify the expected utility model to generate the same predictions about insurance and avoid refutation by other data.

A perhaps more controversial point is that even if we are initially making poor predictions about insurance choices, this might not lead us to abandon our general explanation based on risk aversion, though it forces us to revise our specific model. For example, if we used expected utility and we find that choices of insurance similar to those generating the Allais paradox are causing our model to mispredict, we may find that we can switch to a generalized expected utility model a la Machina (1982) with risk aversion and predict quite well. The observation of Allais type behavior leads us to drop the expected utility aspect of our model, but if our main idea was about risk aversion, not expected utility, this change is not saying the original insight was "wrong."

As this discussion highlights, part of the problem with a focus on refutation is that almost anything can be explained in almost any broadly defined class of models by "tweaking" the model appropriately. If we have no constraints on the way we can choose the factors which influence the utility of the agent, define how the objects in the world map into these factors, and choose a functional form, surely we can rationalize any data.

So when would we abandon our explanation of why A leads to X? The point is that if we have to do a huge number and variety of tweaks, then the resulting model loses all the properties we said were valuable. It ceases to be tractable. It ceases to be intuitive. Unless our new situation is very close to the data our model has been constructed to fit, we lose the ability to make out of sample predictions with any degree of confidence. At this point, we conclude that we need a better model, presumably one based on a completely different explanation of the behavior we observe.

To be clear, it is valuable when a single model is consistent with a wide variety of empirical observations. However, when a model achieves this consistency only by adding a tweak or an additional parameter for each data point, then the model adds no understanding and is unlikely to predict well (except in situations that are very close to our past observations). The only valuable version of a model with such breadth is one which explains a variety of observations with a small number of basic principles.[11]

In the end, choices between models will hinge on a number of considerations. In addition to consistency with data, we value intuitive appeal (in both senses discussed above), tractability, and the range of additional predictions/understanding the model generates, several of which depend at least in part on the story of the model. Hence the choice of a model will depend on the purpose the model is used for, the modeler's intuition, and the modeler's subjective judgement of plausibility. Since all of these things

[11]This is reminiscent of the way Copernicus' model of astronomy replaced the Ptolemaic model.

vary across economists and between economists and psychologists, it should not surprise us to see different models chosen by different social scientists, especially if they wish to focus on different aspects of the issue at hand.[12] We recognize the importance of minimizing the role of subjective judgements and are not arguing that all intuitions are equally valid. One economist may reject another's intuition and, ultimately, the marketplace of ideas will make some judgements. (Given the amount of market failure in this context, perhaps a sociological notion is more apt than our economic metaphor.)

3 Positive Decision Theory: Means

Broadly speaking, there are two branches of work in positive decision theory. One group of papers (see, for example, Gollier (2001) and the references therein) develops new tools, e.g., different measures of risk, for working with existing models. The other, which we focus on henceforth, develops models to introduce new considerations to economics. For example, Schmeidler (1989) and Gilboa–Schmeidler (1989) developed the notion of ambiguity aversion where agents prefer risks based on known probabilities to those based on unknown, a phenomenon impossible in the usual subjective expected utility model.

As discussed above, one identifies the behavior that is intuitively associated with the new issue to be introduced. Since the new behavior is, by hypothesis, inconsistent with standard models, this requires modifying some of the standard assumptions. Typically, one drops as little of the standard assumptions as possible, retaining the rest of the standard assumptions for the purposes of simplification as discussed earlier. This is also useful because it makes it easier to connect the new ideas to the existing literature to see how the model with the "new" ingredient affects our understanding.

A related methodological consideration is that we do not consider multiple new ingredients at the same time. For example, we have mentioned temptation and ambiguity aversion as new ingredients that have been added in recent years. Yet the models which introduced temptation did not include the previously added ingredient of ambiguity aversion. Why not? The point is that these seem like conceptually distinct issues. If so, then the simplest way to achieve understanding of these two issues would be to first understand each in isolation. If we realize some connection between the two issues, perhaps because some phenomenon of interest stems from the interaction of them, then and only then would a model that combines these new ingredients be of interest. Indeed, an important aspect of a new breakthrough could well be pointing out connections between issues which previously seemed unrelated.

[12]We are not experts on the philosophy of science, but our understanding is that the unavoidability of subjective judgements in science is widely acknowledged there.

Turning to specifics, most theorems in decision theory are about the following issues:

1. What is the behavior corresponding to a particular functional form?
2. To what extent does the behavior in question identify parameters in the function?
3. How do intuitive changes in the behavior correspond to changes in the parameters?

The first kind of result is called a representation theorem since it shows that a particular functional form represents (corresponds to) certain behavior. We refer to the second kind of result, for obvious reasons, as an identification theorem or a uniqueness result. Finally, we refer to the last kind of result as a comparative or comparative static. In the sections that follow, we explain the importance of each of these questions.

3.1 Representation Theorems

A representation theorem relates a model of decision making (a class of utility functions) to properties of the implied choices. Choices may be formalized as a choice function/correspondence (a mapping from feasible sets to choices) or as a preference relation (which we interpret as *revealed* preference, a choice function where we only see choices from pairs). More precisely, a representation theorem completely identifies the behavior that corresponds to the model by showing that there is a utility function of the type postulated by the model which predicts the agent's choice if and only if those choices satisfy certain properties often referred to as *axioms*.

The purpose of a representation theorem is to provide better understanding of the model in order to give guidance in choosing among models for applications or other purposes. Naturally, the choice among models hinges in large part on the nature of each model's predictions and representation theorems are nothing more — nor less! — than a precise statement of these.

By definition, of course, the class of utility functions under consideration directly implies the class of behavior it predicts, so what does the representation theorem add? Often it is difficult to recognize the predicted behavior simply from looking at the functional form. Furthermore, a representation theorem may be helpful in identifying which aspects of the functional form should be thought of as convenient simplifications and which are crucial to the particular behavioral phenomenon of interest. The role of the representation theorem is to present the implied choice behavior in a form that is easy to understand, evaluate, and test. It should separate out the implications that are otherwise "standard" (e.g., completeness) from those that are special to the particular issue being studied, e.g., risk, uncertainty, ambiguity, or temptation.

The functional form often corresponds to an intuitive story. While, as discussed above, we find such intuitions useful, they do not tell us what the model predicts. For example, as discussed in the last section, it is common in economic theory to assume that an agent has subjective probability beliefs and maximizes expected utility given them. One interpretation of the model takes this literally: agents have probability beliefs and compute expected utilities with them to make decisions. Since people often find this difficult to do and rarely seem to explicitly follow such a procedure, some skeptics argue that this theory is simply wrong.

However, as noted, the predictions of positive decision theory are predictions of choice behavior. While the story's plausibility may affect our confidence in those predictions, it cannot refute or confirm them. The contribution of Savage (1954) was to identify the predictions of (and thus what would refute) the model by demonstrating that an agent whose behavior satisfies certain easily interpretable and testable properties would behave as if she has probability beliefs and computed expected utility.

One reason this is useful is that it can help us find and/or understand examples that show the limitations of the model. As Ellsberg (1961, page 646) put it, one effect of Savage's axioms is that they gave "a useful operational meaning to the proposition that people do *not* always assign, or act 'as though' they assigned, probabilities to uncertain events." In other words, the identification of the behavior that corresponds to subjective expected utility also identifies the behavior that does *not* correspond to the theory. Ellsberg then identified a set of circumstances under which Savage's theory is unlikely to predict well, and with the help of the axioms, pointed to the behavior which alternative models would need to address. Interestingly, Ellsberg (2001), footnote 2, page 244) observed that he was "considerably surprised" to realize the axiom he was questioning was Savage's sure–thing principle, the analog of the Anscombe–Aumann independence axiom. (See appendix for details.) Without knowing the axioms of subjective expected utility, one could not know what general property of subjective expected utility Ellsberg's example was calling into question and hence it would be much less clear how to understand its implications.

Another way to see the point is to compare the Anscombe–Aumann derivation of subjective expected utility to the Gilboa–Schmeidler multiple priors model (both of which are reviewed in the appendix). In the multiple priors model, the agent has a set of probability beliefs instead of a single belief. In evaluating a given course of action, he uses the "worst case" expected utility across the beliefs in his set. While one can easily compare the story of this model to the story of subjective expected utility, it is not immediately obvious how the behavior predicted by the two models differs. Gilboa and Schmeidler's representation theorem shows that the multiple priors model has two kinds of predictions that differ from subjective expected utility. The first difference is that the independence axiom applies only in certain situations. More specifically, the decision maker's ranking of acts do not change when mixing each with a third act only when the third act is independent of the state of the world and hence does not affect the level of ambiguity of the comparison. The second difference is an ambiguity–aversion property, saying that the decision maker weakly prefers randomizing over indifferent acts because this can reduce the variability of utilities across states and hence reduce ambiguity.

Relatedly, it may be easier to test behavioral implications directly rather than to estimate a general functional–form representation and see if the estimated function is in the class corresponding to the model. This is particularly true if we want to be sure to test the parts of the model we take seriously, not the parts we view as a convenient but inaccurate simplification. If we reject a functional form, it can be difficult to say whether we are rejecting the simplification or the essence of the model. If we directly test the key behavioral predictions instead, we don't have this problem.

The discussion so far only demonstrates the role of *necessity*, not *sufficiency*. A representation theorem proves that behavior satisfies certain properties if and only if it is the result of maximizing a utility function in a certain class. Thus these properties of behavior are necessary and sufficient for the representation. Typically, proving the behavioral properties are necessary (implied by the representation) is very easy — it is not unusual for this part of a proof to be omitted because it is trivial. On the other hand, the sufficiency part of the proof (showing that the behavioral properties imply the representation) is often quite difficult.

Yet the justifications above really only show value to the easy part. To see this, suppose we have a property which we know is necessary for the functional form and don't know whether it is sufficient. If we observe behavior inconsistent with this necessary property, then we reject the model. Similarly, if we understand the intuitive limitations of a necessary property, again, this may enable us to identify the limitations of a model. That is, recognizing that the property conflicting with Ellsberg's example is the sure–thing principle requires only knowing that the sure–thing principle is necessary.

The sufficiency part of a representation theorem is important for several reasons. First, obviously, if we don't know *all* of the implications of a model, we don't know if we are seeing behavior consistent with it. For example, if we only knew that expected utility requires that preferences be transitive and continuous but did not realize it requires the independence axiom, we might fail to realize that the Allais paradox behavior contradicts expected utility. Without knowing that the behavior we have identified is sufficient, we can never know whether there is a necessary property, perhaps a very undesirable one, that we have failed to identify.[13] For example, Harless and Camerer (1994) assessed a

[13]Even if we know the necessary and sufficient conditions for a representation, we may fail to recognize important behavioral implications of these conditions. There could easily be some subtle implication of a combination of two or more axioms that we do not notice simply by examining the axioms in isolation.

large set of experimental studies to determine which of a certain set of theories were consistent with each data point. Without a complete axiomatic characterization of each of these theories, this would not have been possible.

Similarly, we need the sufficiency result if we want to understand the relationships between models. To see the point, recall our discussion above of the multiple priors model. If we only knew that the properties discussed there were necessary but did not know they were sufficient, we would not know if there were other important differences between subjective expected utility and multiple priors.

The importance of knowing we have all the implications is particularly clear when the story of the model is potentially misleading about its predictions. For example, the multiple priors model seems to describe an extraordinarily pessimistic agent. Yet the axioms that characterize behavior in this model do not have this feature. The sufficiency theorem ensures that there is not some unrecognized pessimism requirement.

Similarly, the minimax regret model represents agents as deriving disutility from choosing an action and later realizing that a different action would have been better. Thus the story of the model seems to require the agent to know he will later learn which action would have been best. Yet some axiomatizations[14] identify properties of choice which do not require such ex post observation. See Schlag (2006) for further discussion.

Second, recall our discussion of developing a model of temptation in Section 2. We noted that one proposes a behavioral definition of the phenomenon of interest (temptation in the example), adds some convenient simplifying assumptions, and derives a model which corresponds to that behavior. This development is entirely about the sufficiency part of a representation proof. It is the sufficiency argument which enables us to say that the representation is the more tractable embodiment of our notion of temptation.

Finally, since we do not have the data to distinguish models perfectly, one way to enhance our trust in a model is to see that its results are robust. That is, if we find that a particular conclusion proven for one class of models also holds when we enlarge the set of allowed models, then we have more confidence in this conclusion. Of course, our confidence is enlarged more if the enlarged set of models is "bigger." Conversely, if we find that a "small" change in a model leads some conclusions to fail, then we are less confident in those conclusions. Clearly, to consider such issues, we need some way to measure the size of changes in models. Since changes in models typically involve non–nested changes in functional forms, there is rarely a parameter whose change measures the change in the model. Even when models are nested, it is hard to know how behaviorally significant

There are infinitely many ways to express the behavioral implications of any model and part of the art of axiomatization is bringing out the key implications in the clearest possible way. See Section 4.3.

[14]See, e.g., Milnor (1954), Stoye (2008), or Hayashi (2008).

a given change in a parameter really is. If we know the behavior in question, though, we can assess this. For example, suppose we learn that a conclusion which holds when preferences satisfy the independence axiom continues to hold when we relax independence to betweenness (Chew (1983), Dekel (1986)). That is, the property does not require both parallel and linear indifference curves, just the linearity. This gives us a clear sense of the degree of robustness being demonstrated.[15]

We are not arguing that representation theorems are the only tool for achieving a useful understanding of our models. For example, many interesting papers in behavioral economics (e.g., O'Donoghue and Rabin (1999)) are similar to papers in decision theory in that they note some observations that are difficult or impossible to explain with existing models, propose some alternative model which has an interesting interpretation in light of these observations, and then provide some means to understand the alternative.[16] In our view, this part of the behavioral economics literature takes a different approach than decision theory but with similar purposes and thus complements decision theory. The main differences between the approach of the decision theorist and of the behavioral economist are the completeness of the characterization and the choice environment studied to clarify the models predictions.

Starting with what we prefer about the decision theorist's approach, it generally takes the form of a complete characterization of the model, clarifying both its general predictions and the nature of the simplifications used. By contrast, behavioral economists typically omit any analog of the sufficiency results we argued are important. Indeed, many papers don't provide much in the way of general necessary conditions either, making it hard to see what the model says about individual choice outside the economic environment studied (and the stories one can tell about the representation).[17] In favor of the behavioral approach, it clarifies our understanding of the new model via an exploration of the model in some simple examples of economic environments. The decision theorist's analysis is typically carried out in a simpler but less structured context.

Thus we reject the argument that either approach is inherently superior to the other. Indeed, we think these approaches should be seen as complements, not substitutes. While we have generally taken one of these approaches in our own work, this reflects a combi-

[15]To be fair, it is sometimes possible to study robustness via the functional form as in Machina (1982).

[16]To be sure, many papers in behavioral economics are more focused on using a particular model of decision–making to understand some economic issue and are more analogous to applied work in other parts of economics than to decision theory. Of course, the boundary between these parts of behavioral economics — indeed, between "applied" and "pure" theory more broadly — is not always clear. For example, if we consider two applied models that are identical except for their decision theory component and one is more consistent with the data, this demonstration can be seen as some support for the decision theory model, even if the purpose of the paper was exploring the economic problem, not the decision theory components.

[17]See Spiegler (2008) for a similar argument in favor of a more systematic exploration of the implications of new models.

nation of historical accidents and comparative advantage (or comparative disadvantage as the case may be), not a view that only one approach is legitimate.

Relatedly, representation theorems are not always essential. While we see representation theorems as a valuable way to understand models better, a model can be worth exploring before it has been axiomatized. Also, sometimes the behavioral meaning of the model is sufficiently obvious that a representation theorem adds little. For example, in a model of utility of money, the assumption that $u' > 0$ has an obvious meaning without having to restate this in the language of preferences. More broadly, even if the behavioral meaning of the representation does need clarification, a representation theorem which is too convoluted to provide this is not helpful. Relatedly, much of behavioral economics offers a functional form instead of axioms. In some cases, the translation of this functional form into the behavior it predicts is sufficiently obvious that this does not seem unreasonable. In such cases, we can think of the functional form itself as the axiom, so a representation theorem is not needed. In other cases, it is very difficult to see what this "axiom" says about choice behavior, so we would prefer "real" axioms.

Not only is axiomatization not necessary for a model to be interesting, it is certainly *not* sufficient. A model that is intractable, complex, unintuitive, and makes lousy predictions may have an axiomatization, but we wouldn't want to use the model anyway.

3.2 Identification Theorems and Comparatives

One point which seems not well understood outside the decision theory community (and not globally inside it either) is the importance of identification. A model is supposed to be an intuitive and simplified description of some aspect of reality. The purpose of decision theory is to understand this model and the behavior it predicts/describes. One key to doing so is to determine the extent to which the objects in the representation can be pinned down from the agent's behavior and, in this sense, given behavioral meaning. If the objects in the representation cannot be given such meaning, then the model is, at best, loosely connected to what it is supposed to be describing and, at worst, misleading.

For example, consider state–dependent expected utility. This is a model where the utility function depends on what is consumed but also on the state of the world. It is well–known that we cannot identify subjective probabilities separately from utilities in such models. To see this, suppose there are two states, s_1 and s_2. The agent is represented via an evaluation of *acts*, functions which say what the agent gets in states s_1 and s_2. Suppose the agent's preferences over such acts can be represented by $(1/4)u(f(s_1), s_1) + (3/4)u(f(s_2), s_2)$. That is, if the agent chooses act f, he is interpreted as receiving outcome $f(s_1)$ in state s_1, generating utility $u(f(s_1), s_1)$ and analogously in state s_2. His subjective probabilities are that state s_1 has probability $1/4$, while s_2 has probability

3/4. It is easy to see that the choices predicted by this model would be the same as those predicted by the model where the agent maximizes

$$\frac{1}{2}\left[\frac{1}{2}u(f(s_1), s_1)\right] + \frac{1}{2}\left[\frac{3}{2}u(f(s_2), s_2)\right]$$

since this is the same function. Define $v(x, s_1) = (1/2)u(x, s_1)$ and $v(x, s_2) = (3/2)u(x, s_2)$. Then we are saying that the agent with utility function u with subjective probabilities $(1/4, 3/4)$ is indistinguishable from an agent with utility function v and subjective probabilities $(1/2, 1/2)$. Intuitively, if we see that the agent pays more attention to outcomes in one state than outcomes in the other, we can't tell if this is because his utility function is more sensitive in that state or because he considers that state more likely. Thus we cannot behaviorally identify one object separately from the other.

The usual approach for dealing with objects that cannot be behaviorally separated is to adopt a normalization for one object that yields (conditional) identification for the other. In the case of state–dependent utilities, we could normalize the probabilities to a uniform distribution or, equivalently, simply omit the probabilities altogether and work with the sum of the utilities. Note that this normalization means that the probabilities then do not have any meaning — it cannot be significant to say that the probabilities are uniform when we have to normalized them to be so.

Other normalizations leave some flexibility for both objects, so both can have some meaning. A very useful example of such a normalization is the state–independent model. This is a case where an additional behavioral property, called state independence by Anscombe and Aumann (1963), implies that we can normalize by taking the utility function u to be independent of the state. This is a normalization in the sense that there are state–dependent models that are equivalent. In particular, we have not proved that the utilities are state–independent; we have normalized to make them so.[18] However, this is a normalization which has a nice intuition and is useful.

To see how this can be useful, suppose we have a representation with a state independent u and we know that the agent is weakly risk averse. One can define this from the preferences but for brevity, we focus on an equivalent statement in terms of u — namely, that u is weakly concave.[19] A standard result for this kind of model implies the following. If we compare two acts f and g such that the distribution of outcomes under f (given the subjective probabilities) is a mean–preserving spread of the distribution under g, then it must be true that g is weakly preferred to f, denoted $g \succeq f$.

[18]This point is made forcefully by Karni (2007), though he takes a different perspective on the issue.

[19]The equivalent statement in terms of preferences that we have in mind is for the Anscombe–Aumann model where the set of "prizes" is $\mathbf{R}$, interpreted as money. We can write any given act f as (f_s, f_{-s}) where f_s is the lottery over money in state s and f_{-s} is the profile of such lotteries for other states. Suppose that for every s, $(f_s, f_{-s}) \preceq (\mathrm{E}f_s, f_{-s})$ for every nondegenerate lottery f_s.

Suppose, however, that we take a different normalization. To see the idea, suppose our initial state independent model has two states, a utility function of $u(x) = x$, and probability beliefs $(1/4, 3/4)$. Exactly the argument we gave earlier says that this agent's behavior would be the same as the behavior of an agent with state–dependent utility function $v(x, s_1) = (1/2)x$ and $v(x, s_2) = (3/2)x$ and probabilities $(1/2, 1/2)$. Consider the acts $f = (1, 1)$ and $g = (0, 2)$. Given the belief $(1/2, 1/2)$ that we attribute to the agent in this state–dependent representation, g is a mean–preserving spread of f — it has the same expected value but is not constant and hence is riskier. However, both representations agree on the conclusion that g has strictly higher expected utility than f. Thus the useful property of the state–independent model doesn't hold for an arbitrarily selected but equivalent state–dependent model. Hence the "normalization" of state independence strikes us as a natural one to adopt. As we argue later, even if there was some sense in which the "true" representation was not the state independent one, we'd still prefer this normalization.[20]

One of the key reasons we care about identification is that this is crucial to making comparatives possible. A comparative or comparative static is a result connecting a change in the representation to a change in the behavior it represents. To see the link to identification, note that identification results tell us the extent to which objects in the representation are behaviorally meaningful. Once we understand this, we can see how changes in those objects translate into changes in behavior.

Comparative statics enable us to obtain results connecting behavioral properties (e.g., attitudes to risk or ambiguity) to observables of interest (e.g., investment or insurance decisions). Much of the use of decision theory in economics is either directly of this form or an indirect use where one shows that a particular phenomenon implies a conclusion which does not hold in its absence.[21] A well–known example is the Arrow–Pratt (Arrow (1971), Pratt (1964)) measure of risk aversion.

One reason we think people sometimes overlook or underrate the importance of identification is that achieving identification is not costless. It often requires additional structure and/or additional assumptions. For example, if the domain a particular utility function is defined over is too small, then one may not have enough information to identify aspects of the utility function. Thus it is not unusual for decision theory papers to consider preferences over large and sometimes complex sets of objects. Relatedly, to develop the kind of structure needed for identification often requires stronger assumptions. Naturally, it seems better to have fewer assumptions and/or to focus on preferences over

[20]It is worth noting that one can have models where the Anscombe–Aumann state independence property fails and there is no normalization of utilities for which this result holds. In this sense, even though we can identify probabilities conditional on a normalization, the identification is not as useful in inherently state–dependent models.

[21]The indirect use is also a comparative static in that moving from the absence of the phenomenon to its presence is equivalent to moving from a "zero level" to a positive one.

simpler and presumably more realistic domains.

As we see it, the key question in whether the use of a particular domain is appropriate comes down to whether the objects in the domain can be thought of as reasonable approximations of some actual objects in the world that the agent may consider. For example, in the Savage model, the objects of choice are functions from an infinite set of states of the world to an infinite set of consequences. We don't see real people often making choices between such functions. On the other hand, we do see people viewing options they have in the form of "if x happens, then I'll end up with a, but if y happens, I'll get b." The functions in Savage are just a more formal version of this and so seem reasonable to us. In proposing a large space for use in identification, we think the burden is on the decision theorist to point to such a correspondence.

When multiple models can be used to represent the same behavior, we have a different kind of identification issue.[22] When two different models correspond to intersecting sets of behaviors, this should not be a surprise. Intuitively, a given model is like a particular explanation of the behavior it generates. It is unsurprising that at least some choice behaviors may have multiple explanations.

Furthermore, multiple explanations may be useful. Different explanations will suggest different intuitions, different questions to consider, different comparatives that might be useful. For example, we have a representation of temptation (Dekel and Lipman (2007)) where the agent is uncertain about how/whether he will be tempted in the future and believes his tempted future self will always give in to temptation. We show that the set of preferences for partial commitments that this model generates include all preferences generated by Gul and Pesendorfer's (2001) temptation model where the agent knows his future self's level of temptation and may expect his future self to exert self control.[23]

To see the idea, consider a dieting agent who can commit himself to a healthy dish h, commit himself to a fattening dish f, or leave open the choice between the two, denoted $\{h, f\}$. Then it seems natural that we might have $\{h\} \succ \{h, f\} \succ \{f\}$ where $\{h\}$ denotes committing to h and $\{f\}$ is defined similarly. In Gul and Pesendorfer, this preference is interpreted as saying that the agent knows that if he does not commit to one of the two dishes, then he will be tempted by f but will manage to consume h at the cost of exerting self–control, thus leaving him in between the two commitments in utility terms. In Dekel and Lipman, this preference is interpreted as saying that in the absence of commitment, the agent gives some probability to sticking with h and some probability of being tempted

[22] On the other hand, multiple models can be thought of as a special case of an unidentified parameter in a single model. We can always put two different models into a single class by defining the functional form of α times a function in one class plus $1 - \alpha$ times a function in the other. In cases where there is a function in each class representing the behavior, the α is unidentified.

[23] In a richer behavioral domain involving choices *under* partial commitment instead of choices *of* such commitments, the predictions of the models would differ.

away to f, thus ending up in between the other two options in terms of expected utility.

That different stories about temptation can overlap on at least some predictions should surely not surprise us. Yet these two stories are quite different and so suggest different applications or questions. For example, Gul and Pesendorfer focus on how we can identify the self–control costs and how higher costs would affect behavior, while Dekel and Lipman focus on the degree of uncertainty about future temptations. Thus we do not see this overlap as saying anything about which model is more useful or "better."

4 Implications for How to Do Decision Theory

The perspective on decision theory we have offered leads to a number of conclusions about how we should do decision theory, some of which are not entirely standard. In this section, we discuss several such issues.

4.1 The Role of Psychology

Some economists seem to reject the notion that psychology could be useful to economics and decision theory. In many cases, this view revolves around the distinction between studying the choices people make versus the way people make them. Many economists, notably Gul and Pesendorfer, have argued that we are only interested in the former. Others, such as Rubinstein (1998), suggest that the study of choice procedures may yield better understanding of choice behavior. Yet others, such as Camerer (2008), argue that while choice, not procedure, is the economist's traditional focus, this was an unfortunate compromise forced on the profession in an era when there was no way to observe how people make choices. These economists argue that with the advent of neuroeconomics, we can study the brain and its processing of choice problems.

We agree with Gul and Pesendorfer that what psychologists are interested in is not always the same as what economists are interested in, so that it is not necessarily useful to blend the two perspectives. On the other hand, in our view, the real question is whether psychology can help us generate insights that enable more or better predictions (about variables of interest to economists) or to do so more easily.[24]

To begin thinking about this issue, consider again the theory of risk aversion. As discussed above, the basic idea of taking the concavity of u to represent a property we

[24]We suspect that most economists agree with this point, though they may disagree in evaluating whether a given paper is successful at this.

call risk aversion seems extremely useful in economic theory. Yet this is very different from the way a psychologist might think about risk aversion. As discussed by Gul and Pesendorfer (2008), a psychologist (or a neuroeconomist) might focus more on cognitive or emotional factors, "fear responses," etc. Thus a psychologist might say that the curvature of u is unrelated to "true" risk aversion but is instead related to diminishing marginal utility of money. In our view the fact that the psychologist sees the economist's model as wrong hardly trumps all other considerations by itself. Instead, the key questions are whether the psychologist's model makes more or better predictions (about the variables of interest to economists), whether it is easier to work with, etc.

For a more concrete example, recall the second example in Section 3.2 where a preference could be represented using a state independent utility function and probability beliefs of $(1/4, 3/4)$ or a state dependent utility function and probability beliefs of $(1/2, 1/2)$. As we noted, the state independent representation has the nice feature that we can relate the curvature of the utility function to a theorem on mean–preserving spreads and we do not get this theorem for the equivalent state–dependent model. Suppose a psychologist could measure "utils" and told us that the person's utility function was different in the two states, that he genuinely valued consumption in one state three times as much as consumption in the other state, implying that the person's "true" probability beliefs were the ones in the state dependent representation, namely $(1/2, 1/2)$. We would argue that we should ignore the psychological data here since making our model more psychologically realistic comes at the cost of losing convenient tools. In short, psychological realism is not costless, so the real question is whether the costs exceed the benefits.

So what are the potential benefits of psychology? As we have argued throughout, all models take some observations or intuitions about reality, provide an intuitive explanation for them, and use that explanation to make additional predictions about observable variables of interest. With this in mind, suppose we have two equally tractable models, both consistent with all the data we have and which make different out of sample predictions. Suppose one is consistent with psychologist's views of human decision–making and the other is not. Which would we prefer to use to make predictions? We have already argued that one may well prefer to use the more "intuitive" or the more "plausible" model in such a situation, noting Kreps' (1990) comment that this is like a kind of data. The consistency with psychology can be another way of generating such plausibility.

In addition to affecting the plausibility of competing predictions, psychological considerations can suggest — indeed, have suggested — altogether new predictions. In some cases, such as models of ambiguity aversion, the models economists have developed make little direct use of psychology in motivating the formulation or assumptions. In other cases, though, the "story" of the model explicitly incorporates ideas from psychology.

For example, consider Laibson's (2001) notion of cue–based consumption. He observes that psychologists have identified situations where external influences act as cues for various desires or cravings. He formalizes this idea through the existence of exogenous, random observables which affect the agent's utility function and hence the agent's choices. He notes various new predictions that this model yields. For example, observing a billboard advertising cigarettes may lead someone trying to quit smoking to crave cigarettes and therefore to smoke. Note that Laibson's model is based on a utility function and, in that sense, is an economist's model. The point is that the restrictions on the model — the effect of the random observables — and the resulting predictions about choice are motivated by psychology and would not have been considered otherwise.[25]

4.2 Nonchoice Data

One controversial aspect of the interaction between economics and psychology is the use of brain imaging data or other physiological measures in economics. For example, Gul and Pesendorfer (2008) argue that while psychology may be a source of inspiration for economic models, "economic models are evaluated by their success at explaining economic phenomena. Since . . . brain imaging data are not economic phenomena, economists should not feel constrained to choose models that succeed as models of the brain." Why not? Because "the requirement that economic theories simultaneously account for economic data and brain imaging data places an unreasonable burden on economic theories." While the terminology is not obviously appropriate, we follow the literature in referring to this other data as *nonchoice data*.[26]

As we understand Gul and Pesendorfer, they are *not* arguing that nonchoice data can't help predict choices. Obviously, temperature would be useful in helping to predict the demand for ice cream. Nor are they arguing that we should completely disregard issues outside the realm of choice data. If we are trying to forecast ice cream demand, obviously, it would be useful to know enough about meteorology to know how to forecast temperatures. Instead, we take their argument to be that there is a separation between the "economic" part of a model and "subsidiary" parts of a model, be they meteorological or psychological. As economists, our interest should be in the "economic" part and only economic/choice data really bear on that part.

We disagree on two points. First, as argued earlier, the greater our confidence in even

[25] Gul and Pesendorfer (2008) note that psychologists may desire to break behavior into that which is related to cues and that which isn't in a way which is different from what would be most useful for economists. We see this as a valid concern, but essentially orthogonal to our point.

[26] The problem with this term is that it is not clear why an observation such as an eye movement or response time is less a "choice" than a consumption decision. The distinguishing characteristic of this data seems to be its nontraditional nature more than whether it is "chosen" or not. See Caplin (2008) for a similar critique.

auxiliary assumptions, the greater our confidence in the predictions of the overall model. Second, in a world of limited data, the nonchoice part of the model may be important for prediction or model selection.

To see these points, suppose we have data on the relationship between how the options are presented to subjects in written form, eye movements of the subjects as they examine this information, and final choices by the subjects.[27] It would be natural to consider a model where the eye movements are related to how agents process this information to reach a decision. We acknowledge Gul and Pesendorfer's view that most economists are likely to be primarily interested in the relationship between the form in which information is presented and the final choices, not the eye movements per se. However, if the predictions of our model about eye movements fits that data, this would give us greater confidence in the predictions about the mapping from presentation to choices. By analogy, while we would not reject a macroeconomic model which assumes expected utility because of evidence against expected utility, we would find the macro model more compelling the stronger is each piece of the model. Furthermore, if we have only limited data relating presentation to choices, the use of eye movement data could be very helpful in choosing among models. For example, a model which predicts that a certain aspect of the presentation is key would certainly come into question if eye movement data showed that agents never looked at that. If we have enough data on choices, we would discover the irrelevance of this aspect of presentation eventually, but in the real world of limited data, this information would be valuable.

Finally, while we would not argue that this is a standard that must be met, surely a model which provides a unified explanation of multiple sets of data is a better model. We agree with Gul and Pesendorfer that it is asking a lot, probably too much, to achieve such unity, but we do think unity is a desirable very–long–run aspiration.

4.3 What is a Good Axiomatization?

In the preceding sections, we commented extensively on what makes a good model or, in the context of decision theory, a good representation of behavior. In this section, we take a representation as given and ask what makes a good axiomatization for this representation. Many of the same criteria we use for judging a model are again critical. Recall that the purpose of a representation theorem is to add to our understanding of and ability to test the model of the agent. In particular, the representation theorem gives our statement of the predictions of the model about observable variables.

Hence the first thing a representation theorem must do is to identify the key behavior

[27] This example is similar to Arieli, Ben-Ami, and Rubinstein (2009).

which corresponds to the phenomenon being modeled, which sometimes requires a new domain of behavior. While this step is very basic in a sense, many of the most fundamental developments in decision theory stem from an insightful approach to it. A classic example in the theory of subjective expected utility is Ramsey's (1926) observation that if agents are subjective expected utility maximizers, then their preferences over bets will reflect the subjective probabilities (for the state–independent "normalization"). Specifically, if the agent prefers betting \$1 if A occurs and 0 otherwise to \$1 if B occurs and 0 otherwise, then the subjective probability of A is higher than that of B. This simple insight was a key step in the development of Savage's axiomatization.

As another example, Kreps' (1979) study of the demand for flexibility introduced a new domain and the key property on that domain. Kreps recognized that decisions about flexibility could be thought of in terms of choice of a *menu* — that is, a partial commitment regarding one's future choices. The key axiom is then the formal statement of a desire for flexibility, specifically, that if one menu is a subset of another, then the agent prefers the larger menu.[28] Gul and Pesendorfer (2001) recognized that temptation would induce the opposite — a demand for commitment instead of flexibility — and thus introduced the use of preferences over menus to study temptation. Their key behavioral property is that agents prefer smaller menus in certain situations. Recognizing that non–Bayesian updating induces an intertemporal inconsistency akin to temptation, Epstein (2006) extended the domain to menus over (Anscombe–Aumann) acts and used this richer domain to give a behavioral characterization of such updating.

Finally, we note that a creative development of a domain can itself be the key step. For example, prior to the work of Kreps and Porteus (1978) and Segal (1990), every model involving uncertainty and time implicitly assumed that only the probability distribution over the information received at the time of an action choice was relevant. In particular, as Kreps and Porteus emphasized, this assumes that the timing of the resolution of uncertainty is irrelevant — if an agent can't act till tomorrow afternoon regardless, then it would not matter to him whether the uncertainty he faces is resolved tonight or tomorrow morning. Similarly, as Segal emphasized, this assumes that a sequence of lotteries which determine the agent's consumption are equivalent to the agent to the overall lottery over consumption they imply. Kreps–Porteus and Segal introduce models where one can make these distinctions and hence illustrate the behavioral impact they have.[29]

Identifying the key behavior and the domain is the most essential step, but also the step which is closest to an art. Thus we find it difficult to tell the reader how to do it or how to distinguish "good" choices from "bad." In the rest of this section, we discuss a

[28]Dekel, Lipman, and Rustichini (2001) extend the domain considered by Kreps to allow for menus of lotteries. In contrast to the other examples, this extension does not enable studying a novel form of behavior but instead serves the role of enabling better identification of the parameters in Kreps' model.

[29]The Kreps–Porteus model underlies the framework developed by Epstein and Zin (1989) to study asset demands where risk aversion and consumption smoothing are identified separately.

number of simpler issues where it is much easier to offer some advice.

Since axiomatizations are supposed to state the model's predictions for observables, the first guideline is easy: axioms should be about variables of interest which are at least potentially observable. Unfortunately, it is not always obvious which variables are observable. For example, consider Caplin and Leahy (2001), who consider preferences over lotteries over "psychological states" with a function which relates lotteries over physical outcomes to lotteries over psychological outcomes. It seems to us that psychological outcomes are not directly observed and so the function relating these to physical outcomes cannot be identified. On the other hand, psychologists and the authors may well disagree. Perhaps measurements of psychological outcomes can be examples of the kind of nonchoice data we suggested earlier could be valuable. By contrast, Caplin and Dean (2009) study interesting nonchoice variables that are naturally taken to be observable.

The second point is equally immediate: An axiomatization which does not say more than the representation is not helpful. In some cases, the behavioral meaning of the representation is obvious. We noted earlier that we don't need an axiomatization to tell us that $u' > 0$ represents "more is better." In other cases, the behavioral meaning is not obvious, but the key axioms are little more than a restatement of the representation. For example, we mentioned our 2007 representation of temptation where an agent has some probability beliefs regarding whether his future self will succumb to temptation. Formally, this model represents the agent's evaluation of a menu by the expected utility of a certain probability distribution over items in the menu. We could have axiomatized this model by making our main axiom the statement that for any menu, there is a lottery over the items in the menu such that the agent is indifferent between the menu and the lottery. This is indeed an implication of the model. On the other hand, this axiom just says the representation in different words and so doesn't tell us anything new.

Also, obviously, axioms should be simple and clearly interpretable. Otherwise, again, the axioms have added nothing to our understanding. While it is difficult to define simplicity, we note a few obvious guidelines. First, it is generally better to state axioms in terms of the preferences, not a series of relations derived from the preference. For example, a key in Savage's representation theorem is the "more–likely–than" relation, which is constructed from the preference relation. Yet Savage states his axioms in terms of the preference, not in terms of the derived relation, since the preference is what we are making predictions about. Second, axioms involving existential quantifiers are often too complex to be interpreted in this sense. It is hard to get an intuition about an axiom that says that an object exists with certain properties.[30] Finally, as noted in Section 2, it is desirable to have axioms that are familiar, either because they are standard or because they are intuitive variations on standard axioms. Obviously, it is easier to understand the familiar. Also, this approach makes it easier to compare the new model to existing

[30]Mixture continuity is a notable (though untestable) exception.

work and see what it adds of value to our understanding.

Next, we turn to issues regarding the set of axioms as a whole. Just as individual axioms should be as interpretable as possible, it is important to convey the meaning of the group as well. An obvious point in this regard is that the clarity of a representation theorem typically decreases rapidly with the number of axioms. (Also, we suspect that if a model really requires a large number of axioms, there are probably many issues being combined that should be treated separately.) Second, axioms stated in a conditional form (e.g., if a is preferred to b and b is preferred to c, then a is preferred to c) are often crucial for clarity. By contrast, if all axioms take the form of universal statements (e.g., more is better), the individual axioms may be clear but, except in trivial cases, we cannot imagine avoiding a large and complex set of such axioms. Intuitively, conditional axioms allow us to make many inferences from a few observations. Without such axioms, we have to effectively list all preference statements.

A more difficult principle to adhere to is to state the set of axioms in a way which enables the reader to see the role of each piece in generating the representation. This makes the analysis more informative. However, there is typically a great deal of interaction among axioms, making this hard to do. Indeed, if there were no interaction, arguably, this would indicate that this representation is mixing multiple issues together. An additional tension is that theorems with interactions among the axioms are often more elegant and mathematically interesting.

An easier but still nontrivial principle is to clearly separate the axioms intended to be relatively realistic and/or statements of the main focus from axioms intended primarily as useful simplifications. While this sounds easy, in practice, "innocuous" simplifications can have unexpected substantive implications. For example, in Gul and Pesendorfer's (2001) paper on temptation, it is natural to read their "behavioral" axiom as the property they call set betweenness and to view the other axioms as useful but inessential simplifications. However, they show in Gul and Pesendorfer (2005) that independence buys more than the simplification of a linear structure. Indeed, Dekel, Lipman, and Rustichini (2009) give an example of intuitive temptation behavior that is ruled out in their model not by set betweenness, but by the combination of set betweenness and independence.[31]

Finally, we note that weaker axioms are not *per se* better axioms. If one uses weak assumptions to get a weak result, it's not clear the exercise is useful. On the other hand, if one uses weak assumptions to get a strong result, this is surely either because the result is weaker or the assumptions stronger than is apparent.[32]

[31] For related discussions, see also Fudenberg and Levine (2006) and Noor and Takeoka (2009).

[32] On the other hand, if the misperception of the strength of the assumption or conclusion is widespread, such a theorem could be very instructive.

5 Conclusion: Welfare Economics

We conclude this paper with a few words on a topic we are reluctant to omit but have little to say about. When it comes to prescriptive analysis, prediction of choices is insufficient. We must also have some principled way to identify the welfare implications of choices. Traditionally, economists have assumed that choice reveals utility or welfare directly. That is, if an agent chooses A over B, this implies he is better off consuming A than consuming B. Of course, as soon as we consider models where agents have difficulties in making good decisions and hence may make mistakes, such an assumption seems rather inappropriate.

There seem to be two possible responses to this issue. First, one can reject the idea that economics has anything to say about welfare directly. For example, Gul and Pesendorfer (2008) argue that what is typically called welfare economics should be seen as predictive, not prescriptive. They suggest that results about whether a particular institution leads to Pareto optimal outcomes should not be seen as determining whether the institution is socially valuable. Instead, the statement that there is some other institution which leads to Pareto preferred outcomes, they argue, should be interpreted as raising the question why society does not switch to such an alternative since we predict that each agent would choose to switch if possible.

Alternatively, one can propose a new set of assumptions to identify welfare consequences. Ultimately, though, we cannot imagine how one could ever *prove* what is best for a particular person's welfare. Thus the goal must be seen as finding a hypothesis that we have more confidence in than revealed preference. There are some interesting ideas in the literature, but we have nothing to add. We refer the interested reader to Bernheim and Rangel (2009), Chambers and Hayashi (2008), Koszegi and Rabin (2008), Noor (2009), and Rubinstein and Salant (2009), among other work on this issue.

A Appendix

Subjective probability was originally studied by Savage (1954) but it is very common to work instead with the more tractable formulation of Anscombe and Aumann (1963). In both approaches, there is a set S of states of the world and a set $\mathcal{L}$ of outcomes or consequences. An *act* in either model is a function $f : S \to \mathcal{L}$. The preference relation $\succeq$ is defined over the set of such acts, say F. The difference between the two models is that Anscombe–Aumann interpret $\mathcal{L}$ as a space of lotteries (probability measures) over some other set. Under the natural definition of convex combinations of lotteries, $\mathcal{L}$ is convex, so this enables us to treat the set of acts F as convex where we define the convex combination of $f, g \in F$ pointwise. That is, $\lambda f + (1-\lambda)g$ is defined to be that act h such that $h(s) = \lambda f(s) + (1-\lambda)g(s)$ for every $s \in S$. This convexity property greatly simplifies the analysis. For example, Anscombe and Aumann are able to obtain their representation result when S is finite, which cannot be done under the Savage axioms.

The axioms of Anscombe–Aumann are as follows. First, $\succeq$ is a weak order (complete and transitive). Second, it is required to satisfy a continuity axiom: if $f \succ g \succ h$, then there exists $\alpha, \beta \in (0,1)$ such that $\alpha f + (1-\alpha)h \succ g \succ \beta f + (1-\beta)h$. Third, $\succeq$ is assumed to satisfy the independence axiom. While this has many (equivalent) forms, one is that if $f \succ g$, then $\alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$ for all acts h and all $\alpha \in (0,1]$. Finally, for simplicity, there is a nontriviality axiom (there exists f and g with $f \not\sim g$).

The last axiom used by Anscombe–Aumann is *state independence* which is interpreted as saying that the agent ranks consequences the same in every state. Most recent treatments use an equivalent axiom due to Schmeidler (1989) called *monotonicity*. This axiom says that if $f(s) \succeq g(s)$ for all s, then $f \succeq g$ where, for any act h, we use $h(s)$ to refer to the act that gives outcome $h(s)$ in every state. To see the link between monotonicity and state independence, note that monotonicity implicitly assumes that $f(s) \succeq g(s)$ means that $f(s)$ is better in every state than $g(s)$ and thus embodies (indirectly) an assumption that the agent's ranking of consequences does not vary across states.

Anscombe and Aumann show that a preference satisfies these axioms if and only if there exists $u : \mathcal{L} \to \mathbf{R}$ and a probability measure p over S such that the preference is represented by the function $\sum_{s \in S} p(s)u(f(s))$ where u is affine in the sense that $u(\lambda L + (1-\lambda)L') = \lambda u(L) + (1-\lambda)u(L')$ for all $L, L' \in \mathcal{L}$. Moreover, u is unique up to a positive affine transformation and p is unique. Finally, if monotonicity is dropped, then there is $u : \mathcal{L} \times S \to \mathbf{R}$ such that $\succeq$ is represented by $\sum_{s \in S} u(f(s), s)$.

Gilboa and Schmeidler (1989) generalize this model to incorporate ambiguity aversion. Their axioms include all of the Anscombe–Aumann axioms except that they weaken independence to what they call *certainty independence* and add a new axiom called *ambiguity aversion*. Certainty independence says that the conclusion of independence is required *only* when the act h is a constant act — that is, $h(s) = h(s')$ for all $s, s' \in S$. Intuitively, this ensures that the mixing operation (that is, taking convex combinations with another act) does not change the level of ambiguity or uncertainty involved. To see the idea, suppose $f \succ g$ and that h exactly offsets all the risk in g in the sense that h yields good outcomes when g yields bad ones and conversely. Then $(1/2)g + (1/2)h$ may be "unambiguous" in that the agent no longer cares which state is realized. However, $(1/2)f + (1/2)h$ may still be quite "ambiguous," so that the agent might have $(1/2)g + (1/2)h \succ (1/2)f + (1/2)h$, violating independence. Note, though, that it is crucial that h offsets the risks in g — a constant act could never do this. Thus independence with respect to constant acts seems to be a reasonable property to maintain.

The ambiguity aversion axiom says that if $f \sim g$, then $\alpha f + (1-\alpha)g \succeq f$. The idea is the same as above: by mixing f and g, we take some of the variation out, so the result is "less ambiguous" and hence better than the original acts. Gilboa and Schmeidler prove that a preference $\succeq$ satisfies their axioms if and only if there is a utility function $u : X \to \mathbf{R}$ and a convex *set* of probability measures P such that the preference is represented by the function $\min_{p \in C} \sum_{s \in S} p(s)u(f(s))$. Again, u is affine and unique up to a positive affine transformation. P is unique.

The last model we explain takes a different approach. Gul and Pesendorfer (2001) give a model of preferences over sets of lotteries, which we refer to as *menus*, which can be interpreted as a model of temptation. A menu, which we denote x, is a closed nonempty subset of $\mathcal{L}$. X is the set of all menus and the preference $\succeq$ is defined over X. Their first three axioms again look similar to those of Anscombe–Aumann. First, $\succeq$ is a weak order. Second, it is continuous, though in a different form than mixture continuity. Specifically, their definition says that $\{x \in X \mid x \succeq y\}$ and $\{x \in X \mid x \preceq y\}$ is closed (in the Hausdorff topology) for all $y \in X$. Their third axiom is the independence axiom extended to this setting. Specifically, given menus x and y, let

$$\lambda x + (1-\lambda)y = \{\beta \in \mathcal{L} \mid \beta = \lambda\alpha + (1-\lambda)\alpha', \text{ for some } \alpha \in x, \alpha' \in y\}.$$

Then the independence axiom says $x \succ y$ implies $\alpha x + (1-\alpha)z \succ \alpha y + (1-\alpha)z$ for all menus z and all $\alpha \in (0,1]$. Thus it is the same axiom as before with this definition of what a convex combination is. Their last axiom is called *set betweenness*. It says that $x \succeq y$ implies $x \succeq x \cup y \succeq y$.

They show that a preference $\succeq$ satisfies these properties if and only if there are affine functions u and v such that the preference is represented by the function

$$\max_{\alpha \in x}[u(\alpha) + v(\alpha)] - \max_{\alpha \in x} v(\alpha).$$

References

[1] Allais, Maurice, "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'École Américaine," *Econometrica*, **21**, October 1953, 503–546.

[2] Anscombe, F. J., and Robert J. Aumann, "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, **34**, March 1963, 199–205.

[3] Arieli, Amos, Yaniv Ben-Ami, and Ariel Rubinstein, "Fairness Motivations and Procedures of Choice between Lotteries as Revealed through Eye Movements," Tel Aviv University, unpublished, 2009.

[4] Arrow, Kenneth, *Essays in the Theory of Risk Bearing*, Chicago: Markham, 1971.

[5] Bernheim, Douglas, and Antonio Rangel, "Beyond Revealed Preference: Choice–Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, **124**, February 2009, 51–104.

[6] Bewley, Truman, "Knightian Decision Theory: Part I," *Decisions in Economics and Finance*, **25**, November 2002, 79–110.

[7] Camerer, Colin, "The Case for Mindful Economics," in Caplin and Schotter (2008), 43–69.

[8] Caplin, Andrew, "Economic Theory and Psychological Data: Bridging the Divide," in Caplin and Schotter (2008), 336–371.

[9] Caplin, Andrew, and Andrew Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, Oxford: Oxford University Press, 2008.

[10] Caplin, Andrew, and Mark Dean, "Choice Anomalies, Search, and Revealed Preference," New York University, unpublished, 2009.

[11] Caplin, Andrew, and John Leahy, "Psychological Expected Utility Theory And Anticipatory Feelings," *Quarterly Journal of Economics*, **116**, February 2001, 55–79.

[12] Chambers, Christopher, and Takashi Hayashi, "Choice and Individual Welfare," Cal Tech, unpublished, 2008.

[13] Chew, Soo Hong, "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, **51**, July 1983, 1065–1092.

[14] Dekel, Eddie, "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, **40**, December 1986, 304–318.

[15] Dekel, Eddie, and Barton L. Lipman, "Self–Control and Random Strotz Representations," Boston University, unpublished, 2007.

[16] Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Representing Preferences with a Unique Subjective State Space," *Econometrica*, **69**, July 2001, 891–934.

[17] Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Temptation–Driven Preferences," *Review of Economic Studies*, **76**, July 2009, 937–971.

[18] Dubra, Joan, Fabio Maccheroni, and Efe Ok, "Expected Utility Theory without the Completeness Axiom," *Journal of Economic Theory*, **115**, March 2004, 118–133.

[19] Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, **75**, November 1961, 643–669.

[20] Ellsberg, Daniel, *Risk, Ambiguity and Decision*, New York: Garland, 2001.

[21] Eliaz, Kfir, and Ran Spiegler, "Consideration Sets and Competitive Marketing," University College London, unpublished, 2008.

[22] Epstein, Larry, "An Axiomatic Model of Non–Bayesian Updating," *Review of Economic Studies*, **73**, April 2006, 413–436.

[23] Epstein, Larry, and Stanley Zin, "Substitution, Risk Aversion and the Temporal Behaviour of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, **57**, July 1989, 937–969.

[24] Fudenberg, Drew, "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature*, **44**, September 2006, 694–711.

[25] Fudenberg, Drew, and David Levine, "A Dual Self Model of Impulse Control," *American Economic Review*, **96**, December 2006, 1449–1476.

[26] Gilboa, Itzhak, and David Schmeidler, "Maxmin Expected Utility with Non–Unique Prior," *Journal of Mathematical Economics*, **18**, #2 1989, 141–153.

[27] Gilboa, Itzhak, *Rational Choice*, Cambridge, MA: MIT Press, 2009.

[28] Gollier, Christian, *The Economics of Risk and Time*, Cambridge, MA: MIT Press, 2001.

[29] Gul, Faruk, and Wolfgang Pesendorfer, "Temptation and Self–Control," *Econometrica*, **69**, November 2001, 1403–1435.

[30] Gul, Faruk, and Wolfgang Pesendorfer, "A Simple Theory of Temptation and Self–Control," Princeton University, unpublished, 2005.

[31] Gul, Faruk, and Wolfgang Pesendorfer, "The Case for Mindless Economics," in Caplin and Schotter (2008), 3–42.

[32] Hayashi, Takahashi, "Regret Aversion and Opportunity–Dependence," *Journal of Economic Theory*, **139**, March 2008, 242–268.

[33] Harless, David, and Colin Camerer, "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, **62**, November 1994, 1251–1289.

[34] Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, **47**, March 1979, 263–291.

[35] Kalai, Gil, Ariel Rubinstein, and Ran Spiegler, "Rationalizing Choice Functions By Multiple Rationales," *Econometrica*, **70**, November 2002, 2481–2488.

[36] Karni, Edi, "Foundations of Bayesian Theory," *Journal of Economic Theory*, **132**, January 2007, 167–188.

[37] Koszegi, Botond, and Matthew Rabin, "Revealed Mistakes and Revealed Preferences," in Caplin and Schotter (2008), 193–209.

[38] Kreps, David, "A Representation Theorem for 'Preference for Flexibility,'" *Econometrica*, **47**, May 1979, 565–576.

[39] Kreps, David, *Notes on the Theory of Choice*, Boulder: Westview Press, 1988.

[40] Kreps, David, *A Course in Microeconomic Theory*, Princeton: Princeton University Press, 1990.

[41] Kreps, David, and E. L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, **46**, January 1978, 185–200.

[42] Laibson, David, "A Cue–Theory of Consumption," *Quarterly Journal of Economics*, **116**, February 2001, 81–119.

[43] Machina, Mark, "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, **50**, March 1982, 277–323.

[44] Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, **74**, November 2006, 1447–1498.

[45] Manzini, Paola, and Marco Mariotti, "Sequentially Rationalizable Choice," *American Economic Review*, **97**, December 2007, 1824–1839.

[46] Milnor, J., "Games Against Nature," in R. M. Thrall, C. H. Coombs, and R. L. Davis, eds., *Decision Processes*, New York: Wiley, 1954.

[47] Noor, Jawwad, "Subjective Welfare," Boston University, unpublished, 2009.

[48] Noor, Jawwad, and Norio Takeoka, "Uphill Self Control," Boston University, unpublished, 2009.

[49] O'Donoghue, Ted, and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, **89**, March 1999, 103–124.

[50] Pratt, John, "Risk Aversion in the Small and in the Large," *Econometrica*, **32**, January–April 1964, 122–136.

[51] Ramsey, Frank P., "Truth and Probability," 1926, in Ramsey, F., *The Foundations of Mathematics and other Logical Essays*, ed. by R. B. Braithwaite,London: Kegan, Paul, Trench, Trubner & Co., 1931.

[52] Rubinstein, Ariel, *Modeling Bounded Rationality*, Cambridge, MA: MIT Press, 1998.

[53] Rubinstein, Ariel, and Yuval Salant, "Eliciting Welfare Preferences from Behavioral Datasets," Tel Aviv University, unpublished, 2009.

[54] Sarver, Todd, "Anticipating Regret: Why Fewer Options May Be Better," *Econometrica*, **76**, March 2008, 263–305.

[55] Savage, Leonard J., *The Foundations of Statistics*, New York: Dover Publications, 1954.

[56] Schlag, Karl, "Why Minmax Regret," UPF, unpublished, 2006.

[57] Schmeidler, David, "Subjective Probability and Expected Utility without Additivity," *Econometrica*, **57**, May 1989, 571–587.

[58] Segal, Uzi, "Two–Stage Lotteries without the Reduction Axiom," *Econometrica*, **58**, March 1990, 349–377.

[59] Simon, Herbert, *Models of Bounded Rationality, Volume 2*, Cambridge, MA: MIT Press, 1982.

[60] Siniscalchi, Marciano, "Vector Expected Utility and Attitudes toward Variation," *Econometrica*, **77**, May 2009, 801–855.

[61] Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, **87**, August 1973, 355–374.

[62] Spiegler, Ran, "On Two Points of View regarding Revealed Preferences and Behavioral Economics," in Caplin and Schotter (2008), 95–115.

[63] Stoye, Joerg, "Axioms for Minmax Regret Choice Correspondences," New York University, unpublished, 2008.

[64] Wakker, Peter, "Annotated References on Decisions and Uncertainty," Erasmus University, unpublished, 2009.