

MORE ON RANDOM UTILITY MODELS WITH BOUNDED AMBIGUITY

Charles F. Manski
Department of Economics and Institute for Policy Research
Northwestern University

March 2017, forthcoming in *Theory and Decision*

Abstract

Econometric analysis of discrete choice has made considerable use of random utility models to interpret observed choice behavior. Much empirical research concerns choice problems in which persons act with partial knowledge of the utilities of the feasible actions. Economists use random expected utility models to analyze such choice problems. A common practice is to specify fully the expectations that persons hold, in which case choice analysis reduces to inference on preferences alone. However, the expectations assumptions made in empirical research rarely have much foundation. To enable more credible research, this paper considers inference when one specifies a set of expectations that decision makers may plausibly hold. I first pose the idea in abstraction and then specialize to binary response with linear utilities, where the analysis is particularly straightforward. I initially assume that decision makers possess unique subjective probability distributions on the states of nature and make choices that maximize expected utility. I later consider the possibility that persons place only partial probabilistic structure on the states of nature and make undominated choices.

1. Introduction

Econometric analysis of discrete choice has made considerable use of random utility models (RUMs) to interpret observed choice behavior (McFadden, 1974, 1981). Much empirical research concerns choice problems in which persons act with partial knowledge of the utilities of the feasible actions. Economists have used random expected utility models to analyze such choice problems. A common practice has been to specify fully the expectations that persons hold, in which case choice analysis reduces to inference on preferences alone.

Unfortunately, the expectations assumptions made in empirical research often have little foundation, diminishing the credibility of findings. Consider, for example, analysis of travel mode choice for the journey between home and work, one of the earliest applications of random utility models and still an important subject of empirical research (e.g., Warner, 1962; Domencich and McFadden, 1974). The canonical mode-choice model supposes that, each day, a worker chooses between two alternatives, travel by automobile and by public transit. The utility of each mode depends on its travel cost and travel time.

Empirical researchers have commonly used models of traffic flow on transportation networks to predict the travel times that particular workers would experience by each mode. Researchers have also assumed that these predicted travel times agree with the travel times that workers perceive when they make their mode choices. The accuracy with which researcher-predicted travel times measure travel time expectations is questionable. Transportation network models cannot precisely emulate the circumstances of individual travelers. Moreover, workers typically are uncertain how long the journey will take by each mode. Travel times may vary from day to day due to unforeseen variation in traffic volume and the possibility of accidents.

RUMs with incorrectly predicted expectations of travel times or other attributes of alternatives are misspecified. The generic result for discrete choice analysis is inconsistent parameter estimation, the specifics depending on the case. To enhance the credibility of econometric analysis, I have recommended survey measurement of the expectations that decision makers hold (Manski, 2004). A small but growing body of

empirical research proceeds in this manner, measuring the probabilistic expectations of sampled decision makers and combining choice and expectations data to estimate RUMs. See Lochner (2007), Delavande (2008), van der Klaauw and Wolpin (2008), van der Klaauw (2012), Zafar (2013), Wiswall and Zafar (2015), and Giustinelli (2016).

What can one do in the absence of expectations data? In this case, one can still study how inference depends on the expectations assumptions imposed. Manski (2010) considered inference when one specifies a set of expectations that decision makers may plausibly hold. I first posed the idea in abstraction and then specialized to binary response with linear utilities, where the analysis is straightforward. I mainly assumed that decision makers possess unique subjective probability distributions on the states of nature and make choices that maximize expected utility. I briefly considered the possibility that persons place only partial probabilistic structure on the states of nature and make choices in some manner that uses the available structure.

I referred to the models of choice behavior developed in Manski (2010) as random utility models with bounded ambiguity (RUMBAs). Ambiguity may be observational in that the researcher does not observe the expectations that decision makers hold. Ambiguity may be behavioral in the sense of Ellsberg (1961); that is, the persons under study may not possess complete probabilistic expectations. The adjective “bounded” refers to the fact that meaningful inference on the population distribution of preferences is possible only if the researcher possesses sufficient a priori knowledge of their expectations.

This paper revises and updates the presentation of Manski (2010). Sections 2 and 3 consider observational and behavioral ambiguity respectively.

2. Observational Ambiguity

2.1. Generalities

Let J be a population of decision makers, each of whom chooses an action from a finite choice set C . The standard RUM assumes that person j associates utilities $(u_{jc}, c \in C)$ with the feasible actions and chooses one that maximizes utility. The inferential problem is to learn the distribution of preferences from observation of the choices and covariates of a random sample of decision makers. Let X denote the feasible values of the observable covariates. Assume that the distribution of preferences is continuous and has the form F_θ , where θ belongs to a specified parameter space Θ . Then the equations

$$(1) \quad P(c|x) = F_\theta(u_c \geq u_d, d \in C|x), \quad (c, x) \in C \times X$$

relate the choice probabilities $P(c|x)$ to the distribution of preferences F_θ . The identification region for θ is the set of parameter values that satisfy equations (1). The usual practice is to restrict the parameter space enough to point-identify θ and, hence, fully reveal the distribution of preferences. However, Manski (2007) considers a broad class of problems in which the distribution of preferences is partially identified and Manski (2014) shows how the analysis may be used to study labor supply under alternative income tax schedules.

Most empirical research today concerns choice problems in which persons act with partial knowledge of the utilities of the feasible actions. Economists use random expected utility models to analyze such choice problems. Let Γ be a specified set of states of nature and, for $\gamma \in \Gamma$, let $u_{jc\gamma}$ be the utility of action c to person j in state of nature γ . Suppose that, at the time of decision making, person j does not know what state of nature will be realized. Researchers routinely assume that person j places a subjective probability distribution on Γ , say Q_j , and chooses an action that maximizes expected utility. They assume that the joint distribution of

preferences and expectations is continuous and has the form G_θ , where θ belongs to a specified parameter space Θ . Then the equations

$$(2) \quad P(c|x) = G_\theta(\int u_{c\gamma} dQ \geq \int u_{d\gamma} dQ, d \in C|x), \quad (c, x) \in C \times X$$

relate $P(c|x)$ to G_θ . A common practice is to specify fully the expectations that persons hold, in which case the task of choice analysis reduces to inference on preferences alone.

The idea developed in Manski (2010) is to specify a set of expectations that decision makers may plausibly hold, rather than make assume they hold particular expectations. Empirical research specifying a set of plausible expectations typically cannot point-identify the population distribution of preferences, but it can yield more credible partial-identification findings. It can also make plain the extent to which conventional point estimates rest on untenable expectations assumptions. See Barseghyan, Molinari, and Teitelbaum (2016) for other work of this type.

I assume that the researcher and decision makers agree on the set Γ of feasible states of nature. Let Π_j denote the set of subjective distributions on Γ that the researcher deems plausible for person j to hold. I assume that the researcher is correct in thinking that the set Π_j contains person j 's expectations. I place no cross-person restrictions on expectations; hence, the members of the population can collectively hold any expectations in the Cartesian-Product set $\{\times_{j \in J} \Pi_j\}$.

Suppose that person j chooses action c . This action maximizes expected utility under the expectations Q_j . The researcher knows that $Q_j \in \Pi_j$. Hence,

$$(3a) \quad j \text{ chooses } c \Rightarrow \exists \pi \in \Pi_j \text{ s. t. } \int u_{j c \gamma} d\pi \geq \int u_{j d \gamma} d\pi, \forall d \in C.$$

Suppose that person j does not choose action c . Then there exists another action that yields expected utility at

least as large as c under Q_j . Hence,

$$(3b) \quad j \text{ does not choose } c \Rightarrow \exists \pi \in \Pi_j \text{ and } d \neq c \text{ s. t. } \int u_{jd} d\pi \geq \int u_{jc} d\pi.$$

This logical relationship is equivalent to its contrapositive

$$(3b') \quad \int u_{jc} d\pi > \int u_{jd} d\pi, \forall (d, \pi) \in C \times \Pi_j \text{ s. t. } d \neq c \Rightarrow j \text{ chooses } c.$$

Aggregating across the population, (3a) and (3b') imply these inequalities relating choice probabilities to the distribution of preferences and expectations:

$$(4) \quad G_\theta[\int u_{c\gamma} d\pi > \int u_{d\gamma} d\pi, \forall (d, \pi) \in C \times \Pi \text{ s. t. } d \neq c | x] \leq P(c|x) \\ \leq G_\theta[\exists \pi \in \Pi \text{ s. t. } \int u_{c\gamma} d\pi \geq \int u_{d\gamma} d\pi, \forall d \in C|x], (c, x) \in C \times X.$$

These inequalities provide the basis for inference on θ . The identification region is the set of parameter values that satisfy (4).

2.2. Binary Choice with Linear Utilities

Inequalities (4) describe the inferential problem in generality, but they are too abstract to communicate much. Hence, I now consider the special case of binary choice with linear utilities.

Assume that each member of the population must choose between two actions, labeled 0 and 1. The utility of action c to person j is

$$(5) \quad u_{jc\gamma} = z_{jc}\beta + \alpha y_{jc\gamma} + \varepsilon_{jc}.$$

Here z_{jc} is a K -vector, $y_{jc\gamma}$ and ε_{jc} are scalar, and (β, α) are corresponding parameters. The absence of a γ -subscript on $(z_{jc}, \varepsilon_{jc})$ indicates that the person knows these quantities at the time of decision making. The presence of a γ -subscript on $y_{jc\gamma}$ indicates that this quantity depends on the unknown state of nature. For example, when modeling choice of travel mode, the decision maker may know the travel cost (z) but not the travel time (y) of each alternative.

Assume that the person chooses an action that maximizes expected utility. The linearity of utility in y implies that expected utility varies only with the mean of y , not with its entire distribution. Specifically, let $z_j \equiv z_{j1} - z_{j0}$, $y_{j\gamma} \equiv y_{j1\gamma} - y_{j0\gamma}$, $v_j \equiv \int y_{j\gamma} dQ_j$, and $\varepsilon_j \equiv \varepsilon_{j1} - \varepsilon_{j0}$. Let c_j be the action chosen by person j . Then the decision rule is

$$(6) \quad c_j = 1[z_j\beta + \alpha v_j + \varepsilon_j > 0].$$

Suppose that a researcher draws a random sample of N members of the population. For each sample member $j = 1, \dots, N$, the researcher observes (a_j, z_j) but not (v_j, ε_j) . This would be a standard problem of binary choice analysis if the researcher were to observe v_j . In the absence of data on expectations, the prevailing practice has been to assume that v_j takes some particular value and proceed with standard binary choice analysis. Here I assume instead that $Q_j \in \Pi_j$. I also assume that the sign of α is known; for convenience, the discussion below assumes that $\alpha \geq 0$.

Let $v_{j0} \equiv \inf(\int y_{j\gamma} d\pi, \pi \in \Pi_j)$ and $v_{j1} \equiv \sup(\int y_{j\gamma} d\pi, \pi \in \Pi_j)$. Given (5), (6), and $\alpha \geq 0$, the logical relationships in (3) reduce to

$$(7a) \quad j \text{ chooses } 1 \Rightarrow z_j\beta + \alpha v_{j1} + \varepsilon_j \geq 0,$$

$$(7b) \quad z_j \beta + \alpha v_{j0} + \varepsilon_j > 0 \Rightarrow j \text{ chooses } 1.$$

Aggregating across the population, (7a) and (7b) yield these inequalities relating choice probabilities to the distribution of preferences:

$$(8) \quad G_\theta[z\beta + \alpha v_0 + \varepsilon > 0|x] \leq P(1|x) \leq G_\theta[z\beta + \alpha v_1 + \varepsilon \geq 0|x], \quad x \in X.$$

The observed covariates x include (z_0, z_1, Π) . The parameter θ includes (β, α) plus the parameters needed to describe $P(\varepsilon|x)$, the distribution of ε conditional on x . The identification region is the set of parameter values that satisfy (8).

The size and shape of the identification region depends on the ranges $\{(v_{j0}, v_{j1}), j \in J\}$ within which the expectation $(v_j, j \in J)$ are known to lie. Manski (2010) showed that the identification region is generically unbounded when expectations almost always have indeterminate sign; that is, when $v_{j0} < 0 < v_{j1}$ for almost all $j \in J$. This finding should serve as a warning to empirical researchers studying binary choice under uncertainty. Meaningful inference is possible only if the researcher sometimes knows whether v is positive or negative.

2.3. Monotone-Index Models

Empirical research on binary choice often assumes that ε is statistically independent of x with a specified strictly increasing distribution function, such as the standard normal or logistical distribution. Let F be the assumed distribution function for $-\varepsilon$. Then (8) becomes

$$(9) \quad F(z\beta + \alpha v_0) \leq P(1|x) \leq F(z\beta + \alpha v_1), \quad x \in X,$$

where $\theta = (\beta, \alpha)$. The identification region for (β, α) is the set of parameter values that satisfy inequalities (9) or, equivalently, the linear inequalities

$$(9') \quad z\beta + \alpha v_0 \leq F^{-1}[P(1|x)] \leq z\beta + \alpha v_1, \quad x \in X.$$

This is a monotone-index model with interval regressor data, of the form studied in Manski and Tamer (2002, Section 4). Their Corollary to Proposition 4 shows the following:

- (a) the identification region for (β, α) is convex.
- (b) Let $k \in (1, \dots, K)$. Let $P(z_k|z_{-k}, v_0, v_1)$ have unbounded support, a.e. (z_{-k}, v_0, v_1) . Then β_k is point-identified.
- (c) For each value of v , let there exist no proper linear subspace of \mathbb{R}^K having probability one under $P(z|v)$. Let $P(v_0, v_1) > 0$. Then (β, α) is point-identified.

When the parameter space is compact and some other regularity conditions hold, their *modified minimum-distance* (MMD) method provides a consistent estimate of the identification region.

Manski and Tamer (2002) also study the case in which the researcher knows only that some quantile of $P(\epsilon|x)$ is constant on X . The identification region remains convex, but is larger than the one obtained with a monotone-index model. This region may be estimated using a *modified maximum score* method.

Consideration of a simple monotone-index model illustrates the problem that occurs when expectations almost always have indeterminate sign. Let z have one component, this being an action-specific constant; hence, β is scalar. Let the population contain M distinct values for the range (v_0, v_1) , say (v_{m0}, v_{m1}) , $m = 1, \dots, M$. Let $s_m \equiv F^{-1}[P(1|m)]$. Then the linear inequalities (9') reduce to

$$(10) \quad \beta + \alpha v_{m0} \leq s_m \leq \beta + \alpha v_{m1}, \quad m = 1, \dots, M.$$

When expectations have indeterminate sign, the identification region for β is the entire real line. To see this, consider any conjectured value for β . Holding β fixed at this value, it follows from (11) that α is feasible if it solves these inequalities for all values of m :

$$(11a) \quad (s_m - \beta)/v_{m0} \leq \alpha \leq (s_m - \beta)/v_{m1} \quad \text{if } v_{m1} \leq 0,$$

$$(11b) \quad \max\{(s_m - \beta)/v_{m0}, (s_m - \beta)/v_{m1}\} \leq \alpha \quad \text{if } v_{m0} < 0 < v_{m1},$$

$$(11c) \quad (s_m - \beta)/v_{m1} \leq \alpha \leq (s_m - \beta)/v_{m0} \quad \text{if } v_{m0} \geq 0.$$

The conjectured value of β is feasible if and only if these inequalities have a solution. When expectations have indeterminate sign, (11b) provides the operative inequalities for all values of m . In this case, all $\alpha \geq \max_m \max\{(s_m - \beta)/v_{m0}, (s_m - \beta)/v_{m1}\}$ solve (11). Hence, the conjectured value of β is feasible, along with all such α .

3. Behavioral Ambiguity

In Section 2, I supposed that decision makers have complete subjective probability distributions on the states of nature, which they use to maximize expected utility. The inferential problem was observational, in that the researcher did not know what expectations people have. In this section, I suppose that persons want to maximize expected utility, but do not have unique subjective distributions on the states of nature. Instead, person j has a set Π_j of such distributions, as in Gilboa and Schmeidler (1989), Walley (1991), and much other research on choice under ambiguity.

Decision theorists have suggested various criteria for choice under ambiguity, such as the maximin and minimax-regret expected utility rules. There is no consensus on how persons with incomplete probabilistic expectations should or do behave, but decision theorists do largely agree that persons should not choose actions

that are strictly dominated. In the present setting, this means that person j should not choose an action if there exists another one that yields higher expected utility under all distributions in Π_j . Moreover, if there exists an action that outperforms all other actions under all distributions in Π_j , person j should choose this action.

These dominance conditions yield two logical relationships between choices and decision rules:

$$(12a) \quad j \text{ chooses } c \Rightarrow \exists d \in C \text{ s. t. } \int u_{jd} d\pi > \int u_{jc} d\pi, \forall \pi \in \Pi_j.$$

$$(12b) \quad \int u_{jc} d\pi > \int u_{jd} d\pi, \forall (d, \pi) \in C \times \Pi_j \text{ s. t. } d \neq c \Rightarrow j \text{ chooses } c.$$

Aggregating across the population, (12a) and (12b) imply these inequalities relating choice probabilities to the distribution of preferences and expectations:

$$(13) \quad G_\theta[\int u_{c\gamma} d\pi > \int u_{d\gamma} d\pi, \forall (d, \pi) \in C \times \Pi \text{ s. t. } d \neq c|x] \leq P(c|x) \\ \leq G_\theta[\exists d \in C \text{ s. t. } \int u_{d\gamma} d\pi > \int u_{c\gamma} d\pi, \forall \pi \in \Pi|x], \quad (c, x) \in C \times X.$$

The identification region is the set of parameter values that satisfy (13).

Observe that relationship (12b) is the same as the one (3b') that holds with observational ambiguity. Relationship (12a) is generally weaker than the one (3a) that holds with observational ambiguity. Whereas (3a) requires that the chosen action be optimal for some feasible value of π , (12a) only requires that the chosen action not be strictly dominated. However, (12a) and (3a) are equivalent when the choice set C contains two actions. Hence, inequalities (13) and (4) coincide in this case.

The above discussion assumes that the researcher knows the set Π_j of distributions held by each person j . It may be that a researcher does not know Π_j but can specify a larger set $\Pi_j' \supset \Pi_j$ that contains Π_j . Then the researcher may base inference on Π_j' , which expresses both behavioral and observational ambiguity.

5. Conclusion

This paper offers a second-best approach to econometric analysis of choice under uncertainty or behavioral ambiguity. A better approach is to measure the expectations that decision makers hold and to analyze choice behavior with these data in hand. However, expectations data often are not available. In their absence, empirical researchers typically make strong assumptions about expectations. It is important to understand how these assumptions drive findings. Moreover, researchers should want to learn what inferences are possible using weaker and more credible assumptions. The ideas and methods introduced here serve these purposes.

The analysis in this paper provides simple practical guidance for estimation of RUMs with two alternatives and linear utilities. However, the paper provides only an abstract framework for study of more complex settings with multiple alternatives and/or utilities that are nonlinear in attributes with uncertain values. Future work should aim to make estimation of such models tractable. I also see much scope for empirical research to move away from the expected utility model and study the behavior of persons who make decisions under ambiguity.

References

- Barseghyan, L., F. Molinari, and J. Teitelbaum (2016), "Inference under Stability of Risk Preferences," *Quantitative Economics*, 7, 367-409.
- Delavande, A. (2008), "Pill, Patch, or Shot? Subjective Expectations and Birth Control Choice," *International Economic Review*, 49, 999-1042.
- Domencich, T. and D. McFadden (1975), *Urban Travel Demand: A Behavioral Analysis*, Amsterdam: North-Holland.
- Ellsberg, D. (1961), "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643-669.
- Gilboa, Y. and D. Schmeidler (1989), "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141-153.
- Giustinelli, P. (2016), "Group Decision Making With Uncertain Outcomes: Unpacking Child-Parent Choice of the High School Track," *International Economic Review*, 57, 573-602.
- Horowitz, J. (1993), "Semiparametric Estimation of a Work-Trip Mode Choice Model," *Journal of Econometrics*, 58, 49-70.
- Lochner, L. (2007), "Individual Perceptions of the Criminal Justice System," *American Economic Review*, 97, 444-460.
- Manski, C. (2004), "Measuring Expectations," *Econometrica*, 72, 1329-1376.
- Manski, C. (2007), "Partial Identification of Counterfactual Choice Probabilities," *International Economic Review*, 48, 1393-1410.
- Manski, C. (2010), "Random Utility Models with Bounded Ambiguity," in B. Dutta (editor), *Structural Econometrics, Essays in Methodology and Applications*, New Delhi: Oxford University Press, 272-284.
- Manski, C. (2014), "Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy," *Quantitative Economics*, 5, 145-174.
- Manski, C. and E. Tamer (2002), "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519-546.
- McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (editor), *Frontiers in Econometrics*, New York: Academic Press.
- McFadden, D. (1981), "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden (editors), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, Mass.:M.I.T. Press.
- van der Klaauw, W. (2012), "On the Use of Expectations Data in Estimating Structural Dynamic Models,"

Journal of Labor Economics, 30, 521-554.

van der Klaauw, W. and K. Wolpin (2008), "Social Security and the Retirement and Savings Behavior of Low-Income Households," *Journal of Econometrics*, 145, 21-42.

Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, London: Chapman & Hall.

Warner, S. (1962), *Stochastic Choice of Mode in Urban Travel: a Study in Binary Choice*, Evanston: Northwestern University Press.

Wiswall, M. and B. Zafar (2015), "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 82, 791-824.

Zafar, B. (2011), "How do College Students Form Expectations," *Journal of Labor Economics*, 29, 301-348.