Structure and Uncertainty
in Discrete Choice Models

Curtis S. Signorino

Department of Political Science, 303 Harkness Hall,
University of Rochester, Rochester, NY 14627
e-mail: curt.signorino@rochester.edu

Social scientists are often confronted with theories in which one or more actors make choices over a discrete set of options. In this article, I generalize a broad class of statistical discrete choice models, with both well-known and new nonstrategic and strategic special cases. I demonstrate how to derive statistical models from theoretical discrete choice models and, in doing so, I address the statistical implications of three sources of uncertainty: agent error, private information about payoffs, and regressor error. For strategic and some nonstrategic choice models, the three types of uncertainty produce different statistical models. In these cases, misspecifying the type of uncertainty leads to biased and inconsistent estimates, and to incorrect inferences based on estimated probabilities.

1 Introduction

Social scientists are often confronted with theories where one or more actors—be they individuals, firms, parties, or states—make choices over discrete sets of actions (or options) leading to a discrete set of outcomes. Economics has been the source of many of the earliest discrete choice models, such as those of transportation choices, housing location, and decisions concerning market entry. In American politics, scholars have constructed models to explain why individuals vote for particular parties or candidates, why senators vote for particular bills, and of the choices made by the President and Congress in their often competitive relationship. Comparative research on parliamentary governments includes similar voting and coalition formation models. Finally, the international relations literature is replete with models attempting to explain why nations choose to attack each other, choose to support their allies in times of crisis, and choose to become allies in the first place.

Over the last 30 years, the most common method of statistically analyzing discrete categorical data has been to use one variant or another of logit or probit. Usually, the

---

"appropriate" model is determined by whether the dependent variable takes on more than two values and whether those values have an explicit ordering. Most of these statistical discrete choice models can be derived from assumptions of utility maximization, where the actor's utilities over outcomes include a random element with a known distribution (see, e.g., McFadden 1974a, 1974b, 1976; Hausman and Wise 1978; Alvarez and Nagler 1998). The randomness in the utility is often attributed to one of two sources: (1) regressor error or (2) "bounded rationality" due to misperception or implementation error.

Experimental game theory has recently given rise to statistical equilibrium solution concepts for extensive form games (see, e.g., Chen et al. 1996; McKelvey and Palfrey 1996, 1998; Zauner 1996). These are also based on random utility assumptions, except that the sources of error do not overlap exactly with the statistical literature: (1) incomplete information due to outcome payoff disturbances or, again (2) bounded rationality. These statistical equilibrium models have only been applied in situations where the payoffs are held constant and the variance terms estimated—for example, using data from experiments. Unfortunately, for most social scientific research, we are not at liberty to conduct experiments—for example, by creating new economies, holding elections, starting wars, or overthrowing governments. Instead, we often must test our theories and estimate the effects of substantive variables using available field data.

The role of structural assumptions in statistical inference has been extensively studied in economics. In political science, and in the context of strategic models in particular, Signorino (1999) and Smith (1999) have noted that statistical models of nonstrategic choice are often employed in testing explicitly strategic theories or in analyzing data assumed to have been generated by strategic behavior. Signorino and Yilmaz (2003) formally characterize this "strategic" specification error.

An area that has been given relatively little attention is the source of uncertainty in these models, whether it is theoretically or econometrically motivated, and how that uncertainty interacts with the structural assumptions to produce different probability models. The primary purpose of this article is to examine how and when three sources of uncertainty—regressor error, private information, and agent error—affect our models and inferences. In doing so, I will demonstrate how to derive such statistical models from theoretical first principles. Moreover, in examining the role of structure and uncertainty in discrete choice models, we will also see that the nonstrategic and strategic choice models are all special cases of a general class of models.

The details of deriving the models can at times be somewhat complex. However, the basic "recipe" for doing so is straightforward. Therefore, before jumping into the models, I will begin in the next section by providing an overview of this simple process.

Following that, I generalize a class of statistical discrete choice models that can be specified in extensive form. The statistical model developed here is specified as a general discrete choice model with utility maximizing actors and at least one source of uncertainty: agent error, private information about outcome payoffs, or regressor error. The general model has a number of well-known and new statistical models as special cases. Following that I provide examples of nonstrategic and strategic probit choice models.

In the course of deriving these statistical discrete choice models, we will also examine how the choice structure and error assumptions interact to produce (or not produce) different statistical models for estimation. The observational equivalence of two models is of particular importance for correct inference and for comparative model testing. Observational equivalence may allow for more latitude in model specification, but it negates comparative model testing. On the other hand, observational nonequivalence allows for comparative model testing, but makes model specification all the more important for correct

inferences. As we will see in Section 5, making a seemingly minor misspecification such as the type of uncertainty can at times lead to biased parameter estimates and, hence, incorrect inferences.

2   On "Theoretical" and "Statistical" Models

Let us assume that we, as social scientists, are interested in theory testing and in making correct inferences based on our structural models. There are many ways we might test a theory using statistical analysis. If we have access to all the necessary data, we might test a theory in its entirety. On the other hand, many times we have access to data only for specific actions or outcomes in a theory, and a partial test is all that can be accomplished. Regardless, for our inferences to be valid, we need a statistical model that accurately represents the theory being tested. Doing this for the entire theory ensures it holds for any partial test, so I will henceforth consider only the former. The question then becomes: How can we derive a statistical model from our theoretical model?

I frame the question this way specifically to highlight the ways in which we employ two types of models—*theoretical models* and *statistical models*—in our research. By "theoretical" model, I refer to a model that results from theory construction. Although these may take different shapes and forms, from a choice-theoretic perspective a theoretical model would identify the actor or actors, their sequence of choices, their options at decision nodes, the information available to them, and their incentives for choosing particular actions or outcomes.[1] By "statistical" model, I refer to a model that guarantees positive probability over all outcomes—for example, there is some random component in the model that induces a probability distribution over the outcomes. In political science, we most often use these in statistical regression analysis.

Theoretical and statistical models are certainly not mutually exclusive. They are simply two ways of characterizing a given model. For example, suppose we wanted to model major power war occurrence. A game-theoretic model of war with perfect and complete information would be a theoretical but not statistical model. In contrast, logit models of war where the regressors were randomly selected or selected via stepwise regression only to improve fit would be examples of statistical but not necessarily theoretical models. Finally, a game-theoretic model of war that employed a statistical equilibrium solution concept—such as the logit quantal response equilibrium of McKelvey and Palfrey (1998)—would fall in the category of both theoretical and statistical.

The discussion of theoretical and statistical models helps illuminate the question posed at the beginning of this section: how to translate a theory into a statistical model. In short, if the theory is also statistical, then the translation is done. If the theory is deterministic, then some source of uncertainty (i.e., a random variable) must be added to produce a statistical version of it. In general, for discrete or continuous dependent variables, one derives a statistical model from a theoretical model using the same general steps: (1) specify the theoretical choice model, (2) add a random component (i.e., source of uncertainty) if none exists, (3) derive the probability model associated with one's dependent variable, and (4) construct a likelihood equation based on that probability model. To help clarify these steps, I will briefly discuss the choice model, sources of uncertainty, and likelihood equation.

---

[1]Hereafter, I will not differentiate between "theoretical model," "theory," or "model," when it is clear that the latter refers to a theoretical model.
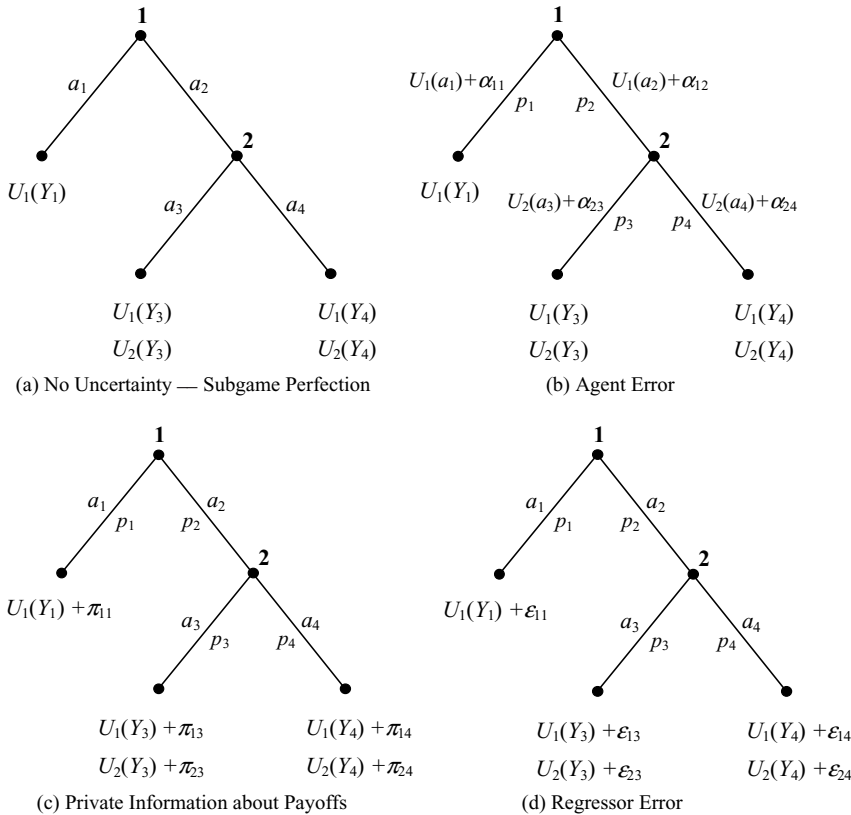
Fig. 1 Four strategic discrete choice models. All of these discrete choice models share the same choice structure and the same observed utilities, $U_m(Y_k)$. The only difference is in the source of uncertainty. In (a), there is no uncertainty. In (b), agent error α_{mj} enters through actions at information sets. In (c), each player has private information π_{mk} concerning his or her outcome payoffs. And, in (d), the players perfectly observe each other's payoffs, but the analyst does not, resulting in regressor error ϵ_{mk}.

2.1 *Choice Model*

The choice model identifies not only the set of players, their sequence of moves, their options, and their incentives, but also the rule players use to choose among options. Take Fig. 1a as an example. Here, players 1 and 2 each make one of two choices. Player 1 must choose between actions a_1 and a_2. If a_1 is chosen, then outcome Y_1 results. If a_2 is chosen, then player 2 chooses between actions a_3 and a_4, resulting in outcomes Y_3 and Y_4, respectively.

A player's "true" utility for an action or outcome will often consist of an observable component, related to the explanatory variables, and an unobservable component, related to the type of uncertainty. Player m's true utility for outcome Y_k will be written as $U_m^*(Y_k)$. The true utility for an action a_j takes the same form: $U_m^*(a_j)$. We will most often specify the observable utility component as a linear function of regressors and their associated parameters, $U_m(Y_k) = X_{mk}\beta_{mk}$, though this need not always be the case.[2] In Fig. 1a, there is no source of uncertainty, so the true outcome utilities are simply $U_m^*(Y_k) = U_m(Y_k)$.

---

[2] For a given data set, there is an implicit observation index i associated with utilities and error terms. For notational simplicity, I will tend to suppress it.

Finally, we need to identify how players choose among different options. In general, we will assume that players choose actions that maximize their expected utility. Formally specifying the choice structure and choice rule precisely relates the dependent variable—whether it is an action or outcome in the model—to the explanatory variables, which enter through the utilities. With these, we can then identify when actions or outcomes will be chosen (or should be chosen if our model is correct). For example, in Fig. 1a, player 2 will choose a_3 if $U_2^*(a_3) > U_2^*(a_4)$, and will choose a_4 otherwise. Player 1 will choose a_1 if $U_1^*(a_1) > U_1^*(a_2)$, and will choose a_2 otherwise. We can also deduce which outcome results in equilibrium:

$$
y = \begin{cases}
Y_1 & \text{if } U_2^*(Y_4) > U_2^*(Y_3) \text{ and } U_1^*(Y_4) < U_1^*(Y_1), \text{ or} \\
 & \text{if } U_2^*(Y_3) > U_2^*(Y_4) \text{ and } U_1^*(Y_3) < U_1^*(Y_1) \\
Y_3 & \text{if } U_2^*(Y_3) > U_2^*(Y_4) \text{ and } U_1^*(Y_3) > U_1^*(Y_1) \\
Y_4 & \text{if } U_2^*(Y_4) > U_2^*(Y_3) \text{ and } U_1^*(Y_4) > U_1^*(Y_1)
\end{cases}
. \tag{1}
$$

In words, Y_1 will result in two situations: (1) if player 2 prefers Y_4 to Y_3 but player 1 prefers Y_1 to Y_4, or (2) if player 2 prefers Y_3 to Y_4 but player 1 prefers Y_1 to Y_3. Similarly, Y_3 will be the outcome if player 2 prefers Y_3 to Y_4 and player 1 prefers Y_3 to Y_1.

With data for the explanatory variables X_{mk}, we could use the above choice equations to predict the actions or outcomes that would be chosen. It is important to note that by assuming no uncertainty in the model, we are assuming that the players are perfectly rational, have complete information, *and* that we as analysts perfectly observe what comprises their utilities. The action and outcome probabilities in Fig. 1a will be degenerate—either zero or one—and will correspond to the subgame perfect equilibrium.[3] Although this is not a statistical model, a comparison could be made of the subgame perfect equilibrium predictions $\hat{y}$ versus the observed outcomes y.

2.2 *Sources of Uncertainty*

To conduct statistical analysis, we require a probability model that puts positive probability over all outcomes. The random component enters as an assumption about (1) the information available to the analyst concerning the players' utilities, (2) the information available to each player concerning the other's utilities, or (3) both. The first point underscores a major difference between our undertaking and that of traditional game theorists: game theorists are omniscient with respect to their models—they know (and specify) each of the players' utilities when they solve for the equilibria. The poor empirical analyst is not so fortunate. We may assume an underlying game-theoretic model. However, we must allow for the fact that we cannot know the players' true utilities.

In this article, I consider three sources of uncertainty: agent error, private information about outcome payoffs, and regressor error.

2.2.1 Agent Error

One source of uncertainty sometimes appealed to in the random utility literature is the "black box" of bounded rationality. It is assumed that players sometimes misperceive each

---

[3]Throughout this article, I will assume ties in utilities over outcomes do not occur. In general, the assumption does not need to be made on any grounds. In a few cases, it is made for mathematical and pedagogical convenience.

other's utilities or that they err in implementing their actions.[4] Consider Fig. 1b. Agent error is assumed to enter the model via choices or actions at information sets (i.e., at decision points), rather than through the outcome utilities. Let player m's true utility for action a_j be $U_m^*(a_j) = U_m(a_j) + \alpha_{mj}$, where α_{mj} represents the error associated with player m's action a_j. This error term is assumed to be private information to m, so player m knows her true utility $U_m^*(a_j)$, but the analyst and other players know only the observable utility component $U_m(a_j)$ and the distribution of α_{mj}. Player m knows that everyone else knows the distribution of α_{mj}, and so on. Each player's equilibrium choice probabilities are, therefore, based on her expected utilities, which take into account the fact that everyone else may make mistakes. Note that this form of uncertainty is purely theoretical—we are making an assumption about what players in the game know about each other's incentives or actions. Hence, if we assume this as part of our theory, our theoretical model is then a statistical model from the start.

2.2.2 Private Information About Outcome Payoffs

Another theoretical source of uncertainty would be to assume that the players have private information concerning their own outcome payoffs. Figure 1c graphically depicts this type of uncertainty. Let player m's true utility for outcome Y_k be $U_m^*(Y_k) = U_m(Y_k) + \pi_{mk}$, where π_{mk} represents the component of player m's true utility that is private information to m.[5] We assume that player m knows his or her true utility $U_m^*(Y_k)$, but the analyst and other players observe only $U_m(Y_k)$ and know only the distribution of π_{mk}. Another way of thinking about this is that the players know m's average outcome payoff, but realize that it varies around that average according to some distribution. As in the case of agent error, player m knows that everyone else knows the distribution of π_{mk}, and so on. Each player's equilibrium choice probabilities are, therefore, based on his or her expected utilities, which take into account the uncertainty concerning outcome payoffs. This form of uncertainty is also purely theoretical and results in a theoretical model that is also statistical.

2.2.3 Regressor Error

The last form of uncertainty pertains not to the players but solely to the analyst. In this case, all of the players perfectly observe each other's utilities, but the analyst cannot perfectly model their utilities with the explanatory variables available. The error term here is similar to that in traditional regression models, where the disturbance is assumed to represent intrinsic randomness, omitted variables, or measurement error, but without any of the usual deleterious consequences. Consider Fig. 1d. Let player m's true utility for outcome Y_k be $U_m^*(Y_k) = U_m(Y_k) + \epsilon_{mk}$, where ϵ_{mk} represents the component of player m's true utility that the analyst fails to capture with the regressors used to specify $U_m(Y_k)$. All players observe the true utility $U_m^*(Y_k)$ for every other player. However, the analyst knows (or assumes) only the distribution of ϵ_{mk}. In this case, the underlying theoretical model in Fig. 1d is not statistical: it is a traditional game of complete information. What translates the theoretical model into a statistical model is an assumption of uncertainty on the analyst's part.

To recap, agent error α_{mj} and private information π_{mk} are components of our theoretical model: the former allows for nonrational behavior or "accidents," whereas the latter allows

---

[4]Opening the bounded rationality black box is beyond the scope of this article. Instead, I will take the stated form of it as a given and examine the resulting statistical models.

[5]Because of the notation, π_{mk} is sometimes also referred to as a "payoff disturbance."

for rational behavior and incomplete information. In contrast, the regressor error ϵ_{mk} is due to the analyst's inability to specify the utilities perfectly. Additionally, the errors are associated with different elements of the model and they differ in whether the analyst versus the players can observe them. For example, ϵ_{mk} and π_{mk} are associated with outcome payoffs, while α_{mj} is associated with actions at information sets; π_{mk} and α_{mj} are private information to player m, while ϵ_{mk} is unobservable only by the analyst. Finally, we do not have to assume that only a single source of uncertainty is present. For example, we might theorize that the players have private information and, at the same time, suspect that we as analysts have not perfectly specified the "observable" utilities with our regressors. Therefore, a model with both π_{mk} and ϵ_{mk} terms might be appropriate. For simplicity, however, I will focus on single sources of uncertainty and the different probability models they produce.

2.3 *Likelihood Equation*

Given a choice model, a source of uncertainty, and an appropriate distribution for the random components, we can derive probabilities for the actions and outcomes in the model. Much of Section 4 will be spent doing exactly this for different probit choice models, so I will postpone details on deriving the probabilities. For now, I simply note that the choice model in combination with an appropriately specified source of uncertainty is enough to guarantee positive probabilities over all outcomes—exactly what we need for statistical analysis. The logic behind this can be seen by considering Fig. 1d, where the uncertainty is due to regressor error. We previously saw that Y_4 is the outcome when $U_2^*(Y_4) > U_2^*(Y_3)$ and $U_1^*(Y_4) > U_1^*(Y_1)$. Substituting for the true utilities, an equivalent statement is that Y_4 occurs when $U_2(Y_4) + \epsilon_{24} > U_2(Y_3) + \epsilon_{23}$ and $U_1(Y_4) + \epsilon_{14} > U_1(Y_1) + \epsilon_{11}$. As analysts, we do not observe the ϵ_{mk}, so we cannot say with certainty whether the condition is satisfied or not. All we can do is state the probability that Y_4 will occur:

$$p_{Y_4} = \Pr[U_2(Y_4) + \epsilon_{24} > U_2(Y_3) + \epsilon_{23}, U_1(Y_4) + \epsilon_{14} > U_1(Y_1) + \epsilon_{11}].$$

Given an appropriately specified ϵ_{mk}, we can derive the probability that Y_4 occurs. The $U_m(Y_k)$ are functions of our explanatory variables, for example, $X_{mk}\beta_{mk}$. Therefore, the outcome probability relates the dependent variable to the explanatory variables and incorporates the structure of the choice model. Probabilities for the other outcomes, Y_1 and Y_3, are derived in a similar manner, as are the probabilities for the actions available to each player. In Fig. 1b–1d, the probability of action a_j is denoted as p_j.

With these action and outcome probabilities, we can construct a likelihood function for our dependent variable. Assume the probabilities have been derived for one of the models in Fig. 1b–1d. Consider the situation where our dependent variable represents which outcome occurred. Let $\delta_k = 1$ if outcome Y_k occurred in the current observation, and zero otherwise. Then the likelihood function is

$$L = \prod^N p_{Y_1}^{\delta_1} p_{Y_3}^{\delta_3} p_{Y_4}^{\delta_4}. \tag{2}$$

The outcome probabilities are functions of the explanatory variables and effects parameters. Therefore, the maximum likelihood estimates can be found by maximizing Eq. (2) with respect to the parameters—and standard statistical tests (e.g., t or likelihood ratio) can be conducted.

If our dependent variable instead consists of actions taken by a player, then we would construct a similar likelihood function, but with the corresponding action probabilities.

For example, for a dependent variable denoting whether a_1 or a_2 occurred, the likelihood function would be

$$L = \prod^N p_1^{\delta_{a_1}} p_2^{\delta_{a_2}}, \tag{3}$$

where $\delta_{a_j} = 1$ if action a_j was chosen, and zero otherwise. The reader familiar with traditional binomial and multinomial choice models will note that Eqs. (2) and (3) share the same general form as in multinomial logit and probit. However, the probabilities p_{Y_k} and p_j incorporate the choice structure of our theory, which may differ from nonstrategic binomial and multinomial choice structures.

3 A General Class of Statistical Discrete Choice Models

The class of models in this section generalizes those statistical discrete choice models in extensive form that allow for agent error, private information about payoffs, and regressor error. Although it is based on a general specification of an extensive form game, it does not require that the model be strategic. In words, we first need to define the choice tree, then assign players, information sets, actions, and outcomes to the nodes of that tree, then assign utilities to the actions and outcomes, then specify the distribution of the errors associated with utilities over actions and outcomes, and, finally, specify the rule players use to make decisions. The notation can be dense at times. However, the core ideas are exactly the same as those just outlined. Concrete examples of these will be given in Section 4.

3.1 *Choice Tree*

The discrete choice tree is specified as a set of nodes with an asymmetric precedence relation defined over the nodes. Let Q be the set of all nodes, with q_0 the initial node. Let $\prec$ be an asymmetric precedence relation, such that $q \prec q'$ if node q immediately precedes node q', and $q' \succ q$ if node q' immediately follows q. For a node q, define the set of its immediate successor nodes as $S(q) = \{q' \mid q \prec q'\}$ and define a terminal node q^t as one where $S(q^t) = \emptyset$. Define the set of terminal nodes as Q^t. A path $P(q^t)$ from the initial node q_0 to a terminal node q^t is a sequence of nodes $(q_0, \ldots, q^t)$ that connects q_0 to q^t such that for each nonterminal node in the sequence, the following node is its immediate successor. All paths are assumed to be countably finite. This rules out games such as an infinitely repeated prisoner's dilemma.

3.2 *Information Sets*

Let M be the set of players. Define the information set containing node q and assigned to player m as $\iota_m(q)$. Define the set of information sets assigned to player m as I_m and the set of all information sets as $I = \bigcup_m I_m$. As all information sets considered here are singletons, it will sometimes be less notationally cumbersome to refer to an information set simply by the single node assigned to it.

3.3 *Actions and Outcomes*

Let $a(q)$ be the label for the action (or option) required to reach q from its predecessor, let $A(q) = \{a(q') \mid q' \in S(q)\}$ be the set of all actions available at q for the player assigned to q, and let $A = \bigcup_q A(q)$ be the set of all actions in the game. Define $g : Q^t \to Y$ as the function that maps terminal nodes Q^t to outcomes Y. Outcomes are simply labels assigned

to terminal nodes. Although different paths must by definition lead to different terminal nodes, they can result in the same outcome. For example, in a crisis model, there may be multiple paths to war.

3.4 *Utilities and "Error" Terms*

Utilities are defined similarly to those of traditional random utility models and statistical equilibrium models.[6] However, here we will specify in detail the types of uncertainty that can enter into the model. Although the utilities and error terms may vary over observations in a data set, I will continue to suppress the observation index to reduce notation.

For terminal node q_j^t mapped to outcome Y_k, we can alternatively speak of m's true utility $U_m^*(Y_k)$ for outcome Y_k or of m's true utility $U_m^*[a(q_j^t)]$ for the action $a(q_j^t)$ leading to the outcome Y_k assigned to node q_j^t. Let player m's true utility $U_m^*(Y_k)$ of outcome Y_k be

$$U_m^*(Y_k) = U_m(Y_k) + \epsilon_{mk} + \pi_{mk}. \tag{4}$$

Here, $U_m(Y_k)$ is the observable component of m's utility for outcome Y_k—for example, as a function of explanatory variables X_{mk} and effects parameters β_{mk}. It is assumed that the analyst and all players observe $U_m(Y_k)$. ϵ_{mk} is the portion of the true utility that is not captured by the regressors. We assume each player observes the ϵ_{mk} of every other player, but the analyst cannot observe ϵ_{mk} (otherwise he or she would have included it in the regressors). π_{mk} is the portion of m's true utility that is private to m. Neither the analyst nor the other players observe player m's π_{mk}.

For nonterminal node q_j, let player m's true utility $U_m^*[a(q_j)]$ for choosing action $a(q_j)$ be defined as

$$U_m^*[a(q_j)] = U_m[a(q_j)] + \alpha_{mj}, \tag{5}$$

where $U_m[a(q_j)]$ is m's expected utility for reaching node q_j and α_{mj} represents agent error associated with action $a(q_j)$. Again, α_{mj} is assumed known only by player m.

3.5 *Distribution of Errors*

Let ϵ_m be the vector of m's regressor error terms and let $\epsilon = (\epsilon_1, \ldots, \epsilon_{\#M})$ be the vector of every players' regressor errors.[7] Similarly, let π_m be the vector of m's payoff disturbances over each of her utilities and let $\pi = (\pi_1, \ldots, \pi_{\#M})$ be the vector of every players' payoff disturbances. Finally, let α_m be the vector of player m's agent errors over each of his or her actions and let $\alpha = (\alpha_1, \ldots, \alpha_{\#M})$ be the vector of every agent's errors. For mathematical convenience, let ϵ, π, and α be distributed according to the same continuous joint density function $f(\cdot)$. Denote the combined error vector as $\Omega = (\epsilon, \pi, \alpha)$.

3.6 *Agent Utility Maximization*

A multiagent representation is assumed.[8] A player is represented by an "agent" at each information set who shares the player's payoff function but makes decisions independently

---

[6]For random utility models, see McFadden 1974a, 1974b, 1976; Hausman and Wise 1978. For statistical equilibrium models, see McKelvey and Palfrey 1996, 1998; Chen et al. 1996; Zauner 1996.

[7]The notation $\#S$ is used to denote the number of elements in the set S.

[8]This is also referred to as "agent normal form" or "agent strategic form."

of the other agents. At each information set $\iota_{mj}(q)$, the agent for player m chooses the action $a(q')$ that maximizes her utility from that node in the tree: $U_m^*[a(q')] > U_m^*[a(q'')]$ $\forall a(q'') \neq a(q') \in A(q)$. Consequently, an equilibrium strategy must be in the interest of every agent, even off the equilibrium path.

3.7 *Statistical Discrete Choice Model*

Given the above, the tuple $(Q, \prec, M, I, A, Y, U, \Omega)$ defines a statistical discrete choice model with well-defined choice probabilities over each action and, hence, over each outcome.

Define the probability of outcome Y_k as

$$p_{Y_k} = \Pr[Y_k]$$
$$= \sum_{\{q^t | g(q^t) = Y_k\}} \Pr[P(q^t)]. \tag{6}$$

If only one terminal node q^t is assigned to a particular outcome Y_k, then Eq. (6) is $p_{Y_k} = \Pr[P(q^t)]$, the joint probability of realizing each of the actions on the path $P(q^t)$. If the same outcome is assigned to multiple terminal nodes, then Eq. (6) is the sum of the probabilities of realizing each of those paths leading to the different terminal nodes $\{q^t \mid g(q^t) = Y_k\}$ mapped to the same outcome.

Assume now that we have N observations of a dependent variable y that takes on the values $y_i \in Y$. Define $j \in Y$ and let the dummies $y_{i,j}$ take on the value

$$y_{i,j} = \begin{cases} 1 & \text{if } j = y_i \\ 0 & \text{otherwise} \end{cases}.$$

In other words, $y_{i,j}$ is a dummy that indicates whether outcome j occurred or not in observation i. Assuming that the utilities are specified in terms of explanatory variables (e.g., $U_{mki}(Y_k) = X_{mki}\beta_{mk}$), then we can denote the probability of Y_k occurring in observation i as $p_{Y_k i}$ and estimate the parameters β_{mk} by maximizing the multinomial likelihood function

$$L = \prod_{i=1}^{N} \prod_{j \in Y} p_{ji}^{y_{i,j}} \tag{7}$$

with respect to the β_{mk}. Again, Eq. (7) should not be interpreted as representing a multinomial logit or probit model. Rather, the data vector $\mathbf{y}$ has a multinomial distribution. Whether the model implied by the tuple $(Q, \prec, M, I, A, Y, U, \Omega)$ is multinomial logit or probit, binomial logit or probit, a sequential model, a strategic model—or something else entirely different—will depend on the elements of the tuple. I now turn to these special cases.

3.8 *Special Cases*

The general model has a number of well-known, as well as new, special cases, depending on what we substitute for the elements of the tuple $(Q, \prec, M, I, A, Y, U, \Omega)$. Table 1 displays a number of these special cases, which are defined based on the number of players $\#M$, the number of information sets $\#I$, the number of terminal nodes $\#Q^t$, the number of

Table 1 Special cases of the general statistical discrete choice model
defined by the tuple $(Q, \prec, M, I, A, Y, U, \Omega)$.

Players M	Info sets I	Terminal nodes Q^t	Outcomes Y	Special cases	
				$\Omega \sim TlEV$	$\Omega \sim N(0, \Sigma)$
				Nonstrategic	
1	1	2	2	Binomial logit	Binomial probit
1	1	≥ 2	$\#Q^t$	Multinomial logit	Multinomial probit
				Nested logit[a]	
1	>1	>$\#I$	>1	Sequential logit	Sequential probit
1 or 2	2	3	3		Heckman selection
				Strategic	
>1	$\geq \#M$	>$\#I$	>1	Logit	Probit

Note. The general model includes a number of well-known, as well as new, statistical models, depending on what we substitute for the elements of the tuple. The special cases in the table are defined based on the number of players $\#M$, the number of information sets $\#I$, the number of terminal nodes $\#Q^t$, the number of outcomes $\#Y$, and whether the errors have type 1 extreme value or normal distributions. Within the class of strategic models are further special cases that depend on which sources of uncertainty are present: agent error, private information, regressor error, or some combination of the three.

[a]$\Omega \sim$ GEV.

outcomes $\#Y$, and whether the errors $\Omega = (\epsilon, \pi, \alpha)$ have type 1 extreme value or normal distributions.

For example, the main difference between nonstrategic and strategic models is that the former usually have only a single decision maker ($\#M = 1$), while the latter have multiple players ($\#M > 1$). Within the nonstrategic models, the main source of differentiation is the number of terminal nodes and the number of information sets. Binomial logit/probit has only one information set and is a special case of multinomial logit. These are differentiated from sequential logit/probit, which have multiple information sets. Heckman-type selection models may have either one or two players. Although not shown in Table 1, as we will see in Section 4, different sequential models also arise depending on our assumptions concerning whether a player's agents have private information concerning their disturbances relative to other agents of the same player (more on this later).

With the strategic models, we assume there are multiple players, at least one information set for each player, and more terminal nodes than information sets. This last assumption ensures that we have at least a binary decision by each player. There are a number of existing and new special cases within the "strategic" class. For example, the logit quantal response equilibrium (QRE) of McKelvey and Palfrey (1998) is the special case with agent errors that are independently distributed type 1 extreme value. A statistical equilibrium model with independently normally distributed agent errors would be equivalent to a probit version of the QRE. Signorino (2002) analyzes a model with normally distributed agent errors that are correlated between two players, synthesizing aspects of traditional

Heckman selection models and QRE-based strategic models. The statistical equilibrium model in Zauner (1996) involves normally distributed payoff perturbations.[9] With the exception of Signorino (1999, 2002) and Signorino and Yilmaz (2003), none of these have been developed for use in regression analysis. Finally, these are just a few of the special cases of the general model. Any number of other models are possible, depending on the assumptions concerning which types of uncertainty are present, the observability of the disturbances by the analyst and players, and the distribution and covariance structure of the disturbances.

4 Examples: Probit Choice Models

The models generalized in Section 3 are illustrated using three examples: (1) nonstrategic nonsequential choice models, (2) nonstrategic sequential choice models, and (3) strategic choice models. Although the first has been addressed in detail elsewhere under the rubric of "discrete choice" (see McFadden 1974a, 1974b, 1976; Hausman and Wise 1978; Maddala 1983; Pudney 1989) and the second has also received some (albeit much less) attention, all three are useful for demonstrating not only how one would apply the general model to specific applications, but also how the choice and outcome probabilities relate to the characteristics of the tuple $(Q, \prec, M, I, A, Y, U, \Omega)$. In fact, the three models provide a nice progression concerning those characteristics.

Throughout this section, I will assume that errors are normally distributed. Therefore, the discrete choice models examined here will be nonstrategic and strategic probit models. For a given discrete choice model, I show how it relates to the general model and how to derive the choice probabilities.

4.1 *Nonstrategic Nonsequential Choice*

The simplest nontrivial discrete choice model is that depicted in Fig. 2a, where a single decision maker (labeled **1**) must choose between $\#Y = 2$ outcomes. Examples might include whether a state votes for a United Nations resolution, whether a citizen votes Democrat or Republican, or whether a commuter takes the bus to work or drives. The tree is defined by the set of nodes $Q = \{q_0, q_1, q_2\}$ and the precedence relation among them: $q_0 \prec q_1$ and $q_0 \prec q_2$. The terminal nodes $Q^t = \{q_1, q_2\}$ are mapped to outcomes $Y = \{Y_1, Y_2\}$. There is only a single information set $\iota_1(q_0)$. The action required to reach node q_j is denoted by a_j. The observation index i will be suppressed unless it is required to highlight that the dependent variable, explanatory variables, utilities, or probabilities vary over observations.

The theoretical story for regressor error ϵ and payoff disturbances π is identical here. What usually differentiates them is that ϵ is observed by all players, while π is private information. As there is only one decision maker, both represent part of a perfectly rational decision maker's utility that is unobserved by the analyst. The story for agent error α is slightly different in that α is assumed to arise because of misperception or implementation error on the decision maker's part. Nevertheless, regardless of the underlying theoretical story for the uncertainty, the resulting statistical models estimated will be identical, as the choice structure in Fig. 2a contains only a single information set.

---

[9]Zauner (1996) develops a two-sided incomplete information model with normally distributed payoff disturbances for the analysis of experimental data from the centipede game, but where payoffs are fixed and only the variance is estimated.
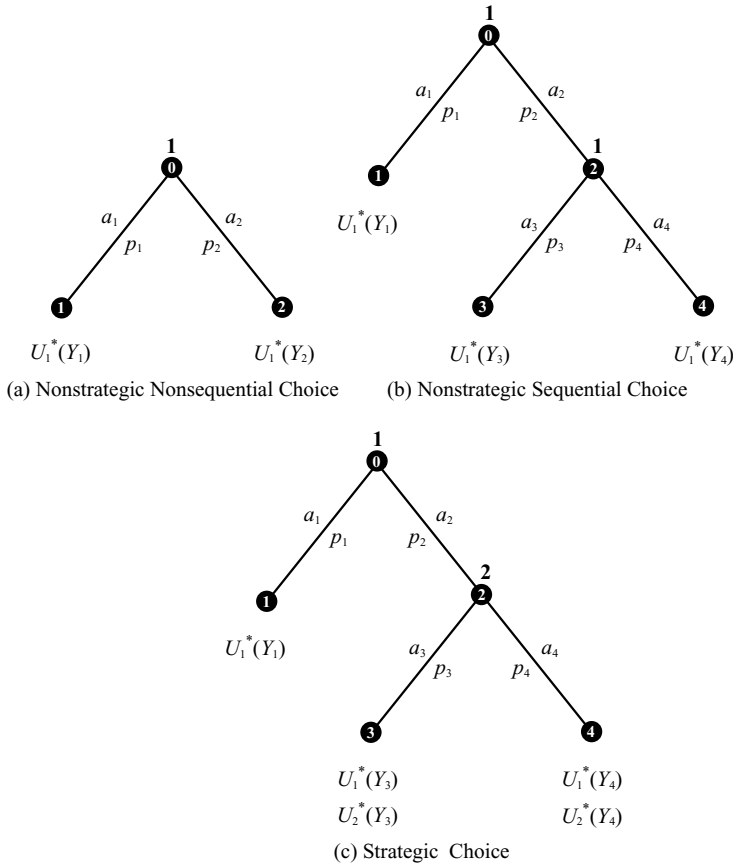
Fig. 2 Discrete choice examples. Parts (a) and (b) are nonstrategic choice models, consisting of a single decision maker. Parts (b) and (c) are sequential choice models. Part (c) is a strategic choice model, where player 1 must condition his or her choice on what he or she expects player 2 will do. Options are denoted by a_j, choice probabilities as p_j, outcomes as Y_k, and player m's utility for outcome Y_k as $U_m^*(Y_k)$.

To see this, consider that the utilities in each case are
Regressor error:

$$U_1^*(Y_k) = U_1(Y_k) + \epsilon_k \tag{8}$$

Private information about outcome payoffs:

$$U_1^*(Y_k) = U_1(Y_k) + \pi_k \tag{9}$$

Agent error:

$$\begin{aligned} U_1^*(a_k) &= U_1(a_k) + \alpha_k \\ &= U_1(Y_k) + \alpha_k. \end{aligned} \tag{10}$$

A utility-maximizing decision maker will choose Y_1 if $U_1^*(Y_1) > U_1^*(Y_2)$. Assuming the disturbances are normally distributed with mean zero and covariance structure $\Sigma = [\begin{smallmatrix} \sigma_1^2 \\ \sigma_{12} & \sigma_2^2 \end{smallmatrix}]$,

then in the regressor error case,

$$
\begin{aligned}
p_{Y_1} &= \Pr[U_1^*(Y_1) > U_1^*(Y_2)] \\
&= \Pr[U_1(Y_1) + \epsilon_1 > U_1(Y_2) + \epsilon_2] \\
&= \Pr[\epsilon_2 - \epsilon_1 < U_1(Y_1) - U_1(Y_2)] \\
&= \Phi\left[\frac{U_1(Y_1) - U_1(Y_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}} \right]
\end{aligned}
\tag{11}
$$

where $\Phi(x)$ is the standard normal distribution. The probability that the decision maker will choose option Y_2 is simply $p_{Y_2} = 1 - p_{Y_1}$.

It is easy to see that the probabilities will be exactly the same (under the same conditions) for the agent error and private information models. Hence, given the same structural and covariance assumptions, the three forms of uncertainty lead to observationally equivalent statistical models.

4.2 *Nonstrategic Sequential Choice*

This section serves as a bridge between the nonstrategic models and the strategic models. On the one hand, the sequential models presented in this section are all nonstrategic and, as we will see, at times reduce to nonsequential models. On the other hand, the sequence of decisions, observability of error terms, and calculations of expected utility enter into the sequential models' choice probabilities in much the same way as they do for the strategic models. Depending on the assumptions, the choice probabilities in this section tend to be similar either to those of the previous section or to those of the next section.

Figure 2b displays a situation where a single decision maker (labeled **1**) makes sequential choices over actions leading to $\#Y = 3$ outcomes. First, he or she must choose between action a_1 and a_2. If a_2 is chosen, then he or she must choose between a_3 and a_4. The tree is defined by the set of nodes $Q = \{q_0, q_1, q_2, q_3, q_4\}$ and the precedence relation among them: $q_0 \prec q_1$, $q_0 \prec q_2$, $q_2 \prec q_3$, and $q_2 \prec q_4$. The terminal nodes $Q^t = \{q_1, q_3, q_4\}$ are mapped to outcomes $Y = \{Y_1, Y_3, Y_4\}$. There are two information sets: $\iota_1(q_0)$ and $\iota_1(q_2)$. The action required to reach node q_j is denoted by a_j. Decision maker 1 has utilities $U_1^*(Y_1)$, $U_1^*(Y_3)$, and $U_1^*(Y_4)$ for the three outcomes.

4.2.1 Regressor Error

In this case, the decision maker is perfectly rational and perfectly observes the utilities, but the analyst does not observe part of the variation in the utility. Hence, $U_1^*(Y_k) = U_1(Y_k) + \epsilon_k$. To calculate p_{Y_k}, we must first determine the conditions for Y_k to be realized. Given the above assumptions, the following identifies the conditions for each of the outcomes to occur:

$$
y = \begin{cases}
Y_1 & \text{if } U_1^*(Y_1) > U_1^*(Y_3) \text{ and } U_1^*(Y_1) > U_1^*(Y_4) \\
Y_3 & \text{if } U_1^*(Y_3) > U_1^*(Y_4) \text{ and } U_1^*(Y_3) > U_1^*(Y_1) \\
Y_4 & \text{if } U_1^*(Y_4) > U_1^*(Y_3) \text{ and } U_1^*(Y_4) > U_1^*(Y_1)
\end{cases}
\tag{12}
$$

As we do not perfectly observe the true utilities, we can only make probabilistic statements concerning the outcomes. So, for example, the probability p_{Y_1} of Y_1 occurring in observation i is

$$p_{Y_1} = \Pr[U_1(Y_1) + \epsilon_1 > U_1(Y_3) + \epsilon_3, U_1(Y_1) + \epsilon_1 > U_1(Y_4) + \epsilon_4]. \qquad (13)$$

Assuming that $(\epsilon_1, \epsilon_3, \epsilon_4) \sim N(\mathbf{0}, \Sigma_\epsilon)$, with

$$\Sigma_\epsilon = \begin{bmatrix} \sigma_{\epsilon_1}^2 & & \\ \sigma_{\epsilon_{13}} & \sigma_{\epsilon_3}^2 & \\ \sigma_{\epsilon_{14}} & \sigma_{\epsilon_{34}} & \sigma_{\epsilon_4}^2 \end{bmatrix} \qquad (14)$$

and letting $\eta_{31} = \epsilon_3 - \epsilon_1$ and $\eta_{41} = \epsilon_4 - \epsilon_1$, then Eq. (13) can be rewritten as

$$
\begin{aligned}
p_{Y_1} &= \Pr[\eta_{31} < U_1(Y_1) - U_1(Y_3), \eta_{41} < U_1(Y_1) - U_1(Y_4)] \\
&= \int_{-\infty}^{U_1(Y_1)-U_1(Y_3)} \int_{-\infty}^{U_1(Y_1)-U_1(Y_4)} \phi(\eta_{31}, \eta_{41}) \, d\eta_{41} \, d\eta_{31}
\end{aligned}
\qquad (15)
$$

where $\phi(\eta_{31}, \eta_{41})$ is the bivariate normal density with mean zero and variance

$$\Sigma_\eta = \begin{bmatrix} \sigma_{\epsilon_1}^2 + \sigma_{\epsilon_3}^2 - 2\sigma_{\epsilon_{13}} & \\ \sigma_{\epsilon_1}^2 - \sigma_{\epsilon_{13}} - \sigma_{\epsilon_{14}} + \sigma_{\epsilon_{34}} & \sigma_{\epsilon_1}^2 + \sigma_{\epsilon_4}^2 - 2\sigma_{\epsilon_{14}} \end{bmatrix}. \qquad (16)$$

The reader will notice that this is equivalent to the traditional multinomial probit probability of choosing outcome Y_1 from the set $Y = \{Y_1, Y_3, Y_4\}$—that is, where there is only a single information set and three outcomes (see, e.g., Hausman and Wise 1978; Alvarez and Nagler 1998). The probabilities of choosing Y_3 and Y_4 yield similar multinomial probit probabilities. Therefore, with a single perfectly rational decision maker and only regressor error, the sequence of choices are irrelevant, and the statistical model is observationally equivalent to a nonstrategic nonsequential model with the same number of options and the same covariance structure.

4.2.2 Private Information About Outcome Payoffs

In Section 4.1, we saw that with only a single information set, payoff disturbances yielded not only the same statistical model as with regressor error, but also the same underlying story: they represent variation in outcome utilities that the decision maker observes, but the analyst does not. With multiple information sets, that may or may not be the case, depending on our assumption of whether payoff disturbances are observable by all "agents" of the same player.

In the typical multiagent model (see Myerson 1991, p. 61; Osborne and Rubinstein 1994, p. 250), each player is represented at each information set by a different agent, who has the same preferences as the player and makes a decision for the player. In the current context, the usual multiagent representation would assume that every agent of a given player observes the disturbances of the other agents of the same player, as they all share the same preferences.

A variation on this is to assume that the error terms are private information for each individual agent of the player. In other words, each agent of the player observes her own disturbance, but does not observe the error terms of the other agents.[10] Why might we do this, given that there is really only a single decision maker? Simply put, decision makers may not know what their true payoffs are "down the road," or, in the agent error models, what errors they might make in future decisions. It may make more sense then to assume that a player perceives her outcome payoff differently at different points in the game or that the agent error associated with actions at an information set is observable only by the agent at that information set. Under these assumptions, an agent must form an expected utility for an action that leads to another information set, even when that information set is assigned to the same player the agent represents. Examples of both the standard multiagent representation and this variation are given in the following.

In the case where the decision maker's agents observe each other's payoff disturbances, the problem is identical to one of regressor error: the decision maker is perfectly rational, the only uncertainty is on the analyst's part, and the uncertainty is with respect to the outcome utilities. The sequence of choices is irrelevant. If we assume that the payoff disturbances are normally distributed, then the resulting choice probabilities are the same multinomial probit probabilities we derived in the regressor error case. Both are observationally equivalent with a nonstrategic nonsequential model with any type of error, assuming the same outcomes, utilities over outcomes, and distribution of errors.

This is not the case when the payoff disturbances are private information to each agent of a player. In fact, this situation is equivalent to a game with multiple players, who all share the same observable utility components $U_m(Y_k)$, but have different payoff disturbances π_{mk}.[11] The probability that player 1 chooses outcome Y_4 is the probability $\Pr[P(q_4)]$ that the path $P(q_4)$ is realized, which is the joint probability of the decisions along it being realized. Denote by p_j the probability that action a_j is chosen at the parent information set of a_j. If we assume that the payoff disturbances are independent of each other, then the joint probability is just the product of the two individual choice probabilities p_2 and p_4.[12] To calculate the choice probabilities, we will work up the tree, as decisions made at information sets earlier in the tree are based on expected utilities that require choice probabilities for actions later in the tree.

As we do not observe the payoff disturbances, the probability p_4 that the agent at information set $\iota_1(q_2)$ chooses a_4 is

$$p_4 = \Phi\left[\frac{U_1(Y_4) - U_1(Y_3)}{\sqrt{\sigma_{\pi_3}^2 + \sigma_{\pi_4}^2}} \right]. \tag{17}$$

The probability p_2 that the agent at information set $\iota_1(q_0)$ chooses a_2 must be written in terms of her expected utility for choosing action a_2, which is $U_1^*(a_2) = p_3 U_1^*(Y_3) + p_4 U_1^*(Y_4)$.

---

[10]Although not expressed in exactly the same way as presented here, nonstrategic "decentralized" sequential choice models share this feature (see, e.g., Pudney 1989, pp. 122–130).

[11]Technically, we should let the true outcome utility of player m's agent n at information set $\iota_m(q_n)$ be denoted by $U_{1_n}^*(Y_k) = U_1(Y_k) + \pi_{1_n k}$. Because this notation is cumbersome, I will not index agents of players. However, the reader should keep in mind that the true utilities and error terms are implicitly indexed this way under the assumption of multiple agents with private errors.

[12]I will hereafter assume that errors are independently distributed; hence, all covariances are zero. This is made purely to simplify the math. However, it does have other implications that we will return to later.

The agent's probability of choosing a_2 is then

$$
\begin{aligned}
p_2 &= \Pr[U_1^*(a_2) > U_1^*(a_1)] \\
&= \Pr[p_3 U_1^*(Y_3) + p_4 U_1^*(Y_4) > U_1^*(Y_1)] \\
&= \Pr[p_3 [U_1(Y_3) + \pi_3] + p_4 [U_1(Y_4) + \pi_4] > U_1(Y_1) + \pi_1] \\
&= \Phi\left[\frac{p_3 U_1(Y_3) + p_4 U_1(Y_4) - U_1(Y_1)}{\sqrt{p_3^2 \sigma_{\pi_3}^2 + p_4^2 \sigma_{\pi_4}^2 + \sigma_{\pi_1}^2}} \right].
\end{aligned}
\tag{18}
$$

Choice probabilities p_3 and p_1 are, of course, $1 - p_4$ and $p_1 = 1 - p_2$, respectively.

Because the error terms are assumed independent, the outcome probabilities simplify to

$$
p_{Y_1} = p_1 \tag{19}
$$
$$
p_{Y_3} = p_2 p_3 \tag{20}
$$
$$
p_{Y_4} = p_2 p_4. \tag{21}
$$

These are clearly not the multinomial probit probabilities we saw in the regressor error case. The probability of Y_1 here is

$$
p_{Y_1} = \Phi\left[\frac{U_1(Y_1) - [p_3 U_1(Y_3) + p_4 U_1(Y_4)]}{\sqrt{p_3^2 \sigma_{\pi_3}^2 + p_4^2 \sigma_{\pi_4}^2 + \sigma_{\pi_1}^2}} \right].
\tag{22}
$$

In contrast to the multinomial probit's joint comparison of $U_1(Y_1)$ against both $U_1(Y_3)$ and $U_1(Y_4)$, here the sequential choice is reflected in the numerator of Eq. (22) by the comparison of $U_1(Y_1)$ to a weighting of $U_1(Y_3)$ and $U_1(Y_4)$, where the weighting is based on the probability that the respective outcomes will be selected. This is directly a result of the agent at information set $i_1(q_0)$ having to form an expected utility concerning action a_2 because of her uncertainty concerning her decision at $i_1(q_2)$. Second, notice that while the multinomial probit probabilities have a constant variance in their denominators, here the variance in the denominator contains probabilities (e.g., p_3 and p_4) that will vary over the observations (as they are functions of utilities that are themselves functions of explanatory variables that will vary over observations).

4.2.3 Agent Error

I consider only the case where agent error over actions at a particular information set is private information to the agent at that information set. The derivation of the outcome probabilities proceeds similarly to those in the model with private payoff disturbances.

As we do not observe the agent error, the probability p_4 that the agent at information set $\iota_1(q_2)$ chooses a_4 is

$$
p_4 = \Phi\left[\frac{U_1(Y_4) - U_1(Y_3)}{\sqrt{\sigma_{\alpha_3}^2 + \sigma_{\alpha_4}^2}} \right].
\tag{23}
$$

The probability p_2 that the agent at information set $\iota_1(q_0)$ chooses a_2 must be written in terms of her expected utility for choosing action a_2. Since the error term comes in through

actions rather than outcomes, the expected utility is slightly different than with the payoff disturbances model:

$$
\begin{aligned}
U_1^*(a_2) &= U_1(a_2) + \alpha_2 \\
&= [p_3 U_1(a_3) + p_4 U_1(a_4)] + \alpha_2.
\end{aligned}
\tag{24}
$$

The agent's probability of choosing a_2 is then

$$
\begin{aligned}
p_2 &= \Pr[U_1^*(a_2) > U_1^*(a_1)] \\
&= \Pr[p_3 U_1(a_3) + p_4 U_1(a_4) + \alpha_2 > U_1(a_1) + \alpha_1] \\
&= \Pr[p_3 U_1(Y_3) + p_4 U_1(Y_4) + \alpha_2 > U_1(Y_1) + \alpha_1] \\
&= \Phi\left[\frac{p_3 U_1(Y_3) + p_4 U_1(Y_4) - U_1(Y_1)}{\sqrt{\sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2}} \right].
\end{aligned}
\tag{25}
$$

The outcome probabilities take the same general form as Eqs. (19)–(21), but substituting the agent error versions of the p_j action probabilities.

As with private payoff disturbances, these are clearly not multinomial probit probabilities. Here, the probability of Y_1 occurring is

$$
p_{Y_1} = \Phi\left[\frac{U_1(Y_1) - [p_3 U_1(Y_3) + p_4 U_1(Y_4)]}{\sqrt{\sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2}} \right].
\tag{26}
$$

This is somewhat similar to the corresponding probability in the (agent–private) payoff disturbance case. However, there the outcome probabilities contained variance terms that varied over observations, whereas here the variance terms [in the denominator of Eq. (26)] are constant across observations.

The import of this section has been to show that the type of uncertainty specified by the theoretical model *can* matter in the resulting statistical model. Although the nonstrategic nonsequential choice models produced observationally equivalent statistical models, the sequencing of decisions and the assumptions concerning uncertainty resulted in a number of different statistical models here. As we will see, the same holds true for the strategic models, to which we now turn.

4.3 *Strategic Choice*

The new element introduced in this section is that there are now not just multiple agents of a single decision maker, but multiple decision makers, each of whom must condition their behavior on the expected behavior of the others. The probabilities are therefore considered *equilibrium* choice probabilities. We have actually already seen one simple example of this. A nonstrategic sequential choice model, where disturbances are assumed private information to each agent, can be viewed as a strategic model where each agent is a different decision maker but all decision makers share the same observed outcome utilities $U_m(Y_k)$. The nonstrategic choice probabilities in the former are the equilibrium choice probabilities of the latter.

Figure 2c, which displays a strategic discrete choice model in extensive form, will be used as the referent example throughout this section. Here, we have two decision makers

(labeled **1** and **2**) who make choices in turn. Player 1 must choose between action a_1 and a_2. If a_2 is chosen, then player 2 must choose between a_3 and a_4. The tree is defined by the set of nodes $Q = \{q_0, q_1, q_2, q_3, q_4\}$ and the precedence relation among them: $q_0 \prec q_1$, $q_0 \prec q_2$, $q_2 \prec q_3$, and $q_2 \prec q_4$. The terminal nodes $Q^t = \{q_1, q_3, q_4\}$ are mapped to outcomes $Y = \{Y_1, Y_3, Y_4\}$. There are two information sets: $\iota_1(q_0)$ and $\iota_2(q_2)$. The action required to reach node q_j is denoted by a_j. Decision maker m has utility $U_m^*(Y_k)$ for outcome Y_k.

4.3.1 Regressor Error

Again, we assume that both players are perfectly rational and view each other's payoffs perfectly, but we (the analysts) do not observe part of the variation in the utility. Hence, $U_m^*(Y_k) = U_m(Y_k) + \epsilon_{mk}$. As in the sequential model, to calculate p_{Y_k}, we ask ourselves: What are the conditions for Y_k to be realized? With perfectly rational players, who have complete information about each other's utilities and make decisions at singleton information sets, that question is equivalent to asking when Y_k will be the subgame perfect equilibrium (SPE) outcome in Fig. 2c.

Without any constraints on the preferences over outcomes, a general specification of the SPE conditions can be quite tedious for more complicated models. However, for the model in Fig. 2c, the conditions are not terribly cumbersome. The SPE is

$$
y = \begin{cases}
Y_1 & \text{if } U_2^*(Y_3) > U_2^*(Y_4) \text{ and } U_1^*(Y_1) > U_1^*(Y_3) \text{ or} \\
& \text{if } U_2^*(Y_4) > U_2^*(Y_3) \text{ and } U_1^*(Y_1) > U_1^*(Y_4) \\
Y_3 & \text{if } U_2^*(Y_3) > U_2^*(Y_4) \text{ and } U_1^*(Y_3) > U_1^*(Y_1) \\
Y_4 & \text{if } U_2^*(Y_4) > U_2^*(Y_3) \text{ and } U_1^*(Y_4) > U_1^*(Y_1)
\end{cases} \tag{27}
$$

As we do not perfectly observe the true utilities, we can only make probabilistic statements concerning the outcomes. So, for example, the probability of Y_1 occurring is

$$
\begin{aligned}
p_{Y_1} &= \Pr[U_2^*(Y_3) > U_2^*(Y_4), U_1^*(Y_1) > U_1^*(Y_3)] \\
&\quad + \Pr[U_2^*(Y_4) > U_2^*(Y_3), U_1^*(Y_1) > U_1^*(Y_4)] \\
&= \Pr[U_2(Y_3) + \epsilon_{23} > U_2(Y_4) + \epsilon_{24}, U_1(Y_1) + \epsilon_{11} > U_1(Y_3) + \epsilon_{13}] \\
&\quad + \Pr[U_2(Y_4) + \epsilon_{24} > U_2(Y_3) + \epsilon_{23}, U_1(Y_1) + \epsilon_{11} > U_1(Y_4) + \epsilon_{14}]. \tag{28}
\end{aligned}
$$

Denote the variance of ϵ_{ij} as $\sigma^2_{\epsilon_{ij}}$ and the covariance of ϵ_{ij} with ϵ_{ik} as $\sigma_{\epsilon_{ijk}}$. If we let $\eta_{ijk} = \epsilon_{ij} - \epsilon_{ik}$, then we can rewrite Eq. (28) as

$$
\begin{aligned}
p_{Y_1} &= \Pr[\eta_{243} < U_2(Y_3) - U_2(Y_4), \eta_{131} < U_1(Y_1) - U_1(Y_3)] \\
&\quad + \Pr[\eta_{234} < U_2(Y_4) - U_2(Y_3), \eta_{141} < U_1(Y_1) - U_1(Y_4)] \\
&= \int_{-\infty}^{U_2(Y_3)-U_2(Y_4)} \int_{-\infty}^{U_1(Y_1)-U_1(Y_3)} \phi(\eta_{243}, \eta_{131}) \, d\eta_{131} \, d\eta_{243} \\
&\quad + \int_{-\infty}^{U_2(Y_4)-U_2(Y_3)} \int_{-\infty}^{U_1(Y_1)-U_1(Y_4)} \phi(\eta_{234}, \eta_{141}) \, d\eta_{141} \, d\eta_{233}. \tag{29}
\end{aligned}
$$

Both terms in Eq. (29) require integration over standardized bivariate normal densities. Given modern computing, this is not difficult. However, if we assume that the disturbances

are independent of each other, then the equation further simplifies to

$$
p_{Y_1} = \Phi\left[\frac{U_2(Y_3) - U_2(Y_4)}{\sqrt{\sigma^2_{\epsilon_{23}} + \sigma^2_{\epsilon_{24}}}}\right] \Phi\left[\frac{U_1(Y_1) - U_1(Y_3)}{\sqrt{\sigma^2_{\epsilon_{11}} + \sigma^2_{\epsilon_{13}}}}\right]
$$

$$
+ \Phi\left[\frac{U_2(Y_4) - U_2(Y_3)}{\sqrt{\sigma^2_{\epsilon_{24}} + \sigma^2_{\epsilon_{23}}}}\right] \Phi\left[\frac{U_1(Y_1) - U_1(Y_4)}{\sqrt{\sigma^2_{\epsilon_{11}} + \sigma^2_{\epsilon_{14}}}}\right]. \tag{30}
$$

Carrying out the same steps to calculate p_{Y_3} and p_{Y_4} gives

$$
p_{Y_3} = \Phi\left[\frac{U_2(Y_3) - U_2(Y_4)}{\sqrt{\sigma^2_{\epsilon_{23}} + \sigma^2_{\epsilon_{24}}}}\right] \Phi\left[\frac{U_1(Y_3) - U_1(Y_1)}{\sqrt{\sigma^2_{\epsilon_{13}} + \sigma^2_{\epsilon_{11}}}}\right] \tag{31}
$$

$$
p_{Y_4} = \Phi\left[\frac{U_2(Y_4) - U_2(Y_3)}{\sqrt{\sigma^2_{\epsilon_{23}} + \sigma^2_{\epsilon_{24}}}}\right] \Phi\left[\frac{U_1(Y_4) - U_1(Y_1)}{\sqrt{\sigma^2_{\epsilon_{14}} + \sigma^2_{\epsilon_{11}}}}\right]. \tag{32}
$$

Examining Eqs. (30)–(32), a number of issues are immediately apparent. First, comparing the above equations to the similar nonstrategic sequential choice Eqs. (12)–(15), we see that the strategic probability model is not a multinomial probit model, as in the nonsequential case with regressor error, even when we assume the disturbances are independent. Eqs. (30)–(32) reflect the observable utilities of two players, not one. In particular, the equilibrium conditions for outcome Y_1 are quite different in the two models, leading to different probability models. Therefore, if the data-generating process is the game shown in Fig. 2c and if the players have perfect and complete information, then employing any of the nonstrategic models will result in specification error.

Second, it was relatively easy to specify the equilibrium conditions and derive the probability model for the game shown in Fig. 2c. However, this becomes increasingly difficult as the complexity of the game increases. In fact, one does not need to add many more information sets for the specification to become much more involved.

Third, and related, the complexity of the underlying game will affect the dimensionality of the integration required for the equilibrium probabilities. In Fig. 2c, each player makes only a single decision. The sequence of choices results in bivariate normal densities, and the independence assumption allows us to write the equilibrium probabilities in terms of univariate normal cdf's. However, even if we assume the ϵ_{mk} disturbances are independent, more complex games (e.g., with more information sets) will result in probability models with multivariate normal densities of higher dimension, which could be computationally intensive. This is an issue not just of the number of choices at information sets (e.g., as in multinomial probit), but also of the depth of the game tree—specifically, the number of times a player makes decisions down a particular path.

4.3.2 Private Information About Outcome Payoffs

We now assume that a decision maker m's outcome utilities are private information: the analyst and other players know (or assume) only the distribution of m's outcome payoffs. Additionally, for the payoff disturbance and agent error models, we will continue to assume

the multiagent representation discussed in Section 4.2, where disturbances are private information to each individual agent, and where the disturbances are independent of each other.

To derive the choice probabilities, we will work "up the tree." The choice probabilities in the strategic model are derived in a similar manner to those in the nonstrategic sequential model. The probability of action a_4 is

$$p_4 = \Phi\left[\frac{U_2(Y_4) - U_2(Y_3)}{\sqrt{\sigma_{\pi_{24}}^2 + \sigma_{\pi_{24}}^2}} \right]. \tag{33}$$

Because player 1 is uncertain about player 2's payoffs, she must assess the probability that player 2 will choose a_3 versus a_4, and then use that probability in her expected utility calculations for her own options. The probability p_2 that player 1 chooses a_2 is

$$\begin{aligned} p_2 &= \Pr[U_1^*(a_2) > U_1^*(a_1)] \\ &= \Pr[p_3 U_1^*(Y_3) + p_4 U_1^*(Y_4) > U_1^*(Y_1)] \\ &= \Phi\left[\frac{p_3 U_1(Y_3) + p_4 U_1(Y_4) - U_1(Y_1)}{\sqrt{p_3^2 \sigma_{\pi_{13}}^2 + p_4^2 \sigma_{\pi_{14}}^2 + \sigma_{\pi_{11}}^2}} \right]. \end{aligned} \tag{34}$$

Because of the independence assumption, the outcome probabilities are the product of the action probabilities along their paths: $p_{Y_1} = p_1$, $p_{Y_3} = p_2 p_3$, and $p_{Y_4} = p_2 p_4$.

Notice that the structure of the model can be seen in (1) the expected utility comparison in the numerator of Equation 34, and (2) the form of the outcome probabilities. Additionally, the equilibrium probabilities above are very similar to the nonstrategic sequential probabilities under the assumption of independent agents. The difference between the two is that the equilibrium probabilities here are a function of two players' preferences, which the analyst will likely specify with different sets of regressors.

4.3.3 Agent Error

We now assume that players are boundedly rational. As in the previous case, to determine the strategic probit equilibrium outcome probabilities, we will proceed by first deriving the equilibrium action probabilities.

The probability p_4 that player 2 chooses action a_4 is

$$\begin{aligned} p_4 &= \Pr[U_2(Y_4) + \alpha_{24} > U_2(Y_3) + \alpha_{23}] \\ &= \Phi\left[\frac{U_2(Y_4) - U_2(Y_3)}{\sqrt{\sigma_{\alpha_{23}}^2 + \sigma_{\alpha_{24}}^2}} \right]. \end{aligned} \tag{35}$$

The probability p_2 that player 1 chooses a_2 must be written in terms of her expected utility for choosing action a_2. As the error term comes in through actions rather than outcomes, the expected utility is slightly different than with the payoff disturbances model:

$$\begin{aligned} U_1^*(a_{12}) &= U_1(a_{12}) + \alpha_{12} \\ &= [p_3 U_1(a_3) + p_4 U_1(a_4)] + \alpha_{12}. \end{aligned} \tag{36}$$

The agent's probability of choosing a_2 is then

$$
\begin{aligned}
p_2 &= \Pr[U_1^*(a_2) > U_1^*(a_1)] \\
&= \Pr[p_3 U_1(a_3) + p_4 U_1(a_4) + \alpha_2 > U_1(a_1) + \alpha_1] \\
&= \Phi\left[\frac{p_3 U_1(Y_3) + p_4 U_1(Y_4) - U_1(Y_1)}{\sqrt{\sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2}} \right].
\end{aligned}
\tag{37}
$$

The equilibrium outcome probabilities are, again, the product of the action probabilities along their paths.

A few observations are in order. First, each of the three strategic models examined here results in a different probability model. Hence, they are not observationally equivalent. The agent error and private payoff models are similar in structure. Their numerators are identical, reflecting the difference in the observed components of the expected utilities. The agent error and private payoff models differ only in the variance terms in the denominators of the probabilities. The variance term in the agent error model's denominator is constant across observations. Moreover, if we assume constant variance (say, σ^2) across disturbances, then a choice earlier in the game has the same amount of uncertainty as a choice later in the game, and the variance will be the same for each observation: here, $\sqrt{2\sigma^2}$.

In contrast, in a model with private information over payoffs, the variance term is a function of the action probabilities, and will therefore vary over observations. Additionally, for a given observation, the uncertainty in the private information model varies according to where a player is in the choice tree. An interesting aspect of this model is that the minimum possible uncertainty is less for choices made earlier in the game.[13] For example, consider Eqs. (33) and (34). Assuming constant variance σ^2 across disturbances, the denominator of Eq. (33) becomes $\sqrt{2\sigma^2}$, while the denominator of Eq. (34) is $\sqrt{\sigma^2(p_3^2 + p_4^2 + 1)}$. The latter will be less than the former for all values of $p_3 = (1 - p_4)$, except zero and one, in which case the two are equal.

It bears reiterating that each of the probability models examined here differ from their nonstrategic sequential counterparts because of the incorporation of multiple players' utilities and expected utility calculations. Therefore, the use of nonstrategic models to analyze data generated by strategic behavior will result in specification error.

Finally, so far the models have been expressed solely in terms of the players' utilities. However, to use these models for regression analysis, the utilities will need to be specified in terms of regressors. There is no set rule as to which regressors should be used to populate the utilities—that should be guided by theory. Moreover, as with nonstrategic discrete choice models, the analyst should pay close attention to model identification when populating the strategic models with constants and regressors.

5 The Type of Uncertainty Matters... To a Degree

We saw in the last section that the theoretical source of uncertainty sometimes generates different statistical models—specifically in the agent form of the nonstrategic sequential

---

[13]Using a centipede game with payoff disturbances, Zauner (1996, p. 15) makes the observation that the uncertainty associated with choices is higher earlier in the game than later. However, this appears to be driven by the structure and payoffs of the centipede game. In general, we can only say that the minimum possible uncertainty is less earlier in the choice tree.
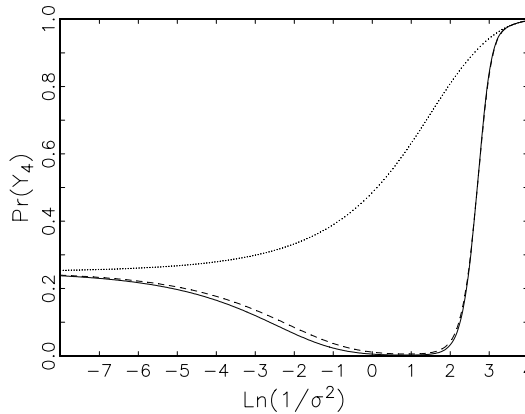
Fig. 3 Differences in equilibrium probabilities. The figure displays $\Pr[Y_4]$ for the private payoff model (solid line), agent error model (dashed line), and regressor error model (dotted line) as a function of σ^2. The far left represents almost complete uncertainty. The far right converges on the subgame perfect equilibrium of the complete information game.

models and in all strategic models. One implication is that, in some circumstances, we should be able to formulate tests to differentiate between the types of uncertainty. It also implies, however, that misspecifying the source of uncertainty may lead to incorrect inferences. In this section, I provide some (limited) sense of the extent to which the sources of uncertainty produce different probability models and affect our inferences.

One obvious question is: How much will the agent error and payoff disturbance probabilities differ, given the same observed utilities $U_m(Y_k)$? Under certain conditions, they will not differ at all. All three strategic probit models have the same limiting conditions. For simplicity, let us continue to assume that the variance is constant across disturbances. As the variance σ^2 of the error terms approaches zero, the choice probabilities approach the complete information SPE. As σ^2 approaches infinity, the observed utilities offer no information to distinguish which actions are more likely, and the action probabilities at information sets become uniformly distributed.

Where the models differ is in the choice probabilities between these limiting conditions. Although it remains that for the given model the choice probabilities tend to be relatively close—especially among the agent error and private information models—the differences in probabilities may at times be large, and will depend on the complexity of the model, on the payoffs, and on the "signal to noise" ratio of the regressors and disturbances.

Consider Fig. 3, which displays the probability of Y_4 based on the private payoff model (solid line), agent error model (dashed line), and regressor error model (dotted line), each as a function of the inverse of the disturbance variance.[14] For illustration, the observable utilities were assumed to be $U_1(Y_1) = 0$, $U_1(Y_3) = -10$, $U_1(Y_4) = 1$, $U_2(Y_3) = 0$, and $U_2(Y_4) = 0.5$. The far left and right of the graph correspond to the limiting conditions mentioned previously. On the far left, the observable utilities are negligible relative to the disturbances. Therefore, probabilities are uniformly distributed over the actions, resulting in a limiting probability of 0.25. On the far right, the uncertainty is vanishingly small, and the probabilities converge on the SPE of the game with perfect and complete information.

---

[14]That ratio is logged for easier interpretation of the graph.

In between, we see that the payoff and agent error models tend to have similar probabilities. However, depending on σ^2, they can differ significantly from the model with regressor error. For example, when $\ln(1/\sigma^2) = 0$, Fig. 3 shows that the private payoff model's $\Pr[Y_4]$ is almost zero, whereas the regressor error model's probability is about 0.5. Between $\ln(1/\sigma^2) = -2.3$ and $\ln(1/\sigma^2) = 2.3$, the model with regressor error deviates from the private payoff probability by anywhere from 0.2 to 0.85.

The reason for the deviation is fairly intuitive. The model with regressor error takes the theoretical model as fixed and incorporates uncertainty only into the statistical (or econometric) model—the disturbances are the parts of the true utilities that we do not capture with our regressors. Therefore, the regressor error model's probability varies monotonically from the state of complete ignorance of the true utilities (the far left side of the graph) to complete certainty about the utilities (the far right).

In contrast, the private payoff and agent error models incorporate uncertainty into both the econometric and theoretical aspects of the model. When there is complete certainty (as in the far right of the graph), both the analyst and player 1 know that player 2 will choose option a_4 (refer to the utilities above). Given that, player 1 will choose option a_2. However, with even a little uncertainty concerning the payoffs (or actions, as in the agent error model), player 1's expected utility calculation heavily leans in favor of option a_1. This is because the observable component of player 2's utilities are relatively close, and player 1's utility for Y_3 is much worse than for Y_1. Hence, there is a nonmonotonic relationship between $\Pr(Y_4)$ and σ^2.

The model in Fig. 2c is rather simple. Under most circumstances, the private payoff and agent error probabilities should be relatively close in this model. However, that is not to say that private payoff and agent error models will always yield virtually identical probabilities. It is fairly easy to construct models where there is more room for divergence, and the reason for the divergence is, again, fairly intuitive. As they share the same numerator, the only way in which corresponding action probabilities can differ is in the denominator. The denominator of the agent error action probabilities is the same, regardless of observation or the information set in the model, assuming disturbances are independent and identically distributed. However, for a given information set, the denominator of the action probabilities in the private payoff model will be a more complicated function of the action probabilities (1) the closer it is to the root node, (2) the deeper the game tree (in terms of sequences of moves), and (3) the wider the tree (in terms of number of options at information sets).[15] Therefore, for the private payoff versus agent error models, although the theoretical source of the uncertainty may not make much difference in "smaller" or "simpler" choice trees, it will be important for more complex choice structures.

5.1 *Bias and Inconsistency*

To assess the effect on our inferences of incorrectly modeling the uncertainty, I conducted a series of Monte Carlo simulations. For the data-generating process, I employed the game in Fig. 2c with the following assumptions:

- The utilities are specified as

$$U_1(Y_1) = 0, \quad U_1(Y_3) = \beta_{13}X_{13}, \quad U_1(Y_4) = \beta_{14}X_{14},$$
$$U_2(Y_3) = 0, \quad U_2(Y_4) = \beta_{24}X_{24}.^{16}$$

---

[15] A more detailed example is available upon request from the author.

[16] $U_1(Y_1)$ and $U_2(Y_3)$ are normalized to zero only for simplicity. The conclusions would not change were we to specify them differently.

- $\beta_{13} = \beta_{14} = \beta_{24} = 1$.
- The uncertainty is due to private information concerning payoffs, where the disturbances are normally distributed with mean zero and variance $V(\pi)$.
- The regressors are each uniformaly distributed with mean zero and variance $V(X)$.

In each iteration of the Monte Carlo simulation, a sample of 2000 observations was generated using the behavioral assumptions of the private payoff model. For a given sample, the private payoff model, the agent error model, and the regressor error model were each estimated, and the estimates were saved. This was repeated 2500 times.

It seems likely that the specification error will depend in part on the extent to which the disturbance "matters" relative to the observable component of the utility. In that regard, Monte Carlo analyses were conducted as detailed above for the following signal-to-noise ratios $V(X)/V(\pi) \in \{\frac{1}{10}, \frac{1}{3}, 1, 3, 5, 7, 10, 13, 15\}$.[17] So, for example, $V(X)/V(\pi) = \frac{1}{10}$ denotes a Monte Carlo experiment where the variance of the disturbance is 10 times that of the regressors, representing generally greater uncertainty. Similarly, an experiment with $V(X)/V(\pi) = 10$ denotes a Monte Carlo experiment where the variance of the regressors are 10 times that of the disturbances, representing generally greater certainty.

Figures 4a and 4b display the estimated mean squared error of the estimators of β_{14} and β_{24}, respectively, as a function of $V(X)/V(\pi)$.[18] The solid line represents the private payoff model, the dashed line represents the agent error model, and the dotted line represents the regressor error model. As Fig. 4a shows, the private payoff estimators have the smallest estimated mean squared error for $\hat{\beta}_{14}$. This is not surprising, as the private payoff model was used to generate the data. The figure shows that $\widehat{\mathrm{MSE}}(\hat{\beta}_{14})$ decreases as $V(X)/V(\pi)$ increases. Not surprisingly as well, the agent error model closely mirrors the private payoff results, albeit with a slightly higher MSE in every case.

Somewhat more surprising is the MSE for the regressor error model (the dotted line). As Fig. 4a displays, the MSE initially falls as $V(X)/V(\pi)$ increases, until $V(X)/V(\pi)$ is somewhere around 2 or 3, at which point it begins increasing in $V(X)/V(\pi)$. This appears to be due to two dynamics. Recall that $\mathrm{MSE}(\hat{\beta}) = V(\hat{\beta}) + \mathrm{Bias}(\hat{\beta})^2$. Although not shown here, the results of the Monte Carlo analysis indicate that the bias decreases until about $V(X)/V(\pi) = 3$ and then starts increasing.[19] As $V(X)/V(\pi)$ increases, the variance accounts for a smaller portion of the MSE than does the bias. However, there is also a small upturn in the variance at around $V(X)/V(\pi) = 7$. Aside from that, it is unclear why we see this particular nonmonotonic dynamic in Fig. 4a. My conjecture is that there is an "optimal" positive amount of uncertainty in estimating strategic models. Too much certainty (e.g., too high a $V(X)/V(\pi)$) produces probabilities too close to zero and one, which in turn produce ill-behaved log-likelihood functions and perhaps computer precision problems, making it difficult for statistical computer programs to converge on a maximum.
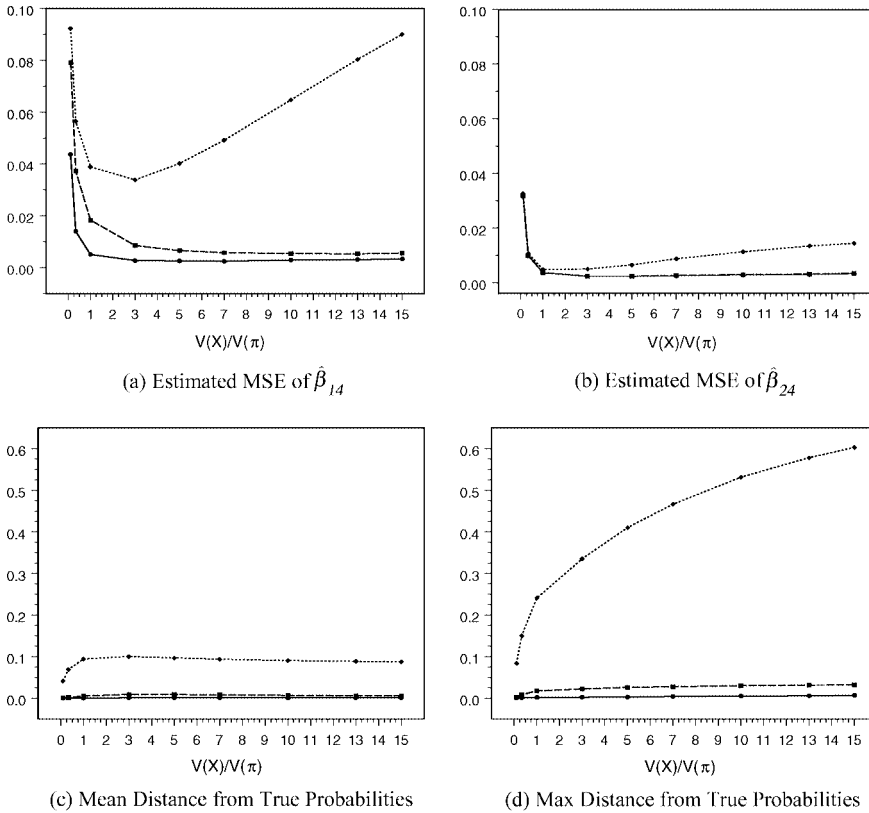
Similar results are seen in Fig. 4b. The MSE for the private payoff model is negligible, except for the lowest $V(X)/V(\pi)$. The agent error model mirrors that. And, we see a similar nonmonotonic relationship in the MSE for the model with regressor error.

All of these results strongly suggest that misspecifying the source of uncertainty results in biased and inconsistent estimates when the estimated model is observationally

---

[17]Perhaps "systematic versus stochastic ratio" would be a more accurate term.

[18]The results for $\hat{\beta}_{13}$ are essentially the same as those for $\hat{\beta}_{14}$.

[19]Results available upon request from the author.

(a) Estimated MSE of $\hat{\beta}_{14}$

(b) Estimated MSE of $\hat{\beta}_{24}$

(c) Mean Distance from True Probabilities

(d) Max Distance from True Probabilities

Fig. 4 Measures of deviation from the true model. In each of the graphs, the private payoff model is assumed to be the data-generating process. Parts (a) and (b) show the estimated mean square error of $\hat{\beta}_{14}$ and $\hat{\beta}_{24}$, respectively, for the three estimated models. Parts (c) and (d) provide a sense of how far off on average and in the worst case, respectively, each model will be in estimating outcome probabilities. The solid line denotes the private payoff model, the dashed line denotes the agent error model, and the dotted line denotes the regressor error model.

nonequivalent with the data generating process. Although the MSEs in Figs. 4a and 4b are fairly small, we would expect the MSEs to increase as the complexity of the data-generating process increases, as the probabilities will diverge as the complexity increases.

5.2 A Broader Comparison of the Estimated Probability Models

Although the signs of parameter estimates are readily interpretable in some discrete choice models, they are not in others. In most cases, to determine the effect of a regressor, researchers interpret the effect of a variable on an outcome probability of interest. It would be nice if we had some measure, like MSE, with which we could assess the extent to which one model broadly diverged from another model in their estimated probabilities.

To do this, I employ the same technique as in Signorino (2002). First, consider a grid of points over the space of our regressors X_{13}, X_{14}, and X_{24}. For any given range of the regressors, suppose we constructed a grid with 50 points along each dimension. That would result in a total of $50 \times 50 = 125,000$ points, evenly distributed across the regressor space. Now consider the private payoff data-generating process and suppose we calculated the

outcome probabilities for each point in that grid. Associated with each point in the grid are three probabilities: ($\Pr[Y_1]$, $\Pr[Y_3]$, $\Pr[Y_4]$). We can treat that triple as a point in a simplex. Now, we can do the same thing with one of our estimated models, say the agent error model. For each point in the regressor grid, we would have the probability triple for the "true" model and the estimated probability triple from the agent error model. To determine how close those points are to each other, I simply take the Euclidean distance between them in the simplex.[20] For each point in the regressor grid, we then have a measure of the "distance" between the two models. We can summarize the distance of the two models (i.e., the true model and the estimated model) in a number of ways.

Figures 4c and 4d provide some sense of how far "apart" the three estimated models are from the data-generating process. As before, the solid line denotes the estimated private payoff model, the dashed line denotes the estimated agent error model, and the dotted line denotes the estimated regressor error model. Figure 4c displays the average distance across the regressor grid for each of the three models, and as a function of $V(X)/V(\pi)$. As one can see, the private payoff and agent error models are extremely close on average to the data-generating process in terms of their estimated probabilities. The average distance between the data-generating process and the estimated regressor error model is generally about 0.1.

Perhaps more striking are the results of Fig. 4d, which shows the maximum distance between the estimated model and the data-generating process. Interestingly, the agent error model's maximum distance is fairly small. On the other hand, the maximum distance for the regressor error model grows with $V(X)/V(\pi)$, and is quite large for all but the very smallest $V(X)/V(\pi)$.[21] For more complex games, one would expect the mean distance to increase for the agent error and regressor error models. Moreover, one would expect the maximum distance to increase, or at least not decrease.

The import of this section then is that the type of uncertainty matters in terms of our statistical inferences, at least in strategic models. The reader should not conclude from this, however, that the private payoff model is always preferred. Those results were due to the fact that the private payoff model was also the data-generating process. If the data were instead generated by a game of perfect and complete information, where we as analysts simply did not correctly account for all variation in utilities via our regressors, then the private payoff statistical model would have produced incorrect inferences. The point to take from this section is that assuming the wrong type of uncertainty can lead to biased and inconsistent parameter estimates and estimated probabilities that are far off the mark.

6 Concluding Remarks

My objectives in this article have been to (1) examine how three plausible sources of uncertainty interact with structural assumptions to produce different statistical models and possibly affect our inferences, (2) demonstrate how to derive such statistical models from theoretical first principles, and (3) generalize a broad class of statistical discrete choice models.

As we have seen, nonstrategic sequential choice models with complete information reduce to multinomial choice with a single information set and are observationally equivalent to the nonsequential models. Assuming agent errors or payoff disturbances are private

---

[20]The farthest two points can be from each other in the simplex is if one is at one corner of the simplex and the other is at another corner, resulting in a distance of $\sqrt{2} \approx 1.414$.

[21]Indeed, the maximum distance is small for extremely small $V(X)/V(\pi)$ only because that reflects almost complete uncertainty, which has the same limit for all three models.

information to a player's agents can lead to different statistical models that are observationally nonequivalent. These latter models have many similarities to the strategic models and can actually be thought of as strategic, where agents are different players but with the same observed outcome payoffs.

Strategic choice results in more complicated models, as the choice probabilities reflect not just the utilities of a single decision maker, but the utilities of multiple decision makers who must condition their behavior on the expected behavior of each other, producing equilibrium choice probabilities. In this case, misspecifying the type of uncertainty leads to biased and inconsistent estimates.

As I have demonstrated, constructing a strategic discrete choice model is not difficult. From a practical perspective, however, some of the probit-based examples examined here are subject to the same constraint as multinomial probit: integration of higher dimension normal densities. Depending on the depth of the choice tree, the strategic probit model with regressor error may be computationally impractical. In fact, any of the probit-based models may be computationally impractical if the number of choices at any information set becomes too large. Nevertheless, most social science strategic models tend to limit the number of choices available at information sets, so the strategic probit models presented here should be widely applicable. Moreover, logit versions of the strategic models are also derivable (e.g., LQRE in McKelvey and Palfrey 1998; see also Signorino 1999).

Much, of course, remains to be done. First, the different variance terms in the strategic probit models—and the fact that one varies over observations while the other is constant—suggest that a specification test can be developed for the different models. A more general strategic probit model that combines agent error and payoff perturbations is easy to derive and may be useful in that respect.[22] However, combining the regressor error assumption with either of the other two types of uncertainty is much more complicated, as it requires differentiating the information available to the players versus the analyst.

Second, as I alluded to earlier, the assumption that disturbances are independent of each other, in combination with the multiagent representation and agent-level private information, is not innocuous. It greatly simplifies deriving the choice probabilities, but at a price: players cannot update in the model. Under these assumptions, a player's agent knows nothing about that player's other agents' disturbances (at subsequent information sets) beyond what the opposing player knows. The agent's choice is not based in any way on those disturbances, other than knowledge of their distribution. Therefore, the opposing player cannot learn from that agent's choice anything about subsequent agents' disturbances.

Again, there may be situations where this is theoretically justified. With payoff disturbances, we might assume that players are not completely certain what their utilities will be at later decisions, but have a pretty good idea what they are likely to be on average. The assumption is perhaps better justified in the case of agent error: players are unlikely to foresee their own misperceptions or mistakes at future decision points—otherwise they could correct for them ahead of time. These assumptions have the virtue of making the math tractable. On the other hand, they create a different—and much more limited—type of incomplete information game than those typically analyzed in the mainstream game theory literature. The assumption of independent, private disturbances could be relaxed. The price for doing so is a more complicated model.[23] Whether the complication is

---

[22]A combined payoff disturbance *and* agent error strategic probit model for Fig. 2c is identical to those in Section 4.3, except that the denominators include the variance terms of both error models.
[23]For one example, see Signorino (2002).

warranted, of course, is an (interesting) empirical question, but it is also a topic for another article.

References

Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science*. 42:55–96.

Chen, Hsiao Chi, James W. Friedman, and Jacques-Francois Thisse. 1997. "Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach." *Games and Economic Behavior* 18:32–54.

Hausman, Jerry A., and David A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice." *Econometrica* 46:403–426.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

McFadden, D. 1974a. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, ed. P. Zarambka. New York: Academic Press, pp. 105–142.

McFadden, D. 1974b. "The Measurement of Urban Travel Demand." *Journal of Public Economics* 3:303–328.

McFadden, D. 1976. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement* 5:363–390.

McKelvey, Richard, and Tom Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *The Japanese Economic Review* 47(2):186–209.

McKelvey, Richard, and Tom Palfrey. 1998. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics* 1:9–41.

Myerson, Rober B. 1991. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge, MA: MIT Press.

Pudney, Stephen. 1989. *Modelling Individual Choice: The Econometrics of Corners, Kinks, and Holes*. Oxford: Basil Blackwell.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93:279–297.

Signorino, Curtis S. 2002. "Strategy and Selection in International Relations." *International Interactions* 28:93–115.

Signorino, Curtis S., and Kuzey Yilmaz. 2003. "Strategic Misspecification in Regression Models." *American Journal of Political Science* 47:551–566.

Smith, Alastair. 1999. "Testing Theories of Strategic Choice: The Example of Crisis Escalation." *American Journal of Political Science* 42:1254–1283.

Zauner, Klaus G. 1996. "A Payoff Uncertainty Explanation of Results in Experimental Centipede Games." The University of New South Wales, Australian Graduate School of Management. Working Paper 96-030.