Estimating hospital choices when the true choice set is unknown

Andrew Sfekas

June 1, 2016

Abstract

Random utility models are frequently used in the health care economics literature to estimate a patient's probability of choosing a hospital. A key assumption of such models is that individuals choose the hospital that would yield the highest utility, given a known set of alternatives. However, managed care networks sometimes do not include all hospitals in a local area, so an individual's choice may reflect both the utility of the hospital and the unobserved limits on the choice set. This paper describes a method of estimating patient choice based on random utility when some sets of hospitals have some probability of being unavailable. The model is tested using both simulated data and data on hospital choices in several markets in Florida.

1 Introduction

The random utility choice model described by McFadden (1974) is a workhorse for a number of settings. It has been used to examine hospital choice (**?**), ambulatory care center choices (Weber 2014), print and online newspaper choices (Gentzkow2007), supermarket goods (Dillon and Gupta 1996), and several other settings. The model offers a number of advantages. It is flexible enough to fit a number of different problems, while remaining empirically tractable. These advantages come with some significant costs, however: there are several unrealistic model assumptions that can pose problems in a number of circumstances (**?**; Shocker et al. 1991).

One fundamental problem with the choice model is described by Manski (**?**). The choice problem observed by the researcher is actually the final step in a process involving

the selection of individuals and the selection of choices. The probability of making a particular choice therefore depends on the probability that it is the best option out of the given set of options and the probability that a given set of options is available to a specific set of individuals. A number of problems may arise. For example, some individuals may have access to all available choices, while others are locked into one single selection (Swait and Ben-Akiva 1987), or individuals may only select from a subset of options because they cannot hold all of the options in their memory at once (**?**).

One particular setting in which choices may be limited is the choice of a hospital for inpatient care. A patient choosing a hospital may appear to have the choice of any hospital within a reasonable driving distance. However, the patient's available choice set is limited to those covered by the patient's insurance plan. Managed care insurance plans frequently do not contract with all providers in a market, and in some markets up to 40% of the possible plan-provider combinations do not result in a contract (Ho 2006). A possibility, in that case, is that an individual choosing hospital i may have wanted to choose hospital j–hospital j actually yielded a higher utility–but could not because hospital j was not present in the choice set. The individual might, for example, want to visit the closest hospital, and j is the closest hospital. Since hospital j is not actually present in the choice set, that individual had to choose a more distant hospital, making it appear that distance was not as important to the choice.

Several previous studies of hospital markets have dealt with this problem by limiting analyses to patients who are expected, *a priori*, to have their choice of most hospitals in the local area–specifically, patients with PPO or commercial indemnity insurance, rather than HMO insurance (e.g. Capps, Dranove, and Satterthwaite 2003; Ho 2006). An additional factor that could help mitigate the problem is that individuals may choose their managed care plan based in part on whether their preferred hospital is available (Ho 2006). However,

the assumption that limiting the sample to PPO and commercial indemnity patients is sufficient is generally not tested. Additionally, commercial indemnity insurance is rare (**?**), and even PPO patients may face limited networks in some markets.

This study describes and implements an approach to estimating a choice model when some choices may be absent. The approach retains the assumptions of McFadden's (1974) random utility conditional logit model, but allows for the possibility that certain blocks of choices are absent for some individuals. The model is structured to fit the market for hospital inpatient services, where hospitals in systems (such as HCA or Tenet) are likely to bargain as a unit (**?**), and may therefore have a common probability of being absent. I test the model for a variety of specifications using simulated data. I then estimate the model in a real-world market: hospitals in south Florida for the years 1995, 2000, and 2005.

The proposed method works as follows. Individuals will have some set of choices that is always present (though this set need not be the same for all individuals). The remaining choices are grouped into nests, each of which has a certain probability of being present in the individual's choice set. The probability of a choice in the always-present nest is the probability that it is the highest-utility choice in the highest-utility nest, plus the probability that it is the highest-utility choice in the second-highest utility nest and the highest-utility nest is missing, and so on until the choice is in the lowest-utility nest and all of the other nests are missing. For a choice in a set that is not always present, the probability is adjusted to account for the fact that the choice is not always available. The parameter estimates will consist of the standard coefficient estimates and a vector of probabilities.

The results from simulated data demonstrate the effectiveness of the approach and its limitations. When the model assumptions are true, the proposed model yields more accurate coefficient estimates than a standard conditional logit. Further, the probability that a

block of options is present can also be estimated. The largest difference in estimates comes from variables that only vary across blocks, and not within them. In a real application, this could be something like the ownership status of a hospital system (nonprofit, for-profit, or government), which would not vary within the system but would vary across systems. On the other hand, the model performs poorly when it includes choice fixed effects and some blocks are small. For example, in the simulations where a single choice had its own associated probability, and the model included fixed effects, the estimated probability was about one–the combination of fixed effect and probability ended up in the fixed effect. This result is not surprising, since there was little to separately identify the parameters in that case. However, this problem can be reduced substantially through the use of a ridge RIDGE CITE HERE. When the ridge was added, the model estimates came close to the true estimates, though standard errors could be fairly large in some cases.

The simulations established that the model performs better than conditional logit in some circumstances and no worse than conditional logit in others, which suggests that it can be applied to actual choice data. The results from hospital choice data suggest that, in the years examined, some sets of choices were sometimes absent. However, absences were greatest for commercial indemnity patients, not for HMO patients–a result that is contrary to common practice. Probabilities of being present were generally in a reasonable range when hospital fixed effects were included (though not when only hospital characteristics were included).

Several previous studies have examined the random utility choice model, through both theoretical and empirical lenses. An early study by Manski (**?**) develops the basis of the the model used in this study. In Manski's framework, choice probabilities represent a series of steps: first, the set of individuals and the set of choices are selected; and second, the individuals choose from the options determined in the first stage.

The marketing and transportation literatures have engaged with the problem of limited choices in the context of "consideration sets": the options that an individual actually considers, out of the full range of potential options (Shocker et al. 1991). Consideration sets may arise if, for example, some individuals are "captive" to a particular choice, i.e. if they only have access to that choice (Swait and Ben-Akiva 1987). In that case, the choice model is a mixture in which the captive fraction is estimated along with the conditional logit parameters. Consideration sets may also arise when a consumer cannot readily hold all available options in memory (Nedungadi 1990), in which case the choice set may change depending on environmental conditions. Consideration sets can limit the effects of price changes, since changes to products outside of a consideration set will not affect consumer choice (Bronnenberg and Vanhonecker 1996).

The model presented here is also similar to models with misclassification error. Hausman *et al.* (**?**) describe how to deal with misclassification error in a binary choice model by explicitly estimating the misclassification probabilities (false positive and false negative). A more closely related problem (though not a more closely related method) is the multinomial logit with misclassification examined by Poterba and Summers (1995). In their case, the choice set is known, while the individual's actual choice is subject to misclassification error. In the problem presented here, the individual's actual choice is known, but the choice set is not. In the case of both the Hausman *et al.* and the Poterba & Summers models, the misclassificatiion probabilities are independent of individual characteristics. The same assumption will be made here.

A final related strain of the literature on choices deals with the question of how individuals learn about the options in their choice sets. Moorman et al.: people search near prior beliefs for choice info.

The paper proceeds as follows. Section 2 describes the basic random utility model,

in which one set of options is present with certainty for all individuals. This section also discusses a method of testing the model assumptions, and a simple way to determine if the misclassification model is necessary. Section **??** then adds the complication that different groups of individuals may have different sets of probabilities and a different reference (always present) group. Section 3 presents the results of a number of simulations, designed to test the model's performance for a range of different specifications. Section **??** presents the analysis of the south Florida hospital market. Finally, section 6 concludes with a discussion of the model implications.

2 The model

The model employed here follows the random utility framework described by McFadden (1974). This framework is used extensively to describe hospital choices, as well as individual choices for other types of discrete alternatives. The framework begins with the assumption that people assign utilities to different choices according to a linear utility function:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

where V_{ij} is a stable preference term that will be a function of individual and choice characteristics, and ε is an idiosynchratic preference term. The individual will choose the alternative j that yields the highest level of utility, given the alternatives. The specific functional form is usually the following:

$$U_{ij} = Z_{ij}\beta + \delta_j + \varepsilon_{ij} \tag{2}$$

where Z is a set of variables that vary across individual choices and δ_j is a fixed effect for choice j. The probability that the individual makes a particular choice j is:

$$P_{ij} = Pr\left(U_{ij} > U_{ik} \quad \forall k \neq j; k, j \in J\right) \tag{3}$$

where J is the choice set. If ε is i.i.d., P_{ij} is simply equal to the product of the probabilities that j is greater than each of the other k options:

$$P_{ij} = \int_{-\infty}^{\infty} f(\varepsilon_{ij}) \prod_{k \neq j} \left[F(V_{ij} - V_{ik} + \varepsilon_{ij})\right] d\varepsilon_{ij} \tag{4}$$

If ε follows the double exponential distribution, this becomes the conditional logit framework, and the probability simplifies to:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{k \in J} \exp(V_{ik})} \tag{5}$$

The above is McFadden's (1974) choice model, in which the parameters of the utility function are estimated based on observations of the individual's choice set and the choice the individual actually made from those alternatives. Excluding a choice that is actually available does not present a problem for the model–the same parameter values will be estimated. This is the assumption of the independence of irrelevant alternatives (IIA). On the other hand, including options that are not available does present a problem for estimation. Consider what would happen if the researcher were to add an option that would yield high utility, but that is not available to most of the individuals in the sample. Few people would choose the option, even though its utility value is very high. According to the model assumptions, this would not mean that it was unavailable, but that it yielded a lower utility than the other options. The result is that the estimated utility for that

option would be biased downward.

The estimation approach presented here begins with the random utility framework. It next adds the possibility that for a given individual, one set of choices is available all the time, while other sets of choices are available with probabilities q_k. For example, there might be 3 choices available to everyone, 2 choices which are present or absent as a pair, and one choice which is present or absent on its own. Individuals with all choices available will pick the choice that offers the highest utility. Individuals with some choices missing may not be able to select the choice with the highest utility; instead, they choose the highest available utility.

As a concrete example, consider a patient choosing a hospital, in a city with six hospitals. While there are six hospitals present, the patient's insurance plan may only cover 3 or 4 of them in a given year.

The discussion begins with the simplest case of such a problem: one set of choices has a single probability q of being missing from the choice set, i.e. some choices are always present and the rest are either all there or all missing with probability q. In that case, some individuals who would like to choose j will not be able to, because j is absent from their choice set. Define the full choice set as L, consisting of the set K that is always present and the set J that is either present or missing as a block. Utility will be the utility function in equation 2. One way to estimate the β coefficients (and at least some of the δ coefficients) would be to throw out the choices in J, as well as the individuals who chose an option from J. One could then estimate the model using a standard conditional logit model:

$$Pr(k|K) = \frac{e^{Z_{ik}\beta+\delta_k}}{\sum_{m\in K} e^{Z_{im}\beta+\delta_m}} \tag{6}$$

The above conditional probability would not contain any of the terms that bias the estimates when the options in J are used. Hausman (**?**) uses this approach to develop a test

of the IIA assumption for the conditional logit model. The estimates from the restricted choice set will be consistent. However, the probability q that the J choices are absent is also of interest, since it may tell us something about market structure. Additionally, any variables that do not vary within set K (such as fixed effects for choices in J) will not be identified, so it would not be possible to estimate $P(j \in J)$.

Given equation 6, one can see a way to get get q and the parameters that do not vary withn K. The probability of choosing an option j in K, the part of L that is absent with probability q, would be:

$$P_j = q\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L} e^{Z_{ik}\beta+\delta_k}} \tag{7}$$

i.e. it is the probability that K is present (and therefore all choices are present) and that j is the best choice out of all options in L. Alternatively, if j were in $L \setminus K$, then it is either the best choice out of all options in L or the best choice out of all options in $L \setminus K$ (but possibly not better than some options in K). In that case, the probability of choosing j is:

$$P_j = q\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L} e^{Z_{ik}\beta+\delta_k}} + (1-q)\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L\setminus K} e^{Z_{ik}\beta+\delta_k}} \tag{8}$$

Equations 7 and 8 actually have parallel structures, since there is a probability associated with nest L: it is assumed to be one. The logit probability in the first term in equation 7 is actually multiplied by $q \cdot 1$, and the second, which is absent from equation 7 is multiplied by $(1-1) = 0$. This is useful for thinking about how the model extends to additional nests.

Coefficients will be identified in the above model, even if they do not vary within nests. Consider a set of variables Z that only vary across nests. The Z variables will disappear from the second term in equation **??**, but not the first term. The parameter q, on the other hand, is present in both terms. Thus, q and the coefficients on Z are present in different

locations in the estimation equation; it is not possible to optimize the equation by sending one set of parameters to $-\infty$ and the other to ∞.

2.1 Multiple nests

The above framework extends to multiple nests in a fairly straightforward manner. Equations **??** and **??** show that the probability of choosing a hospital is the sum over all possible choice sets of the probability that the hospital is the best choice out of a particular choice set, multiplied by the probability that the individual has that choice set. If there are three different groups (K_0, K_1, and K_2), and group K_0 is always present, then there are four possible choice sets: all present, one or the other missing, and both missing. Retaining L as the set of all choices, the above equations would become:

$$P_j = q_1 q_2 \frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L} e^{Z_{ik}\beta+\delta_k}} + q_2(1-q_1)\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L\setminus K_1} e^{Z_{ik}\beta+\delta_k}} \tag{9}$$

$$+ q_1(1-q_2)\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L\setminus K_2} e^{Z_{ik}\beta+\delta_k}}$$

$$+ (1-q_1)(1-q_2)\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L\setminus\{K_1,K_2\}} e^{Z_{ik}\beta+\delta_k}} \tag{10}$$

when j is in the main nest, and:

$$P_j = q_1\left(q_2\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L} e^{Z_{ik}\beta+\delta_k}} \tag{11}$$

$$+ (1-q_2)\frac{e^{Z_{ij}\beta+\delta_j}}{\sum_{k\in L\setminus K_2} e^{Z_{ik}\beta+\delta_k}}\right) \tag{12}$$

when j is in nest K_1. For j in K_2, the equation is analogous to the above. The above are the sums over all configurationss of the probability that the configuration happened

multiplied by the probability that j was the best choice for that configuration.

Adding nests creates some more complicated notation, but the idea remains the same:

$$P(j) = \exp X_{ij}\beta \left(\prod_{K \in L \setminus J} q_K \left(\frac{1}{\sum_{l \in L} \exp X_{il}\beta} \right) + \sum_{m=1}^{|L|} \prod_{k \in K \subset L \setminus J \ni |K|=m} q_j \prod_{n \in L \setminus J, K} (1 - q_l) \frac{1}{\sum_{l \in L} \exp X_{il}\beta} + \prod (1 \right. \tag{13}$$

$$P(j) = q_J \exp X_{ij}\beta \left(\prod_{K \in L \setminus J} q_K \left(\frac{1}{\sum \exp X_{ik}\beta} + \sum \prod q_j \prod (1 - q_l) \frac{1}{\sum \exp X_{ik}\beta} + \prod (1 - q_k) \frac{1}{\sum \exp X_{ik}\beta} \right. \right. \tag{14}$$

The above equations retain the basic intuition of the less complicated 3-nest case. The probability of choosing hospital j is the sum over all possible cases of the probability that a set of nests is in the market multiplied by the probability that j is the best choice within that set. When the choice is not in the base set, a few terms drop out because the probability of the base set being present is assumed to be one.

2.2 Improving estimation by adding a ridge

As the simulations will demonstrate, some coefficients will be poorly identified under certain circumstances. When choice fixed effects are included, the estimated qs are sometimes very large–well over 2 for a significant percentage of the simulations. Not surprisingly, the problem is particularly acute when a nest has only one or two choices in it. A partial solution would be to restrict q to be less than or equal to 1, but that would cover up the problem, rather than solve it.

Instead of restricting the values of q, I use a ridge to impose a penalty for large parameter values (Hoerl and Kennard 1970). With the ridge term added, the log likelihood

becomes:

$$LL = \sum_{i \in N} \ln P_{ij} + \lambda \theta' \theta \tag{15}$$

where:

$$\theta = (1 - q, \beta)' \tag{16}$$

i.e. θ is a vector of all model parameters, but with $1 - q$ substituted for q, and λ is a parameter supplied by the researcher that determines the size of the penalty. The effect of the ridge is to shrink the coefficients towards zero (or towards 1, in the case of the qs). This reduces the variance of the coefficients, at the cost of some bias. As λ is increased, bias will increase and variance will decrease; for very high values of λ the model would simply predict the same probability for all choices. Initially, however, at low levels of λ, the bias can be very small in comparison to the reduction in variance.

3 Monte Carlo simulation

This section presents the results of several Monte Carlo simulations, intended to demonstrate the above model under a range of different circumstances. There are several possible situations to address: (1) one occasionally missing set of choices vs. several missing sets; and (2) models with sets of choice characteristics (e.g. product type or size) vs. models with choice fixed effects. I create several datasets to show when the model works well and when it can be expected to fail or work poorly.

All datasets will include measures of the linear distance between individual and choice. This distance is constructed by assigning two uniform random variables (varying between 0 and 1) to each individual and each choice. These variables represent a point in a 1-by-1 square. The distance will then be the length of the line between them.

I begin with the simplest case: one set of misclassified choices. A draw from a uniform

distribution for each individual will determine whether the nest is present. The remaining choices will be present with certainty.

The paired datasets will consist of 3 individual characteristics each: one binary variable (which could represent gender, e.g.), one integer variable that varies between 1 and 80 (which could represent age), and one continuous variable that varies between 1 and 100. It also has two choice-specific variables: one indicator variable and one integer variable that varies between 1 and 100. The choice-specific indicator variable will vary within each nest. Finally, it has one distance term, which will differ across individuals and choices based on their locations in the square. The data consist of 60,000 individuals and 6 choices, with two of the choices in the nest that is sometimes missing. Table **??** contains the true parameters and estimation results from conditional logit and misclassification logit models. Columns (2) and (3) contain estimation results from the dataset with double exponential idiosyncratic terms, while columns (4) and (5) contain estimation results from the dataset with normally distributed idiosyncratic terms.

— TABLE **??** ABOUT HERE —

The next set of datasets involves four nests, but is the same as the above in terms of choices, variables, and observations. Since one nest is present with probability 1, three probabilities will be estimated. In this case, I treat the nests as though they are hospital systems, so the indicator variable will not vary within nests–it will be identified only by cross-nest variation. As in the case of the one-missing-group results, the method was able to get close to the true parameters, and the main difference with conditional logit was in the estimation of the NFP indicator (which only varied across groups).

— TABLE **??** ABOUT HERE —

The next pair of datasets is similar to the above in number of choices, nests, and observations, but in this case the true model consists of one distance measure, a set of

individual-choice interactions, and a set of choice fixed effects. The groups are set up to push the limits of the estimation technique: the base group and group 1 contain 3 choices each, group 2 contains 2 choices, and group 3 contains only 1 choice. For group 3 in particular, this should lead to fairly poor identification of some parameters. With this in mind, I add a ridge to the model, with the ridge parameter λ set to several different values.

Table **??** contains the true parameters and estimation results from these models. When the ridge parameter is 0, the misclassification model performs poorly.

— TABLE **??** ABOUT HERE —

As noted in section XXXXXXX, the true parameters are obviously not known outside of the simulated results.

4 Other ways of modeling the probability of being in the choice set

The model described so far in this paper has been relatively simple in the assumptions on the probabilities of being in the choice set. With independent probabilities, the absence of one choice does not affect the likelihood that another choice is absent. In a number of cases, this may not be a reasonable assumption. For example, in hospital markets, the omission of one hospital or set of hospitals from a managed care network may affect the value of the remaining hospitals. In that case, the probability that two nests are absent from the network would not be equal to the product of their associated probabilities.

In some cases, there may be a well-defined theoretical model that suggests a structure for the exclusion probabilities. There are several structural models currently in use to describe managed care hospital networks. When such a structure exists, it may be possible to impose additional restrictions on the probabilities, allowing for overidentification of the

model.

Perhaps the simplest alternative approach to dealing with this problem is to estimate the model using a mutually exclusive set of categories, corresponding to each potential configuration. This would involve a potentially large number of probabilities to be estimated, but would be the most flexible approach to the problem. Restrictions could then be imposed and tested afterwards through an approach like that in Chamberlain (1982).

5 Application: Hospital choices in southern Florida

The simulations showed that the model could produce estimates of the β parameters and the inclusion probabilities, though it was not well-identified in the hospital fixed effects models. This section applies the model to a real-world set of data: hospital choices in south Florida. I choose the years 1995, 2000, and 2005 to demonstrate. Between these two years, two hospital systems merged. One assumption of hospital merger simulations (e.g. CITES) is that all hospitals in a system bargain as a single unit with managed care organizations. Thus, in 2000 and 2005, hospitals in the merged systems should have the same probability q, while in 1995, they should not.

5.1 Data sources

Data for this part of the study come from Florida hospital inpatient discharge records and hospital financial disclosure records for the years 1995, 2000, and 2005. The data were obtained from the Florida Agency for Health Care Administration (AHCA). I focus on patients from Miami-Dade and Broward counties in south Florida. These are two of the larger counties in Florida by population, but the area is also small enough to be tractable analytically. The market area is bounded on three sides by either water or by a low-density population area, but is not bounded to the north. However, recall that the IIA assumption

means that offering fewer choices than are present is not a problem.

Hospital discharge records are a brief description of a hospital inpatient stay, including patient age, race, gender, payment sources, and lists of diagnoses made and procedures performed during the stay. No information about patient income is included; this paper will take a fairly standard approach of including the median income of the patient's zip code, measured by the American Community Survey. The patient's zip code can be used together with the hospital's address (obtained from the financial disclosures) to determine travel times to each hospital in the county. Travel times come from Google Maps.

The analysis will focus on three groups of privately insured patients: those with commercial indemnity insurance, those with preferred provider organization (PPO) managed care plans, and those with health maintenance organization (HMO) managed care plans. Indemnity plans are rare: they are an option for less than 2 percent of privately insured individuals in the US (Claxton et al.). However, they account for a larger share of the discharge data (potentially because they chose these plans with the expectation of needing hospital services). These plans do not restrict the choice of hospital. However, they may not offer full coverage, and there may also be other factors, such as emergency room diversion, that keep a commercial indemnity patient from having access to all possible choices.

The other two types of insurance, PPO and HMO, involve limited networks, i.e. insurers may not contract with all providers in a market. Additionally, providers and insurers may have periodic disputes in which the provider is temporarily out of the network. For example, an ongoing dispute in Pittsburgh, PA between UPMC and Highmark has resulted in a number of changes to patients' available hospital networks (Delano). Patients in both types of plan face higher prices if they go to providers not in the network, with HMOs in particular paying out very little for out-of-network providers.

The Miami-Dade county hospital market for these two years contained a mix of for-profit, nonprofit, and government hospitals. There were three major for-profit chains in 1995: Tenet, HCA, and OrNda. Tenet acquired OrNda in 1997, leaving two major for-profit chains by 2000. There was also one government system, the Broward Hospital District.

I examine whether the for-profit systems were absent for some patients. For 1995, there will be three probabilities to estimate; for the other two years there will be two, but I also estimate models in which the former OrNda and Tenet hospitals are treated as though they remain separate. I then test whether the probabilities are equal. I do not examine the possibility that other hospitals are missing, including the Broward public system. I estimate separate models for commercial indemnity, PPO, and HMO patients.

The empirical model follows the random utility framework described above. The specific estimation equation is:

$$U_{ij} = X_i H_j' \beta + T_{ij} X_i \tau + \varepsilon_{ij} \tag{17}$$

where X_i is a vector of patient characteristics that includes a constant term, H_j is a vector of hospital characteristics, and T_{ij} is the travel time between patient i's residence and hospital j. Finally, ε_{ij} is an idiosyncratic component of utility that is distributed as Type I extreme value.

The variables in X are age, white race, female, weighted Charlson index, and emergency status. These are interacted with the hospital characteristics for-profit, not-for-profit, and number of licensed beds. Finally, I include travel time (in minutes) and its interactions with patient characteristics and with a set of indicators for the patient's type of primary diagnosis.

Table **??** contains summary statistics for the estimation samples.

5.2 Estimation results

The results of the estimation demonstrate several patterns. First, HMO patients do not face the most restrictive choice sets; instead, the commercial indemnity patients appear to have the most limitations. This pattern was similar for the models with and without hospital fixed effects. However, the fixed effect results yielded much more realistic probability estimates. The model without hospital fixed effects clearly did not include a sufficient number of hospital characteristics, and the result was that some of the unexplained variation in choices became part of the inclusion probabilities.

Table 4 contains the estimated probabilities and selected coefficients for the models without fixed effects. Estimated qs for 1995 show that commercial indemnity patients may have frequently had access to only a subset of the potentially available choices. This is also true for the PPO patients in the case of the OrNda hospitals. By contrast, HMO patients appear to have generally had access to their preferred hospitals. The results are similar for the 2005 sample. In that sample, PPO patients and HMO patients generally appear to have had access to their preferred choices, but commercial indemnity patients again may not have. Because the model assumed that the only potentially absent hospitals were those in the Tenet, HCA, or (for 1995) OrNda systems–all for-profit systems–the most affected coefficient was the coefficient on for-profit status. The remaining coefficients, including travel time, were similar between the conditional logit and misclassification conditional logit models.

Table 5 replaces the hospital characteristics with hospital fixed effects. Here the patterns are somewhat similar, but the inclusion probabilities change quite a bit. It is still the case that commercial indemnity patients faced the most restrictive choice set in both 1995 and 2005. However, in contrast with results from table 4, the estimated probabilities indicate that all hospitals were present for most patients. The principal effect on the coef-

ficient estimates is for the hospital fixed effects in the three systems. The PPO results did not converge and are not included in the table; they will be added to a future version of the paper, as will the results from 2005.

6 Conclusion

Choices that appear to be present but are actually missing can bias the estimation of choice models. In the hospital choice literature, the general approach has been to either estimate choice models on subsets of patients who may have the entire set, or to accept the bias that comes with the incomplete set. This study describes a method that estimates the inclusion probabilities of sets of choices along with the model coefficients. Applying the model to hospital choices showed that PPO and indemnity patients may not have access to the entire choice set, while HMO patients appeared to have the choice of most hospitals. One possible explanation is that HMO patients were more careful in selecting networks with their preferred hospitals, while indemnity patients may have faced prices greater than zero.

References

Bronnenberg, BJ, and WR Vanhonecker. 1996. "Limited choice sets, local price response, and implied measures of price competition." *Journal of Marketing Research* 33 (2): 163–173.

Capps, C, D Dranove, and M Satterthwaite. 2003. "Competition and market power in option demand markets." *RAND Journal of Economics* 34 (4): 737–763.

Chamberlain, G. 1982. "Multivariate regression models for panel data." *Journal of Econometrics* 18 (1): 5–46.

Claxton, G, M Rae, M Long, N Panchal, A Damico, K Kenward, and H Whitmore. "Employer Health Benefits: 2015 Annual Survey." Kaiser Family Foundation/Health Research & Educational Trust Report.

Delano, J. UPMC, Highmark Reach Deal On Highmark Access To UPMC Facilities In 2015.

Dillon, WR, and S Gupta. 1996. "A segment-level model of category volume and brand choice." *Marketing Science* 15 (1): 38–59.

Gentzkow2007.

Ho, K. 2006. "The welfare effects of restricted hospital choice in the US medical care market." *Journal of Applied Econometrics* 21 (7): 1039–1079.

Hoerl, AE, and RW Kennard. 1970. "Ridge regression: Biased estimation for nonorthogonal problems." *Technometrics* 12 (1): 55–67.

McFadden, D. 1974. "Conditional logit analysis of qualitative choice behavior." Chapter 4 of *Frontiers in Econometrics*, edited by P Zarembka. Academic Press.

Nedungadi, P. 1990. "Recall and consumer consideration sets: Influencing choice without altering brand evaluations." *Journal of Consumer Research* 17:263–276.

Poterba, JM, and LH Summers. 1995. "Unemployment benefits and labor market transitions: A multinomial logit model with errors in classification." *The Review of Economics and Statistics* 77 (2): 207–216.

Shocker, AD, M Ben-Akiva, B Boccara, and P Nedungadi. 1991. "Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions." *Marketing Letters* 2 (3): 181–197.

Swait, J, and M Ben-Akiva. 1987. "Empirical test of a constrained choice discrete model: Mode choice in Sao Paulo, Brazil." *Transportation Research B* 21 (2): 91–103.

Table 1: Simulated results, one missing group, no fixed effects

		Single run, std errors using inv hessian		1000 runs, std errors from bootstrap	
		Cond logit	Msc cond logit	Cond logit	Msc cond logit
Probability	0.8		0.7628		0.807
			(0.0584)		(0.141)
NFP	0.5	0.0446	0.4852	0.066	0.448
		(0.0361)	(0.1369)	(0.016)	(0.265)
Size	0.1	0.0966	0.0949	0.087	0.087
		(0.0176)	(0.0119)	(0.014)	(0.014)
Male · NFP	-0.1	-0.044	-0.0523	-0.074	-0.086
		(0.0176)	(0.0209)	(0.02)	(0.026)
Var1 · NFP	0.2	0.1488	0.1608	0.159	0.171
		(0.042)	(0.013)	(0.012)	(0.017)
Var2 · NFP	0.1	0.0832	0.0791	0.089	0.085
		(0.04)	(0.0122)	(0.012)	(0.014)
Var1 · size	0.2	0.1552	0.1551	0.169	0.169
		(0.0225)	(0.0118)	(0.013)	(0.013)
Var2 · size	0.25	0.1915	0.1915	0.212	0.212
		(0.0742)	(0.0118)	(0.013)	(0.013)
Distance	-1	-0.808	-0.8332	-0.813	-0.833
		(0.0234)	(0.018)	(0.027)	(0.026)
N	50,000				

Standard errors in parentheses. Bootstrap contained 1000 runs.

Weber, E. 2014. "Measuring welfare from ambulatory surgery centers: A spatial analysis of demand for healthcare facilities." *Journal of Industrial Economics* 62 (4): 591–631.

Table 2: Simulated results, three missing groups, no fixed effects

		Cond logit	Msc cond logit
P1	0.8		0.78
			(0.046)
P2	0.85		0.841
			(0.01)
P3	0.7		0.665
			(0.063)
NFP	0.5	0.142	0.538
		(0.014)	(0.12)
Size	0.1	0.14	0.08
		(0.01)	(0.01)
Male · NFP	-0.1	-0.085	-0.098
		(0.018)	(0.021)
Var1 · NFP	0.2	0.158	0.197
		(0.01)	(0.014)
Var2 · NFP	0.1	0.069	0.099
		(0.009)	(0.012)
Var1 · size	0.2	0.166	0.183
		(0.008)	(0.009)
Var2 · size	0.25	0.215	0.23
		(0.008)	(0.009)
Distance	-1	-0.918	-0.954
		(0.046)	(0.028)
N	50,000		

Standard errors from bootstrap in parentheses. Bootstrap contained 1000 runs.

Table 3: Simulated results, three missing groups, no fixed effects

				Misclassification cond logit		
	Group	True values	Cond logit	$\lambda = 0$	$\lambda = 0.5$	$\lambda = 1$
P1		0.8		0.747	0.77	0.793
				(0.277)	(0.136)	(0.116)
P2		0.85		0.738	0.803	0.842
				(0.149)	(0.173)	(0.156)
P3		0.7		2.901	0.857	0.801
				(3.977)	(0.182)	(0.107)
Male · NFP		-0.1	-0.087	-0.101	-0.1	-0.096
			(0.019)	(0.024)	(0.031)	(0.03)
Var1 · NFP		0.1	0.097	0.117	0.114	0.109
			(0.01)	(0.015)	(0.018)	(0.017)
Var1 · size		0.2	0.195	0.205	0.203	0.201
			(0.01)	(0.011)	(0.015)	(0.015)
Distance		-1	-0.911	-0.948	-0.943	-0.935
			(0.037)	(0.041)	(0.047)	(0.044)
FE 1	0	0.7	0.661	0.663	0.662	0.66
			(0.022)	(0.022)	(0.027)	(0.028)
FE 2	0	0.9	0.826	0.829	0.828	0.825
			(0.022)	(0.022)	(0.028)	(0.028)
FE 3	1	-0.2	-0.589	-0.077	-0.162	-0.222
			(0.035)	(0.334)	(0.283)	(0.23)
FE 4	1	1.1	0.684	1.202	1.118	1.055
			(0.026)	(0.335)	(0.284)	(0.229)
FE 5	1	-0.1	-0.474	0.719	-0.043	-0.107
			(0.034)	(0.273)	(0.284)	(0.228)
FE 6	2	0.5	0.265	0.719	0.603	0.523
			(0.024)	(0.273)	(0.291)	(0.252)
FE 7	2	0.2	-0.021	0.432	0.314	0.236
			(0.026)	(0.272)	(0.291)	(0.252)
FE 8	3	-0.2	-0.613	-1.088	-0.413	-0.365
			(0.033)	(1.017)	(0.275)	(0.152)
N	25,000					

Standard errors from bootstrap in parentheses. Bootstrap contained 1000 runs. The omitted choice is also in group 0.

Table 4: South Florida results, no fixed effects

	Indemnity		PPO		HMO	
	Condit logit	Misc logit	Condit logit	Misc logit	Condit logit	Misc logit
1995						
HCA		0.753***		0.790***		1.067***
		(0.013)		(0.013)		(0.008)
OrNda		0.646***		0.073***		1.027***
		(0.015)		(0.006)		(0.011)
Tenet		0.849***		0.856***		1.060***
		(0.012)		(0.012)		(0.004)
Not for profit	0.976***	1.077***	1.143***	1.257***	1.217***	1.227***
	(0.060)	(0.062)	(0.048)	(0.049)	(0.040)	(0.040)
For profit	0.736***	1.077***	-0.017	0.445***	0.871***	0.837***
	(0.057)	(0.061)	(0.048)	(0.051)	(0.038)	(0.039)
Licensed beds	0.228***	0.241***	0.093***	0.124***	0.168***	0.169***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)
Travel time	-0.113***	-0.114***	-0.117***	-0.117***	-0.113***	-0.111***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
N	752710	752710	1021825	1021825	2085755	2085755
2005						
HCA		0.456		1.144		0.944
		(0.020)		(0.020)		(0.008)
Tenet		0.504		0.733		0.848
		(0.020)		(0.022)		(0.010)
Not for profit	0.487***	0.462***	-0.209***	-0.227***	-0.194***	-0.189***
	(0.096)	(0.113)	(0.055)	(0.062)	(0.041)	(0.047)
For profit	-0.176	0.477***	-0.870***	-0.821***	-0.205***	-0.105***
	(0.105)	(0.132)	(0.047)	(0.058)	(0.033)	(0.040)
Licensed beds	0.209***	0.216***	0.062***	0.063***	0.106***	0.107***
	(0.011)	(0.012)	(0.006)	(0.007)	(0.005)	(0.005)
Travel time	-0.117***	-0.128***	-0.115***	-0.114***	-0.103***	-0.104***
	(0.007)	(0.008)	(0.004)	(0.004)	(0.003)	(0.003)
N	341280	341280	925260	925260	1763250	1763250

Standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%.

Table 5: South Florida results, fixed effects

	Indemnity		HMO	
	Condit logit	Misc logit	Condit logit	Misc logit
1995				
Travel time	-0.114***	-0.116***	-0.117***	-0.120***
	(0.004)	(0.004)	(0.003)	(0.003)
Factor = 100009	-0.613***	-0.539***	0.337***	0.336***
	(0.067)	(0.088)	(0.034)	(0.042)
Factor = 100029	-0.646***	-0.543***	-0.317***	-0.240***
	(0.058)	(0.069)	(0.030)	(0.032)
Factor = 100187	-0.344***	-0.267***	-0.037	0.095***
	(0.062)	(0.077)	(0.035)	(0.038)
Factor = 100209	-0.602***	-0.478***	-0.480***	-0.303***
	(0.067)	(0.090)	(0.038)	(0.043)
N	752710	752710	2085755	2085755

Standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%.