Why Do Previous Choices Matter for Hospital Demand? Decomposing Switching Costs from Unobserved Preferences*

Devesh Raval

Federal Trade Commission

draval@ftc.gov

Ted Rosenbaum

Federal Trade Commission

trosenbaum@ftc.gov

December 19, 2017

Abstract

Since patients frequently revisit the same health care provider, their previous choice of provider is highly predictive of their future choices. Using data on women's choice of hospital for childbirth in Florida, we find that women return to the same hospital approximately 70% of the time. However, it is difficult to disentangle how much of this behavior reflects structural state dependence ("switching costs") as opposed to unobserved preference heterogeneity. We separate these two explanations using a panel data fixed effects estimator, and find that switching costs account for approximately 40% of the difference between predictions from models that do and do not include a lagged dependent variable. We then show how correctly identifying the relative magnitude of switching costs affects policy conclusions for product exit and for the dynamic effects of entry on competition. The welfare effects of excluding a hospital from a payer's network are smaller in the short run, but higher in the long run, given our estimates of switching costs compared to estimates that do not control for unobserved heterogeneity. The dynamic effects of entry on competition are also significantly smaller with lower switching costs.

*The views expressed in this article are those of the authors. They do not necessarily represent those of the Federal Trade Commission or any of its Commissioners. We thank editor Amitabh Chandra, Dan Hosken, Nathan Petek, Dave Schmidt, Mark Shepard, Nathan Wilson, and an anonymous referee for their comments on this paper, and Massimiliano Piacenza for his discussion of an earlier version of this paper.

1 Introduction

The provision of health care often involves multiple interactions between patients and health care providers. Since patients frequently revisit the same provider, a patient's previous choice is likely to be highly predictive of future choices.

However, empirical demand models of provider choice typically do not use patients' previous choices in order to predict current ones (Gowrisankaran et al., 2015; Ho and Lee, 2017), in part because it is difficult to disentangle the reasons why patients revisit the same provider. As Heckman (1981) demonstrates more generally, there are two major reasons for such behavior. First, visiting a provider could increase a patient's utility from revisiting that provider relative to others. We call this effect, which Heckman (1981) refers to as structural state dependence, *switching costs*. Second, patients may revisit the same provider because of persistent unobserved tastes for that provider.

Distinguishing between these two mechanisms is important for understanding the incentives of providers and insurance companies along several dimensions. First, the bargaining position of an insurer relative to a health care provider depends upon switching costs through how excluding that provider from the insurer's network affects patients. Thus, to give one example, one must know the extent of switching costs to understand the welfare effects of a transition to a narrow network insurance plan. In addition, when switching costs are large, provider entry will be more difficult because an entering firm will find it hard to win over captive consumers. Providers may also find it difficult to attract new patients by improving quality, repositioning in product space, or advertising if consumers fail to respond to such costly efforts.

Using panel data from Florida from 2006 to 2014, we study why patients' previous hospital choice is predictive of women's choice of hospital for childbirth. We examine childbirth because the services that the patient needs are similar across multiple childbirths, but patients are not required to return to the same hospital where they previously received care. Also, childbirths are relatively similar in clinical treatment, which facilitates comparisons across multiple admissions.

Consistent with Shepard (2016), patients' previous choices are an important predictor of their current choices. A model that includes women's previous choice as a covariate predicts that women will return to the same hospital 72% of the time. In contrast, a model that does not include previous choices predicts a return to the same hospital 40% of the time.

We then separate switching costs from persistent unobserved tastes for hospitals using a panel data fixed effects approach from Honoré and Kyriazidou (2000). Our identification is based upon a woman's choice of hospital for her third birth, using variation from women who switch providers between their first and second births. Using the panel structure of the data, we can hold persistent unobserved tastes constant across births. This allows us to infer switching costs from the proportion of women who choose the same hospital for their third birth and their second birth, as opposed to returning to the same hospital that they used for their first birth or going to a hospital she has not previously visited.

The fixed effects estimator implies that having previously visited a hospital increases its probability of visitation by 89% for a hospital with a low probability of being selected, which is 45% to 55% lower than under the estimates from the lagged dependent variable model. In addition, we find that about 40% of the changed predictions of models after including a lagged dependent variable is due to switching costs.

We then demonstrate the importance of distinguishing between the switching cost and preference heterogeneity explanations through two applications. For each application, we compare policy predictions given three models: the standard logit model, with no switching costs, the lagged dependent variable model, with the full effect of the lagged dependent variable as switching costs, and the fixed effects model.

First, we examine the welfare effects of an insurer excluding a hospital from its network. While the welfare loss due to switching hospitals is a one time cost, the welfare loss from women going to a less preferred hospital will be persistent over the long run. Thus, assuming that the full magnitude of the lagged dependent variable is switching costs will tend to overstate welfare effects in the short run, and understate welfare effects in the long run. We examine these effects by simulating the exclusion of a hospital from patient choice sets. Indeed, we find that, in the short run, the change in welfare from the network exclusion is smaller under the fixed effects estimates, at 10.5%, than 9.2% under the lagged dependent variable estimates. However, in the long run, the decline in welfare is larger under the fixed effects estimates, at 7.1%, relative to 5.6% under the lagged dependent variable estimates.

Second, we assess whether entry of a hospital can alleviate the effects of a merger on competition. We first examine the initial impact of entry for a merger of two hospital systems across all three models, using the percent change in WTP for the merged system – an indicator of post-merger competitive harm based upon Capps et al. (2003) – as a proxy for the effects of entry on competition. The different models imply substantially different entry effects on competition. An entrant hospital with unobserved quality a half standard deviation below the mean would alleviate the competitive harm for the lagged dependent variable and standard logit models, but not for the fixed effects model. In addition, greater switching costs imply greater dynamic effects of entry on competition, which magnify the initial differences between models. When the entering hospital is one standard deviation higher quality than the average quality in each zip code, the percent change in WTP falls by 8 percentage points under the lagged dependent variable model compared to 3 percentage points under the fixed effects model.

While we study health care markets in this paper, economists and marketers have documented the presence of switching costs in a wide array of markets (Shum (2004), Dubé et al. (2009), Dubé et al. (2010), Goettler and Clay (2011)). In addition, some work has focused on separating estimates of switching costs from unobserved heterogeneity in other markets. Chintagunta et al. (2001) apply the Honoré and Kyriazidou (2000) estimator to estimate switching costs in the context of yogurt demand. Handel (2013) examines health insurance choice when consumers are forced to make an active choice in one period because the set of plans available changes, while Sudhir and Yang (2014) use auto rental upgrading due to stockouts to separate switching costs from consumer heterogeneity.

The paper proceeds as follows. In Section 2 we outline a model of provider choice and our econometric approaches to identification. In Section 3 we describe the data that we used. In Section 4, we detail our structural estimates of switching costs, and in Section 5 we apply these estimates to two policy applications. We conclude in Section 6.

2 Identification of Consumer Choice Model

We begin by reviewing a workhorse model of a patient's hospital choice. This model is the cornerstone of a broader empirical model of hospital and insurer bargaining (Capps et al. (2003), Gowrisankaran et al. (2015), Ho and Lee (2017)). We then demonstrate how to identify switching costs and distance in this framework.

2.1 Baseline Model

A patient i becomes pregnant at time t in market m. She has a set of hospitals H ($j = 1, ..., N$) that are available to her, as well as an outside option $j = 0$. Patient i's utility from care at hospital j at time t is given by:

$$u_{ijt} = \underbrace{\beta x_{ijt} - \alpha d_{ijt} + \xi_{ij} + \gamma I[j = H_{it-1}]}_{\delta_{ijt}} + \epsilon_{ijt} \tag{1}$$

where

$$H_{it-1} = \arg\max_{k=0,...,J} u_{ikt-1}$$

In equation (1), δ_{ijt} is the expected utility of hospital j for patient i at time t , and ϵ_{ijt} is a patient-hospital-time Type-I extreme value i.i.d. error term that reflects a patient's idiosyncratic hospital preferences. We normalize $\delta_{i0t} = 0$ for all i, t.

We parametrize δ_{ijt} to include distance, switching costs, and observed and unobserved components of an individual's tastes for hospitals. In equation (1), x_{ijt} are a set of observable characteristics of the patient, the hospital, and interactions between them.[1] The distance from patient i's residence to hospital j at time t is d_{ijt}.

Since $I[j = H_{it-1}]$ is a dummy variable indicating whether the patient attended hospital j on her previous visit, γ represents the switching costs of attending a different hospital on different visits. Persistent patient preferences for a given facility are represented by ξ_{ij} . These unobservables could be the result of persistent doctor or friend referral patterns, patient specific preferences for hospital amenities, or proximity to another location to which

[1] See Raval et al. (2017a) for a description of the different types of models that have been used in this context.

the patient frequently travels.

2.2 Identification

Because a patient's previous choice $I[j = H_{it-1}]$ is a function of the patient's distance to the facility in time $t-1$ (d_{ijt-1}) and her persistent preference for that facility (ξ_{ij}), one cannot identify switching costs by including a lagged dependent variable in the demand specification. This problem is well known in the literature; for example, Shepard (2016) includes a lagged dependent variable but acknowledges that he is unable to separately identify preference heterogeneity from switching costs.

To separately identify switching costs γ from unobserved patient heterogeneity, we use a panel data approach from Honoré and Kyriazidou (2000) which allows us to identify both α, the effect of distance, and γ, the effect of switching costs, when both effects may be conflated with unobserved heterogeneity. Both γ and α are identified given data from at least three births of patients that switch at least once between providers, and have one birth after a switch without moving residence.[2]

Using the utility function from equation (1) and the results from Honoré and Kyriazidou (2000), the expression for the conditional probability is

$$\begin{aligned}&\Pr((A,B,j)|(A,B,j) \text{ or } (B,A,j)) \\ &\qquad = (\Pr(A,B,j)/\Pr(B,A,j))/((\Pr(A,B,j)/\Pr(B,A,j))+1) \quad (2)\end{aligned}$$

where j is any facility and the equality follows from the definition of conditional probability. Using the parameterization in equation (1),

$$\Pr(A,B,j)/\Pr(B,A,j) = \exp(\alpha((d_{iA1}-d_{iB1})-(d_{iA2}-d_{iB2}))+\gamma(I[j=B]-I[j=A])). \quad (3)$$

We provide a derivation of equation (3) in Appendix A.1.

For intuition on our identification strategy, consider Figure 1, which shows a woman who

[2]See Raval and Rosenbaum (2016) for an example of identifying another variable, the steering effect of Medicaid Managed Care, using the same framework.

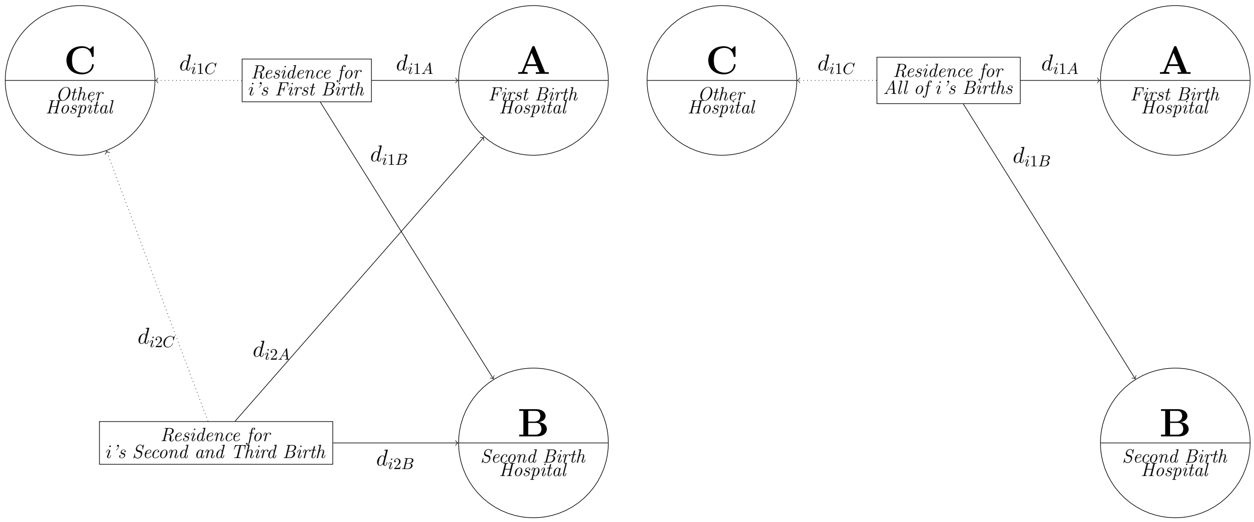
has hospitals A, B, and C in her choice set for all three births. The left figure in Figure 1 examines a woman who moves between her first and second birth. For her first birth, she lived closer to hospital A than hospital B. Between the first two births, she moves residences, such that for her second birth she is closer to hospital B than hospital A. She remains at the same residence for her third birth. Since hospital A is located closer to her for the first birth, and hospital B is located closer to her for the second birth, her likelihood of going to A for the first birth and B for the second birth increases as the distance elasticity rises. As switching costs rise, her likelihood of going to hospital B (compared to hospitals A, C, or any other hospital) for the third birth rises if she went to hospital B for the second birth as opposed to the first birth.

The right figure in Figure 1 examines a woman who remains at the same residence for all three births, but switches between hospital A and hospital B between her first and second birth. Her likelihood of switching hospitals for her third birth decreases as switching costs rise, thus providing identifying information for the switching cost parameter γ. However, because she does not move, we cannot learn anything about the distance coefficient from her behavior using the conditional logit estimator.

The conditional logit estimator is based upon the probability that she went to hospital A for the first birth, and then hospital B for the second birth, and then any hospital j for the third birth, compared to the opposite order for the first and second birth and hospital j for the third birth. Since hospitals A, B, and j were each chosen once, the fixed effects approach differences out any time-invariant hospital-patient interactions. Switching costs for the second birth are also differenced out, because in either case the patient incurs a switching cost γ in the second period, but switching costs for the third birth are included. We are then able to identify the marginal impact of changing distance using the likelihood terms from the first two births based upon the patient's move, and the marginal impact of switching costs from the patient's choice of hospital for the third birth.

We use equation (3) to estimate α and γ using a conditional likelihood function using women who have three children, switch hospitals between their first and second birth, and do not move houses between their second and third child. The third birth yields variation in switching behavior over time for women visiting one of the same hospitals as for the first

Figure 1 Identification Intuition



(a) Movers (b) Non Movers

two births, as some women will go to the same hospital for their second and third birth while others will go to a different one. Women who, in addition, move between their first and second birth provide identifying variation for the distance coefficient α.

Our estimation approach assumes that only a woman's immediately previous choice affects her decision making. If, however, whether that hospital was *ever* chosen is the relevant determinant of switching costs, then we are likely to underestimate the extent that switching costs determine choices relative to preference heterogeneity. Our estimate of switching costs detects the switching cost of going to a hospital for a woman's immediately prior birth versus her first birth. An additional switching cost associated with going to a hospital that one has never gone to would add to the switching cost we identify.

3 Data

We use hospital discharge data obtained from the Florida Agency for Health Care Administration (AHCA) from 2006 to 2014. The data includes the zip code of residence for each patient, which allows us to compute the travel time from the patients' residence to each hospital. Patient identifiers allow us to match births by the same woman over time.[3]

[3]The limited data set was obtained from the AHCA, but that agency bears no responsibility for any analysis, interpretations, or conclusions based on this data. We use information on the age in days and

For our analysis of switching costs, it is important to have the full medical history of the women. Since our data starts in 2006, we only include women who were no more than 21 in 2006, which eliminates the initial conditions problem for most of the women.[4] The under 21 age restriction leaves about 35 percent of the women from the full sample of all births in Florida. We define the choice set for our standard estimator as all hospitals within 45 minutes driving time.[5] The hospital chosen for each birth has to be in the choice set for all births.

Of women in this group where the previous hospital is within a 45 minute drive time, 74% of the deliveries for second births and later were in the same hospital as the previous birth. This is relatively constant across births and is similar for normal deliveries and deliveries with more complications.

The fixed effects estimator requires women to have at least three births, switch hospitals between their first and second birth, have both of those hospitals within a 45 minute drive for both births, and not move residences between their second and third births. From the set of women who are under 21 in 2006, 3,883 fit these criteria and form the Honore dataset that we use for our fixed effects estimates. Of these women, 59% of women went to the same hospital for their second and third birth (the "ABB" group) and 23% went to the same hospital for their first and third birth (the "ABA" group).

While we use the full set of 3,883 women in the estimation, we can gain further intuition by focusing on the 2,302 women in that group who never moved residences. For those women, their distance to each hospital does not change over time, so the difference in the order in which they went to hospitals is solely informative about the extent to which structural state dependence drives decisions. Without state dependence, an approximately equal number of women should be in the ABA and ABB groups. For this group of women, 26% of women

admission date of each patient in order to check that the patient identifiers correctly identify unique patients, and remove all patients that do not have unique identifiers. We use the Diagnosis Related Group (DRG) codes provided in the data to indicate pregnancy. For DRG version 24 and earlier, these are DRGs from 370 to 375. For DRG version 25 and after, these are DRGs 765 to 768 and 774 to 775.

[4] Nationwide, about 87 percent of first births are of mothers age 20 and above and 96 percent of first births are of mothers age 18 or above (Hamilton et al. (2015))

[5] See Ho and Pakes (2014) for a similar choice set restriction. This is a different assumption from the fixed effects estimator, where we assume that the choice set (which does not need to be fully specified) is constant across all births.

follow an ABA pattern and 56% of women follow an ABB pattern. Since many more women are in the ABB group than the ABA group, switching costs must be substantial.

We report summary statistics in Table I for the three datasets described above: the full sample of all births in Florida, all births for women that meet our age restriction, and births for women who meet our age restriction and contribute to the likelihood in the fixed effects estimator (the "Honore" sample). The average age is much higher for the full sample, at 27.6 years, than for the datasets that impose the age restriction, both of which have an average age of about 21.7. As we impose more restrictions, the fraction of admissions that are black rises from 23 percent for the full sample to 48 percent for the Honore dataset. The fraction of admissions from patients on Medicaid rises from 51 percent for all births to 83 percent for the Honore sample.

Table I Summary Statistics

	All	Age Restriction	Honore
Age	27.58	21.73	21.73
White	0.66	0.61	0.45
Black	0.23	0.30	0.48
Hispanic	0.20	0.20	0.17
Medicaid	0.51	0.72	0.83
Commercial	0.43	0.22	0.13
Metropolitan	0.94	0.92	0.96
Normal Birth	0.78	0.80	0.79
Number of Births	1,794,598	667,310	13,099
Number of Women	1,247,610	439,636	3,883

Note: The full sample includes childbirths of all women in Florida between 2006 and 2014. The age restriction sample includes all women in Florida that were of age 21 or less in 2006. The Honore sample only includes women in the age restriction sample who switch hospitals between their first and second birth and have a third birth, and do not move residences between second and third birth.

4 Results

In this section, we first show that a lagged dependent variable has significant explanatory power in predicting patient choices. Then, using a fixed effects estimator, we show that both preference heterogeneity and structural state dependence play an important role in the explanatory power of the lagged dependent variable.

4.1 Reduced Form Estimates

A reduced form way to account for any kind of dependence in choices over time, either due to state dependence or unobserved heterogeneity, is to include a lagged dependent variable in the demand model. In Table II, we display the estimates of a multinomial logit model of women's choice of hospital for obstetrics that includes all women that meet our age restriction. We both include and exclude a lagged dependent variable from the specification, and include linear travel time and hospital dummy variables as additional covariates. While much of the literature includes more detailed controls for the type of patient, our sample is extremely homogeneous, consisting of young women with the identical procedure.

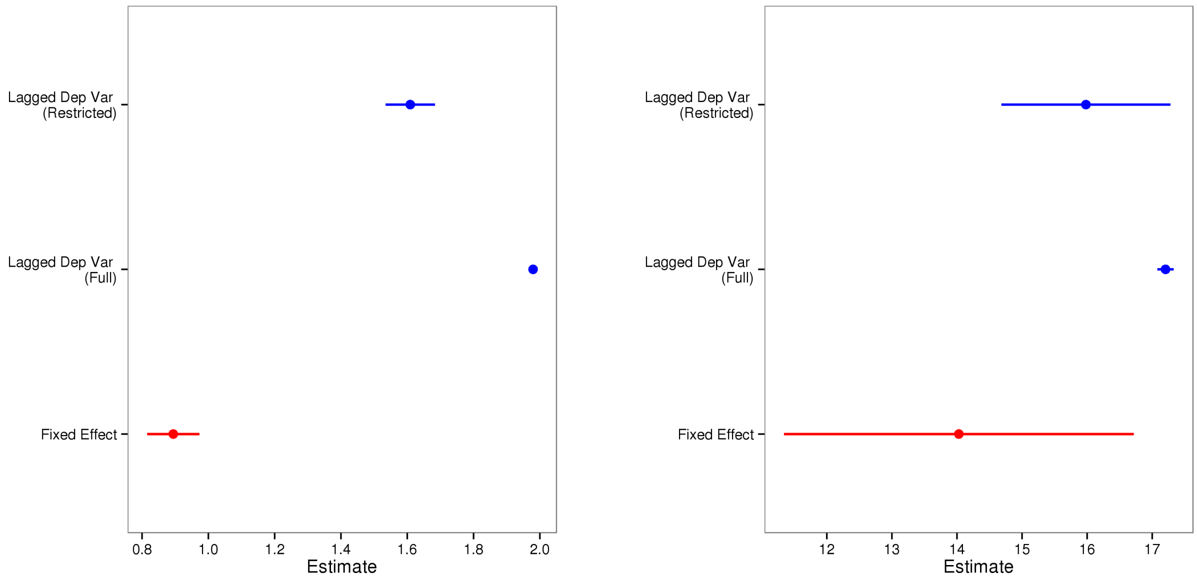
Including the lagged dependent variable has explanatory power and leads to different predictions of hospital choices. When hospital choices depend solely on travel time and a hospital fixed effect, the "standard logit" specification, approximately 40% of women in our sample would be expected to return to the same hospital as their previous birth. When a lagged dependent variable is included, this share increases to 72%.[6]

Table II Baseline Estimation

	Standard Logit	Standard with Lagged Dep Var
Travel Time	-0.12	-0.12
	(0.00)	(0.00)
Previous Choice		1.98
		(0.01)
Predicted Share to Previous	39.75%	72.14%
N	626,738	626,738

Note: The parameter estimates in the first two rows are derived from a multinomial logit model that includes linear travel time and hospital dummy variables, and either includes or excludes a lagged dependent variable. The "Predicted Share to Previous" gives the expected number of people who would use the same hospital as their previous birth, conditioning on the previous sequence of hospital choices.

[6] By comparison, Shepard (2016) estimates that people are likely to go to a hospital where they have a previous relationship about 40% of the time, which is five times larger than what would be expected based upon other covariates. However, that analysis includes all hospital admissions and we focus only on childbirth.



(a) Switching Costs

(b) Switching Costs (min)

Figure 2 Estimates of Switching Costs

Note: The lagged dependent variable full sample specification is based on all women meeting our age restriction, while the lagged dependent variable restricted sample is only on third births in the "Honore" sample. For each specification, the dot is the point estimate and the lines are the 95% confidence interval. Switching costs in minutes is the switching cost estimate divided by the distance estimate. See Table III for a table of the estimates and standard errors used to generate this figure.

4.2 Structural Estimates

While patients' previous choice has significant explanatory power in predicting future choices, it is not clear how much of the importance of patients' previous choice is due to preference heterogeneity or state dependence. In order to disentangle the two, we estimate the fixed effects model described in Section 2. To ensure that our results are not driven by sample selection, we compare our results from the full set of women who meet the age restriction – the women used in the model estimates in Table II – to the restricted set of women that contribute to the likelihood for our fixed effects estimation.

Figure 2a depicts the parameter estimates and 95% confidence intervals for switching costs. While both lagged dependent variable estimates are of similar magnitude, the switching cost parameter from the fixed effects specification is 45% to 55% smaller than that of the lagged dependent variable approach. For a hospital with a minimal probability of being chosen, a patient having visited that hospital before increases demand by 89%, compared to

198% for the lagged dependent variable model for the full sample and 161% for the lagged dependent variable model for the restricted sample.[7] The statistical difference between these coefficients is well beyond that which can be explained by sampling variation. Thus, people are substantially more willing to switch hospitals than would be implied from treating the lagged dependent variable as a measure of structural state dependence.

Figure 2b examines the tradeoff between switching costs and distance by dividing the switching cost estimate by the distance estimate. For switching costs in terms of travel time, the difference between the lagged dependent variable and fixed effects specifications is much smaller. Switching costs in minutes fall by 12% to 18% under the fixed effects specification compared to the lagged dependent variable specifications, to about 14 minutes of travel time. The reason for this is that, as we demonstrate in Raval and Rosenbaum (2017), distance estimates also decline after controlling for unobserved heterogeneity. The difference between the fixed effects specification and lagged dependent variable specification for the full sample is barely statistically significant at the 5% level, while the difference with the lagged dependent variable specification for the restricted sample is statistically insignificant.

4.3 Decomposing a Lagged Dependent Variable Effect

Using these estimates, we measure the extent to which the inclusion of a lagged dependent variable in a standard logit framework reflects switching costs as opposed to persistent preference heterogeneity. Since the fixed effects estimator we use relies on differencing out the fixed effects, we cannot recover estimates of ξ_{ij} for each individual/hospital pair in the estimation. Therefore, we cannot compute the individual specific choice probabilities under the fixed effects model in a way that allows for preference heterogeneity and switching costs. Nevertheless, we can still gain insight into the relative importance of preference heterogeneity and switching costs by comparing the predicted share of women who return to the same hospital as their previous birth across three models: the standard logit without a lagged dependent variable, the standard logit with a lagged dependent variable, and the fixed effects logit estimator where we compute choice probabilities under the assumption of no preference heterogeneity.

[7] Given the logit structure, this is given by $\gamma(1 - Pr(h_{it} = j))$.

We recover the choice probabilities for the fixed effects estimator through a two step process. First, we recover the distance and switching cost parameters using the fixed effects estimator outlined above. Then, we rerun a full logit model plugging in the parameter estimates for distance and switching costs to recover the mean hospital quality. In the first step, we allow for unobserved preference heterogeneity in order to consistently recover an estimate for switching costs. In the second step, we use those parameters to compute mean hospital quality under the assumption of no preference heterogeneity. For this calculation, we focus our attention on the restricted Honore sample.

Under the standard logit approach without a lagged dependent variable, 30% of women are predicted to return to their previous hospital, while under the fixed effects approach without preference heterogeneity 42% of women are predicted to return. The difference in predictions between these two approaches reflects the impact of switching costs on predicted choices, because the standard logit model does not include switching costs. Under the lagged dependent variable logit, 59% of women are predicted to return to the same hospital. The predictions from the lagged dependent variable specification include both preference heterogeneity and switching costs. Therefore, the difference between the fixed effects predictions and the lagged dependent variable predictions can be thought of as reflecting preference heterogeneity. Using this approach, approximately forty percent of the difference in predicted choices is due to structural state dependence.[8]

4.4 Robustness

Our results remain robust to alternative specifications that relax some of the model's assumptions. Our baseline model assumes that all patients have the same ex-ante treatment complexity, that we observe all births of the women, and that there is no time-varying change in patient tastes for providers that can account for their switching behavior.

To address the concern that patients may vary in treatment complexity, we restrict our attention to women that had a normal delivery for all three deliveries. To address the concern that some women may have had births prior to 2006, the start of our dataset, we further

[8]The calculation is $(42-30)/(59-30) \approx 40\%$. In Appendix B, we show that we cannot account for this preference heterogeneity by including additional observables into our demand estimation.

restrict our sample to women who were 18 and under in 2006.

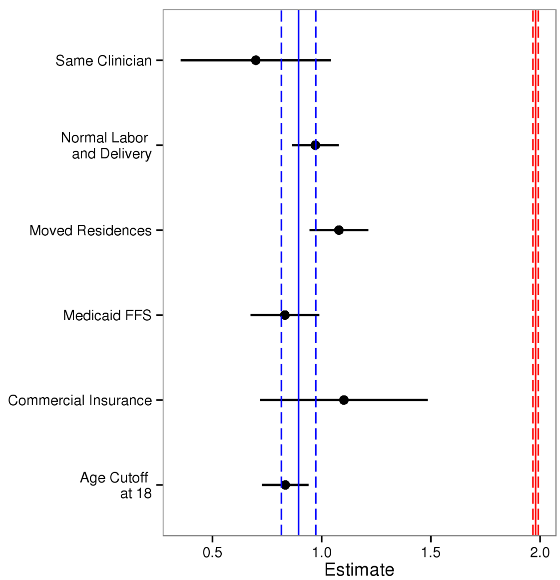
We also examine three reasons why time-varying patient/hospital unobservables could arise. First, women may shift delivery physician over time to physicians with different preferences over hospitals for delivery, yielding a time-varying patient/hospital unobservable. We thus estimate a specification where we only include women who have the same delivery physician for each of the three births. Second, women may have shifted insurance between births. Therefore, we examine two specifications that restrict the sample to women who had the same insurance type for all deliveries. In one specification we look at women who had commercial insurance for all births and in the other we look at women who had Medicaid Fee-for-Service insurance for all births. Finally, we consider a specification to account for the fact that women may have had a negative experience at their first birth hospital that drove them to switch to a different hospital for their second birth. Therefore, we estimate a specification where we only include women who move residences between her first two births. For these women, it is more likely that the switch took place because of their move rather than because of a negative experience.[9]

Figure 3 depicts the results from these alternate specifications.[10] In general, the estimates of switching costs are close to the estimates from the baseline fixed effects approach, and much lower than the estimates from the lagged dependent variable estimator.

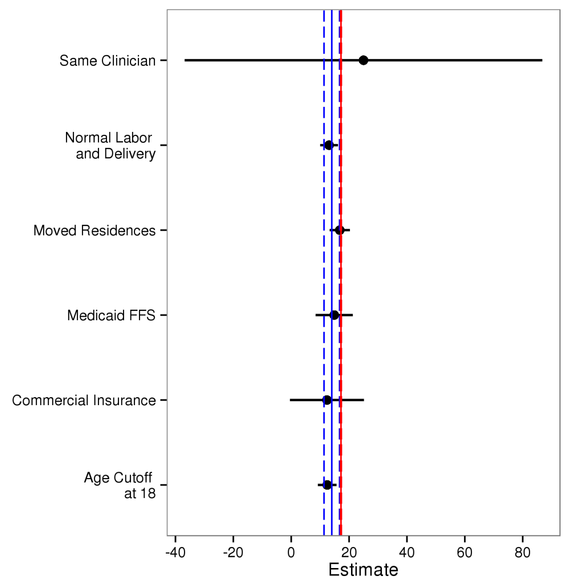
In addition, we conducted a set of Monte Carlo experiments to examine the bias in our estimates when the underlying model is misspecified due to time varying shocks. Appendix C contains the results of these experiments. We find that the model is approximately unbiased when it is correctly specified, or when the variance of the time varying shock is small. When the variance of the time varying shock is relatively large, estimates of switching costs decline when the time varying shock is slightly autocorrelated over time and rise when the time varying shock is substantially autocorrelated over time. However, the magnitude of the bias is less than 0.15 in all cases, compared to a true value of 0.89, and much smaller than

[9]One type of measurable negative experience is having a delivery with complications. We find that among women meeting our age restriction those who switch hospitals between their first and second birth had an uncomplicated first birth 66% of the time as opposed to 69% of women who went back to the same hospital. However, since our estimates that restrict to only women with all normal deliveries look very similar to our baseline estimates, we do not think that this type of time-varying shock would significantly bias our results.

[10]Figure 10 in Appendix E examines several of these same specifications using the lagged dependent variable estimator, and finds similar point estimates to the baseline estimates using this estimator.



(a) Switching Costs



(b) Switching Costs (min)

Figure 3 Robustness Checks for Estimates of Switching Costs

Note: The red vertical lines are the parameter estimates from the lagged dependent variable model while the blue vertical lines are the parameter estimate from the fixed effects model. The solid lines are the point estimates while the dashed lines are the 95% confidence interval. The black horizontal lines are the parameter estimates from our robustness checks. The dot is the point estimate and the lines are the 95% confidence interval. See Table IV for a table of the estimates and standard errors used to generate this figure.

the difference between the lagged dependent variable and fixed effect estimates of switching costs.

5 Applications

In this section, we examine two ramifications of these differing estimates of demand: the welfare impact of a network exclusion and the effect of entry on competition after a merger. For each application, we examine the standard logit model, with no switching costs, the lagged dependent variable model, with the full magnitude of the lagged dependent variable as switching costs, and the fixed effects model. In addition to the differences in switching costs, the fixed effect model separates the structural impact of distance from preference heterogeneity correlated with distance, while in the standard logit and lagged dependent variable models the distance coefficient conflates the two.

Since the fixed effects estimator we use relies on differencing out the fixed effects, we impose an additional assumption restricting patient heterogeneity ξ_{ij} to vary at the zip code level for patients' first births. Under this assumption, we recover ξ_{ij}. We apply this assumption to all three models using the estimated demand parameters from the previous section.[11] For both applications, we use the same service area in Florida, based upon the 17 zip code service area stipulated by HCA for its planned move of Plantation General Hospital (PGH) to the campus of Nova Southeastern University (NSU) in Broward County, Florida. In Appendix D.1, we provide further details of this service area.

5.1 Network Exclusion

Insurance companies have begun to offer narrow network plans that exclude several hospital providers in the market as one way to reduce costs by excluding more costly providers (Robinson (2003)). The use of narrow network plans has accelerated with the health care reforms of the Affordable Care Act (ACA); ACA marketplaces have a higher prevalence of narrow network plans (Haeder et al. (2015)). For example, in New Hampshire, the only

[11] Details are in Appendix A.2.

insurer in the ACA marketplace excluded 10 of the state's 26 hospitals from its network.[12]

Holding all prices fixed, consumer welfare is clearly lower following a network exclusion. However, estimates of the long and short run effects of that exclusion depend critically on distinguishing between switching costs and consumer preference heterogeneity. In particular, using a lagged dependent variable model that overstates the importance of switching costs is likely to overstate the short run welfare loss and understate the long run welfare loss. The component of the welfare loss that is due to switching costs is a one time cost. However, the component of the welfare loss that is due to women going to a less preferred hospital is going to be persistent over the long run (Shepard (2016)).

With that in mind, we examine the effect of a network exclusion on consumers by examining the welfare implications of removing PGH in its current location from an insurer's network. In particular, we conduct simulations that exclude PGH from women's choice sets and examine how the ex-ante utility for hospitals changes over time under each estimate of switching costs. We compare a woman's ex-ante utility when she has PGH in her network to her ex-ante utility when she does not.[13] For this analysis, we study women who have had previous births for whom switching costs are relevant.

Figure 4 compares women's ex-ante utility if the hospital is excluded from her choice set to her utility if it is included. When switching costs are equal to the full magnitude of the lagged dependent variable, there is a dramatic change in the fall in utility over time, from -10.5 percent in Year 0 to -5.6 percent in Year 10. Under the fixed effects switching cost estimates, the differences in the fall in utility over time are much less sharp, falling from -9.2 percent in Year 0 to -7.1 percent in Year 10.[14]

[12] See https://www.bostonglobe.com/news/nation/2014/01/20/narrow-hospital-networks-new-hampshire-spark-j2ufuNSf9J2sdEQBpgIVqL/story.html.

[13] We exclude the possibility that the women could switch insurers and maintain access to PGH from these welfare calculations. See Appendix A.3 for the formula for the change in ex-ante utility from a change in the choice set.

[14] Using the standard logit specification, the utility decline is constant over time at 8.7 percent, and so is higher than under both the lagged dependent variable and fixed effects models in the long run. The intuition for this is that, by assuming away switching costs, the utility from the excluded hospital is larger relative to the utility of all other hospitals. Thus, excluding the hospital is a much larger utility loss than under the other two models.

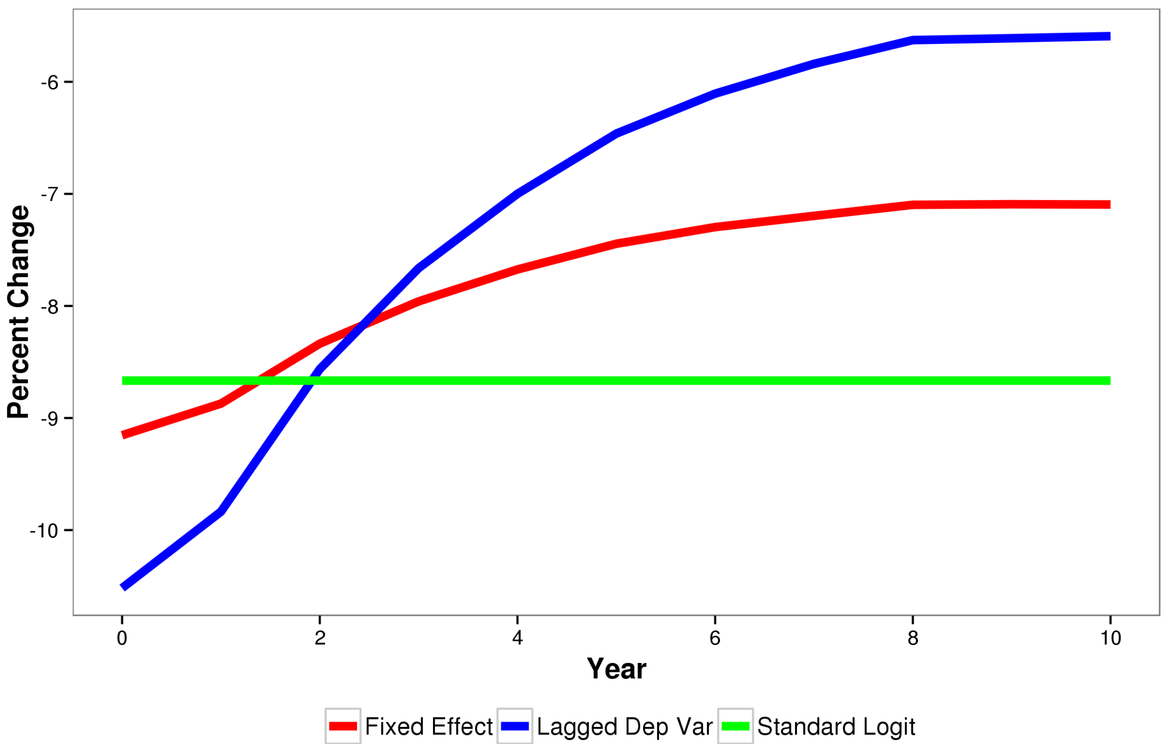


Figure 4 Percent Change in Utility from Hospital Exclusion: Previous Birth Mothers

Note: The figure depicts the change in ex-ante expected utility for each year for all hospitals after Plantation General Hospital is removed from all patients' network from the case in which it remains in the patients' choice set at its existing location for all zip codes. All estimates average across 100 simulation draws.

5.2 Entry and Post-Merger Competition

The Horizontal Merger Guidelines state that the agencies should consider the effects of post-merger entry when evaluating the potential effects of a merger on competition. The reduction in competition from a merger might attract entrants, which could mitigate part or all of the loss in competition from the merger itself. The ability of entrants to attract demand, and thus replace the lost competition due to the merger, depends upon switching costs; when switching costs are large, it might be difficult for the entrant to gain a substantial customer base quickly.

We examine how switching costs affect the ability of entrants to replace lost competition from a merger through the simulation of a merger in the PGH service area. We model a merger between PGH at its present location and Memorial Hospital System. Our simulated entering hospital enters at the proposed new site for PGH in 2015.[15] We then simulate competitive outcomes over the next ten years under different assumptions on the quality of this new hospital.

We allow the entering hospital to have five different levels of quality. In the first entry scenario, the new hospital has the mean unobserved quality ξ in each zip code. The other four scenarios take the standard deviation of ξ across all hospital-zip code pairs, and add or subtract either half or one standard deviation to the mean ξ in each zip code. For example, the best case scenario for entry has the entering hospital have the mean ξ plus one standard deviation of ξ in each zip code.[16]

To evaluate competition post-entry and merger we look at the post-merger change in "willingness to pay" ("WTP"), a commonly used metric for post-merger harm used for hospital mergers (Capps et al. (2003)).[17] In Section D.3, we describe an alternative approach using the entering hospital's share of admissions in the service area.

The left figure of Figure 5 depicts the percent change in WTP for 2015, the first year of entry, across all models and counterfactual scenarios. Without entry, all of the models predict similar increases in the percent change of WTP of about 16 to 19 percent. These

[15]See Appendix D.1 for a map of the service area.
[16]The details are in Appendix D.
[17]Appendix A.3 contains a detailed description of how we calculate the change in WTP.

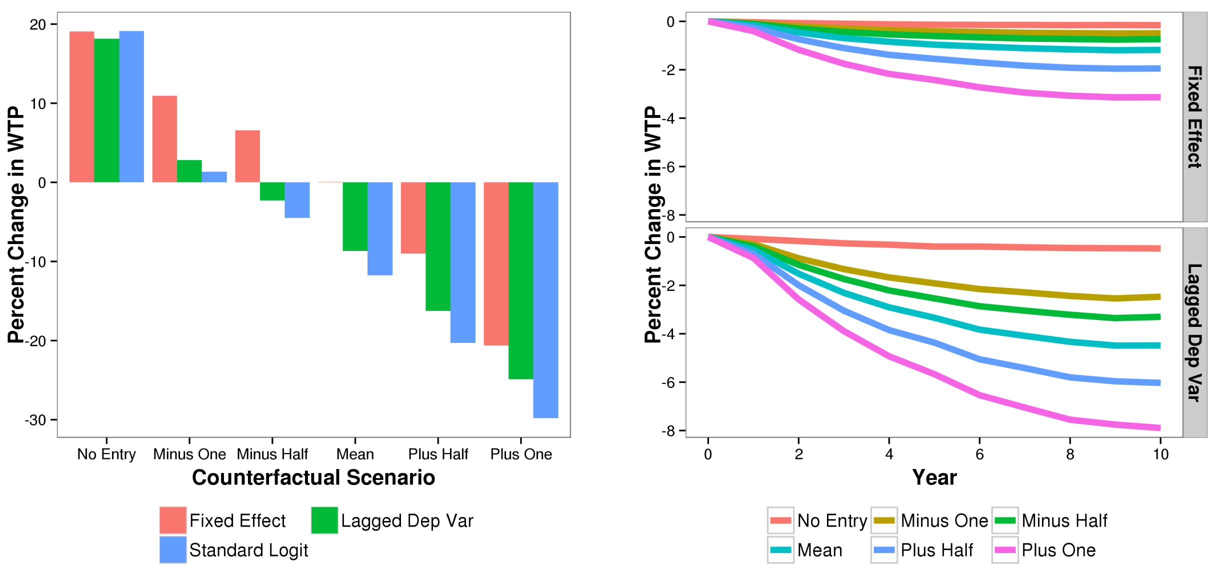
changes in WTP are large; for example, Bob Town suggests to the American Bar Association that antitrust authorities should use a 10% change in WTP as a merger screen.[18]

Different models of demand make sharply different predictions on the percent change of WTP upon entry. Under the standard logit and lagged dependent variable models, the entry of a hospital with unobserved quality a half standard deviation below the mean would still more than restore competition, with a fall in the percent change in WTP between 2 to 5 percentage points. By contrast, under the fixed effects model, the entry of a hospital with unobserved quality a half standard deviation below the mean would not fully restore competition, with a predicted rise in the percent change in WTP of 6.6 percentage points. The entry of a hospital with mean unobserved quality barely restores competition from the merger under the fixed effects model.

The reason for these differences is due to the different estimates of distance across the models; the standard logit and lagged dependent variable models have larger estimates of the effect of distance. Thus, given the entering hospital would be situated in the middle of the service area, these two models imply that it can capture substantial demand due to its location. Under the fixed effects model, unobserved quality matters more relative to distance, so the entrant has to have higher unobserved quality in order to restore competition.

The right figure of Figure 5 depicts changes over time in the percent change in WTP compared to its 2015 value. The percent change in WTP falls more sharply over time for the lagged dependent variable model, magnifying the initial differences in share between the fixed effect and lagged dependent variable models. Over 10 years, the percent change in WTP falls by 8 percentage points for the plus one standard deviation quality hospital, and 6 percentage points for the plus half standard deviation quality hospital, for the lagged dependent variable model, compared to only 3 percentage points and 2 percentage points under the fixed effects model. Since switching costs are less important under the fixed effects model, the dynamic effect of entry on competition is much smaller.

[18] See http://apps.americanbar.org/antitrust/at-committees/at-hcic/pdf/past-programs/20100601_town.pdf.



(a) At Entry

(b) Over Time

Figure 5 Percent Change in WTP for Different Counterfactual Entry Scenarios at Initial Entry and Over Time

Note: The left figure depicts the percent change in WTP in 2015, the year of entry, for the merged hospitals under six different counterfactual entry scenarios. The right figure depicts changes over time in the percent change in WTP relative to 2015. All estimates average across 100 simulation draws.

6 Conclusion

In this paper, we have first shown that a patient's previous choice is strongly predictive of her future choices. We then decomposed how much of the effect of the patient's previous choice on demand was due to persistent patient heterogeneity as opposed to switching costs by applying a fixed effects approach to the market for the hospital for childbirth. We showed that switching costs account for about 40% of the difference in predictions between models that do and do not include a lagged dependent variable.

We then applied our estimates to two applications. First, we simulated the exclusion of a hospital from a network and found smaller short run, but larger long run, welfare effects given the lower estimates of switching costs obtained from a fixed effects model. Second, we examined the effect of entry on competition after a merger. An entrant has to have higher average unobserved quality to replace lost competition from a merger under the fixed effects model. In addition, increases in competition over time are also much smaller under the lower estimates of switching costs. Thus, correctly estimating demand is important to understand the effect of entry on post-merger competition.

For future research, it would be useful to relax some of the assumptions underlying our empirical work. First, we assume that switching costs are the same whenever switching away from a hospital that one went to immediately prior. However, it is possible that the switching costs are lower when one is switching to a hospital that one went to previously, but not immediately prior. Further, while the specific conditions associated with labor and delivery make this a particularly convenient area to study these questions, it is important for future research to aim to separate preference heterogeneity from structural state dependence for patient's choice of providers beyond labor and delivery, in order to see how switching costs vary across different types of hospital admissions.[19]

Our results have implications for health policy and beyond. As managed care organizations sell more narrow network products, it is important to distinguish between preferences for a hospital from switching costs. If a provider goes out of network, patients only pay the switching cost one time, but the utility loss from going to a less preferred plan can remain for

[19] A recent paper, Irace (2017), finds evidence of structural state dependence across all conditions.

many years. In addition, the degree of switching costs inform the incentives of health care providers to enter, reposition in product space, or upgrade quality, as well as the incentives for insurers to steer patients to efficient providers.

References

Berry, Steven T, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Capps, Cory, David Dranove, and Mark Satterthwaite, "Competition and Market Power in Option Demand Markets," *RAND Journal of Economics*, 2003, *34* (4), 737–763.

Chintagunta, Pradeep, Ekaterini Kyriazidou, and Josef Perktold, "Panel data analysis of household brand choices," *Journal of Econometrics*, 2001, *103* (1), 111–153.

Dubé, Jean-Pierre, Günter J Hitsch, and Peter E Rossi, "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 2009, *46* (4), 435–445.

_ , _ , and Peter E. Rossi, "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 2010, *41* (3), 417–445.

Farrell, Joseph, David J. Balan, Keith Brand, and Brett W. Wendling, "Economics at the FTC: Hospital Mergers, Authorized Generic Drugs, and Consumer Credit Markets," *Review of Industrial Organization*, 2011, *39* (4), 271–296.

Goettler, Ronald and Karen Clay, "Tariff Choice with Consumer Learning and Switching Costs," *Journal of Marketing Research*, 2011, *48* (4), 633–652.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town, "Mergers when Prices are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 2015, *105* (1), 172–203.

Haeder, Simon F, David L Weimer, and Dana B Mukamel, "California Hospital Networks Are Narrower In Marketplace Than In Commercial Plans, But Access And Quality Are Similar," *Health Affairs*, 2015, *34* (5), 741–748.

Hamilton, Brady E., Joyce A. Martin, Michelle J.K. Osterman, Sally C. Curtin, and T.J Mathews, "National Vital Statistics Reports," *National Vital Statistics Reports*, 2015, *64* (12).

Handel, Benjamin, "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *The American Economic Review*, 2013, *103* (7), 2643–2682.

Heckman, James J, "Heterogeneity and State Dependence," in "Studies in Labor Markets," University of Chicago Press, 1981, pp. 91–140.

Ho, Kate and Ariel Pakes, "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, 2014, *104* (12), 3841–3884.

_ and Robin Lee, "Insurer Competition in Health Care Markets," *Econometrica*, 2017, *85* (2), 379–417.

Honoré, Bo E and Ekaterini Kyriazidou, "Panel Data Discrete Choice Models with Lagged Dependent Variables," *Econometrica*, 2000, *68* (4), 839–874.

Irace, Mariano, "Patient Loyalty in Hospital Choice: Evidence from New York," *manuscript*, 2017.

Raval, Devesh and Ted Rosenbaum, "Decomposing the Welfare Impacts of Medicaid MCO Steering," *mimeo*, 2016.

_ and _ , "How Strong is Gravity? Separating Home Bias from Transport Costs Using Hospital Choice," *mimeo*, 2017.

_ , _ , and Nathan E. Wilson, "Using Disaster Induced Hospital Closures to Evaluate Models of Hospital Choice," *mimeo*, 2017.

_ , _ , and Steven A. Tenn, "A Semiparametric Discrete Choice Model: An Application to Hospital Mergers," *Economic Inquiry*, 2017, *55* (4), 1919–1944.

Robinson, James C, "Hospital tiers in health insurance: balancing consumer choice with financial incentives," *Health Affairs*, 2003, p. W3.

Shepard, Mark, "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange," 2016.

Shum, Matthew, "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market," *Journal of Economics & Management Strategy*, 2004, *13* (2), 241–272.

Sudhir, K and Nathan Yang, "Exploiting the Choice-Consumption Mismatch: A New Approach to Disentangle State Dependence and Heterogeneity," 2014.

A Derivations

A.1 Derivation of equation (3)

We first outline the conditional logit estimator, based on women who switch between their first and second birth between hospitals A and B, do not move between their second and third birth, and go to hospital j at third birth.

$$\Pr(A,B,j) = \frac{\exp(\alpha d_{iA1} + \xi_{iA})}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik1} + \xi_{ik})} \times \frac{\exp(\alpha d_{iB2} + \xi_{iB})}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik2} + \xi_{ik} + \gamma I[k = A])} \times \frac{\exp(\alpha d_{ij2} + \xi_{ij} + \gamma I[j = B])}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik2} + \xi_{ik} + \gamma I[k = B])}$$

and similarly:

$$\Pr(B, A, j) =$$

$$\frac{\exp(\alpha d_{iB1} + \xi_{iB})}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik1} + \xi_{ik})} \times$$
$$\frac{\exp(\alpha d_{iA2} + \xi_{iA})}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik2} + \xi_{ik} + \gamma I[k = B])} \times$$
$$\frac{\exp(\alpha d_{ij2} + \xi_{ij} + \gamma I[j = A]))}{1 + \sum_{k=1}^{N} \exp(\alpha d_{ik2} + \xi_{ik} + \gamma I[k = A])}$$

Therefore,

$$\Pr(A, B, j)/\Pr(B, A, j)$$
$$= \frac{\exp(\alpha d_{iA1} + \xi_{iA}) \times \exp(\alpha d_{iB2} + \xi_{iB}) \times \exp(\alpha d_{ij2} + \xi_{ij} + \gamma I[j = B]))}{\exp(\alpha d_{iB1} + \xi_{iB}) \times \exp(\alpha d_{iA2} + \xi_{iA}) \times \exp(\alpha d_{ij2} + \xi_{ij} + \gamma I[j = A]))}$$
$$= \exp(\alpha((d_{iA1} - d_{iB1}) + (d_{iB2} - d_{iA2})) + \gamma(I[j = B] - I[j = A])) \quad (4)$$

A.2 Parametrization of Patient Heterogeneity

For our counterfactual experiments, we have to place further restrictions on patients' unobserved preferences for hospitals. We assume that ξ_{ij} only varies with the zip code where women live at the time of their first birth. Therefore, we can rewrite $\xi_{ij} = \mu_{gj}$, where μ_{gj} is the mean within woman i's zip code at the time of their first birth.

Using this assumption, we recover μ_{gj} by focusing on the full population of first time mothers. Rewriting equation (1) for this population of women, we obtain:

$$u_{ijt} = -\alpha d_{ij1} + \mu_{gj} + \epsilon_{ij1}$$

Since we only observe distance at the zip code level, we can write this as:

$$u_{ijt} = \underbrace{-\alpha d_{gj1} + \mu_{gj}}_{\tilde{\mu}_{gj}} + \epsilon_{ij1}$$

Since ϵ_{ij1} is distributed type I extreme value, we obtain:

$$Pr_{ij1} = \frac{\exp(\tilde{\mu}_{gj})}{\sum_j \exp(\tilde{\mu}_{gj})}$$

Using the transformation discussed in Berry (1994), we can recover:

$$\mu_{gj} = \log(s_{gj1}) - \log(s_{g01})$$

where s_{gj} is the share of first time mothers in zip code g giving birth in hospital j, and s_{g01} is the share of women giving birth outside of their service area.

A.3 Ex-Ante Utility and WTP

In the logit model, the expected decline in patient i's welfare from excluding a set of hospitals $S \subset J$ is as follows:

$$WTP_{iS} = -\ln(1 - \sum_{j \in S} s_{ij}). \tag{5}$$

A patient's WTP is an increasing function of the probability she will select a hospital in set S, and equals zero when that probability is zero. Overall WTP is obtained by adding up patient-level WTP across all patients.

The antitrust agencies have used WTP to assess the expected harm from a merger of two hospital systems (Farrell et al. (2011)). The combined system's bargaining position changes post-merger, since it can now threaten to exclude both systems simultaneously from the provider's network. Let WTP_{12} represent the WTP for the combined system, and WTP_1 and WTP_2 for System 1 and System 2 individually. If the two systems are substitutes, then the loss in welfare from simultaneously excluding both systems exceeds the sum of the losses from individually excluding each system. The percentage increase in WTP resulting from a merger between the two systems can then be calculated as follows:

$$\Delta WTP_{12} = \frac{WTP_{12}}{WTP_1 + WTP_2} - 1. \tag{6}$$

This measure has the property that it equals zero when the two systems are not substitutes, and is an increasing function of the level of substitution between the two systems. In our entry application, we calculate WTP_{12} under each counterfactual entry scenario and $WTP_1 + WTP_2$ under the pre-merger, pre-entry scenario. Thus, the percent change in WTP is the change in WTP post-merger, post-entry compared to the pre-merger, pre-entry case.

B Using More Observables to Estimate Switching Costs

The gap between switching cost estimates in the lagged dependent variable and fixed effects models reflects preference heterogeneity's role in driving choices. However, the question remains whether it is possible to narrow that gap by allowing for more preference heterogeneity by using more observable variables. If all of the unobserved preference heterogeneity could be captured using observable variables, the coefficient on the lagged dependent variable after including those variables in the logit estimation should be equal to the fixed effects (structural) estimate of switching costs.

To test this, we allow for preference heterogeneity for hospitals to vary across different groups of patients, where those groups are defined by patients' location of residence, DRG, payor type, and ethnicity. Most state hospital discharge datasets contain this level of detail on patients. Within each of those groups, we recover group specific δ_js. To compute these δ_js, we only use data on women's first birth so that there are no switching costs. Therefore, under the assumption that this preference heterogeneity is at the group level, these estimates give us preferences for the hospitals within each narrowly defined group of women. To construct these groups, we use the procedure in Raval et al. (2017b).[20]

With these group specific δ_js in hand, we estimate the coefficient on a lagged dependent variable using a standard logit model with an "offset" for these group specific δ_js. In this way, we control for group level heterogeneity. For this calculation, we use the third births of the "Honore" sample of women.

We find that including these group specific δs only reduces the coefficient on the lagged dependent variable in the estimation from 1.61 to 1.45. However, the fixed effects estimate of switching costs is 0.89, so significant individual preference heterogeneity remains unaccounted for even after accounting for group specific preference heterogeneity using a rich set of observable covariates.

C Monte Carlo Simulations

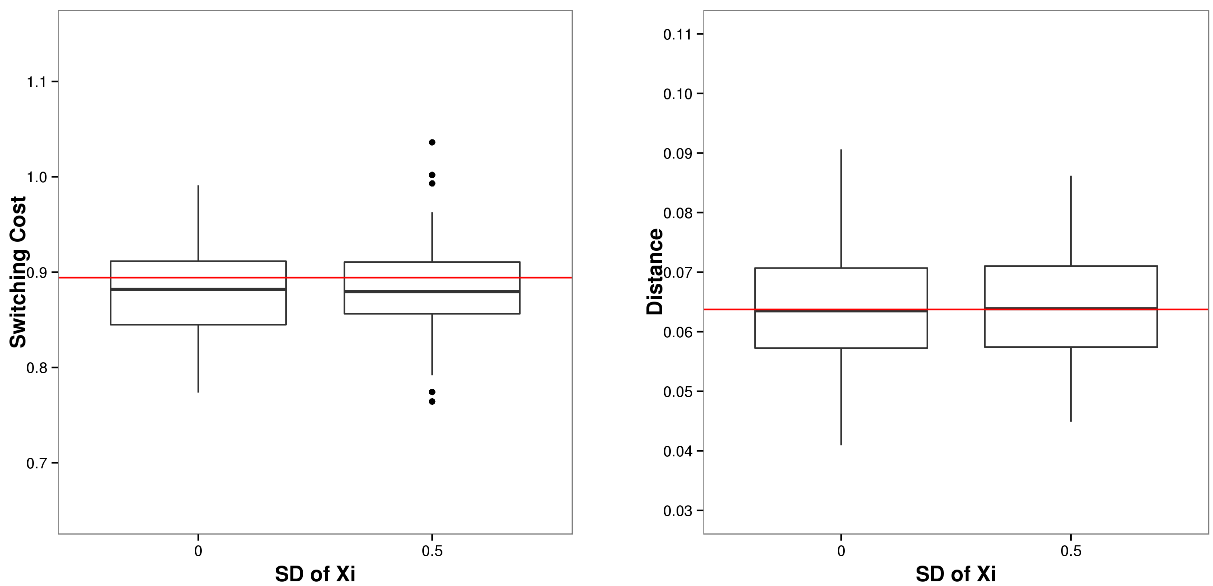
In order to examine the performance of the Honore estimator under a misspecified data generating process, we conduct a series of Monte Carlo experiments. We specify the following data generating process for utility:

$$
\begin{aligned}
u_{ijt} &= \underbrace{-\alpha d_{ijt} + \xi_{ij} + \psi_{ijt} + \gamma I[j = H_{it-1}]}_{\delta_{ijt}} + \epsilon_{ijt} \\
\xi_{ij} &= \Xi_{zj} + s_\xi u_{ij} \\
\psi_{ij1} &= \frac{s_\psi}{\sqrt{1-\rho^2}} v_{ij1} \\
\psi_{ijt} &= \rho \psi_{ijt-1} + s_\psi v_{ijt}
\end{aligned}
$$

where ϵ_{ijt} are i.i.d. type I extreme value shocks and u_{ij} and v_{ijt} are i.i.d. standard normal shocks. We set α to 0.0635 and γ to 0.8944, our baseline fixed effect estimates. We set fixed unobserved preferences ξ_{ij} to equal Ξ_{zj}, which are the fixed unobserved preferences at the zip code-hospital level for first births required to rationalize first birth zip code-hospital shares given our distance estimate α, as in Appendix A.2, and an i.i.d. zero mean normal shock with standard deviation s_ξ.

We then allow a time-varying shock ψ. In the first period, ψ is an i.i.d. normal shock with standard deviation $\frac{s_\psi}{\sqrt{1-\rho^2}}$. In subsequent periods, ψ is the sum of the previous ψ shock multiplied by ρ, and a new i.i.d. normal shock with standard deviation s_ψ. This setup keeps the time-varying shock ψ to have the same standard deviation across periods, $\frac{s_\psi}{\sqrt{1-\rho^2}}$, with correlation across time

[20]In particular, we use the variables: patient county, patient zip code, DRG, payor, and Hispanic. We use a minimum group size of 100 (for first births) and the group ordering listed in the text.



(a) Switching Costs (b) Distance

Figure 6 Monte Carlo Simulations with Fixed, Unobserved Preferences ξ

Note: The red, horizontal line is the true value in the Monte Carlo simulation. The x axis is the standard deviation of ξ at the zip code-hospital level. All boxplots are across 100 Monte Carlo simulations.

periods ρ.

For a given set of parameters for the standard deviation of the shock to ξ, s_ξ, the standard deviation of the shock to ψ, s_ψ, and the correlation of ψ across periods, ρ, we simulate hospital choices for all women with three births in Florida from 2006 to 2014. We then develop the "Honore" sample for each simulation and estimate distance and switching cost coefficients on that sample. For each Monte Carlo, we run 100 such simulations, keeping the same random seeds across different Monte Carlo experiments.

We first examine the model without misspecification, i.e. when the time-varying shock is zero, and vary the degree of heterogeneity in fixed unobserved preferences, using a standard deviation of ξ at the zip code-hospital level of 0.0 and 0.5. Figure 6 contains boxplots for switching costs in the left figure and for distance in the right figure. In each figure, the x-axis is the standard deviation of ξ at the zip code-hospital level.

At median, estimates of distance and switching costs are very close to their true values when the model is specified correctly. Compared to a true value of 0.89, the median switching cost estimate is 0.88 with a standard deviation of ξ of 0.0 and 0.5. Compared to a true value of 0.0635, the median distance estimate with a standard deviation of ξ of 0.0 is 0.0635, and with a standard deviation of ξ of 0.5 is 0.0638. Thus, on average, the Honore estimator exhibits very little bias when estimating distance and switching costs when it is specified correctly.

We then allow time-varying shocks. We examine both three different values of the standard deviation of ψ, at 0.1, 0.5, and 1.0, and three different values of the correlation between periods ρ, at 0.1, 0.5, and 0.9. We keep the standard deviation of ξ at 0.5, given, as shown above, changing the standard deviation of ξ has little effect on the distribution of estimates. Figure 7 contains boxplots for switching costs in the left figure and for distance in the right figure. In each figure, the x-axis is the standard deviation of ψ; each panel is a different value of the correlation ρ.

We find substantial deviations from the truth when the standard deviation of the time varying shock is large, with a decrease in estimated switching costs when the autocorrelation across births is low and an increase in switching costs when the autocorrelation across births is high. When the standard deviation of the time-varying shock ψ is low, at 0.1, the median estimate of switching costs lies between 0.87 and 0.88 across different values of the autocorrelation. However, when the standard deviation of ψ is high, the median estimate of switching costs is 0.77 when the correlation is 0.1, compared to 0.92 when the correlation is 0.5 or 0.9. A standard deviation for the time varying shock ψ of 1 is quite large; a one standard deviation increase in the shock is larger in magnitude than the effect of switching costs. Given the large variance in the shock, the average bias is relatively small at 0.13 for a low autocorrelation and 0.03 for a high autocorrelation.

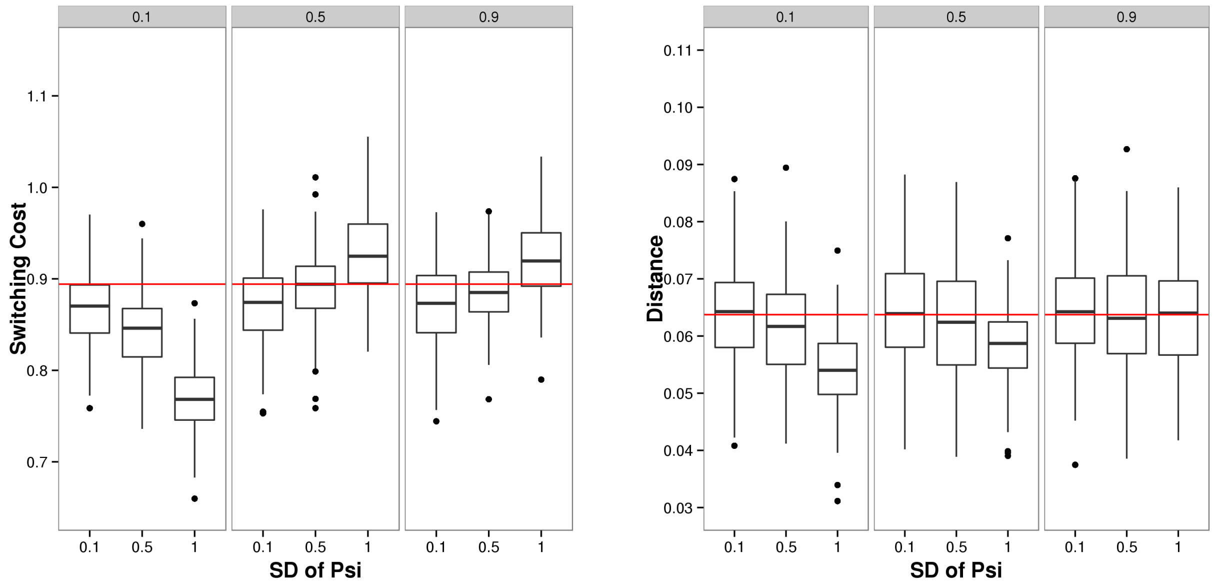
For distance, we also only find substantial deviations from the truth when the standard deviation of the time varying shock is large, but distance effects fall in magnitude for all values of the autocorrelation parameter. When the standard deviation of the time-varying shock ψ is low, at 0.1, the median estimate of distance is about 0.064 across different values of the autocorrelation. However, when the standard deviation of ψ is high, the median estimate of distance is 0.054 when the correlation is 0.1, 0.059 when the correlation is 0.5, and 0.064 when the correlation is 0.9. Again, the magnitude of the bias is larger when the correlation is low, but the bias shrinks the distance coefficient towards zero regardless of the size of the correlation coefficient.

D Supplemental Simulation Information

D.1 Service Area For Simulation

For our service area, we use the service area defined by HCA in its proposed move of Plantation General Hospital (PGH). In 2014, HCA proposed to relocate its existing location of PGH to the campus of Nova Southeastern University (NSU) in the town of Davie in Broward County, Florida. In Figure 8, the existing site is hospital number one and the new hospital site the purple star. The relocated hospital would have 200 beds after the move, including 32 dedicated OB beds, down from 264 at the original hospital site, and would be 6.7 miles or 13 to 20 minutes drive from the original hospital site.

Since Florida has a Certificate of Need (CON) Law, the construction of a new hospital required approval from the Florida Agency for Health Care Administration (AHCA), which had rejected an earlier plan by HCA to build a new 100 bed hospital at the same site after five hospital systems in the area filed statements of opposition. As part of the CON application process, HCA defined



(a) Switching Costs (b) Distance

Figure 7 Monte Carlo Simulations with Time Varying, Unobserved Preferences ψ

Note: The red, horizontal line is the true value in the Monte Carlo simulation. The x axis is the standard deviation of ψ. The three different panels reflect three values of the correlation between ψ over time at 0.1, 0.5, and 0.9. All boxplots are across 100 Monte Carlo simulations.

a service area consisting of 17 zip codes for the relocated hospital. Figure 8 displays these zip code areas, with the closer areas in yellow, constant areas in purple, and farther areas in green comprising the service area.

For our first application, we simulate the exit of PGH from this service area at its current location (i.e., hospital number one). For our second application, we simulate the merger of PGH with the Memorial Health system, which constitutes hospitals 11 through 15 in Figure 8. Only Memorial Regional Hospital, Memorial Hospital West, and Memorial Hospital Miramar are competitors for OB services in our service area. The entering hospital is located at the proposed site for PGH, i.e. the star in Figure 8.

D.2 Details of Simulation

We simulate cohorts of mothers and their hospital choices in order to examine how demand and utility changes for mothers in the service area given the network exclusion or hospital entry. Our simulations simulate the period from 2015 to 2025; for each year, we simulate new mothers based on the estimated fraction of new mothers multiplied by predicted total demand for each zip code from the CON application. We then randomly sample previous birth mothers to have a birth in a given year using the estimated fraction of previous birth mothers with previous births from one to eight

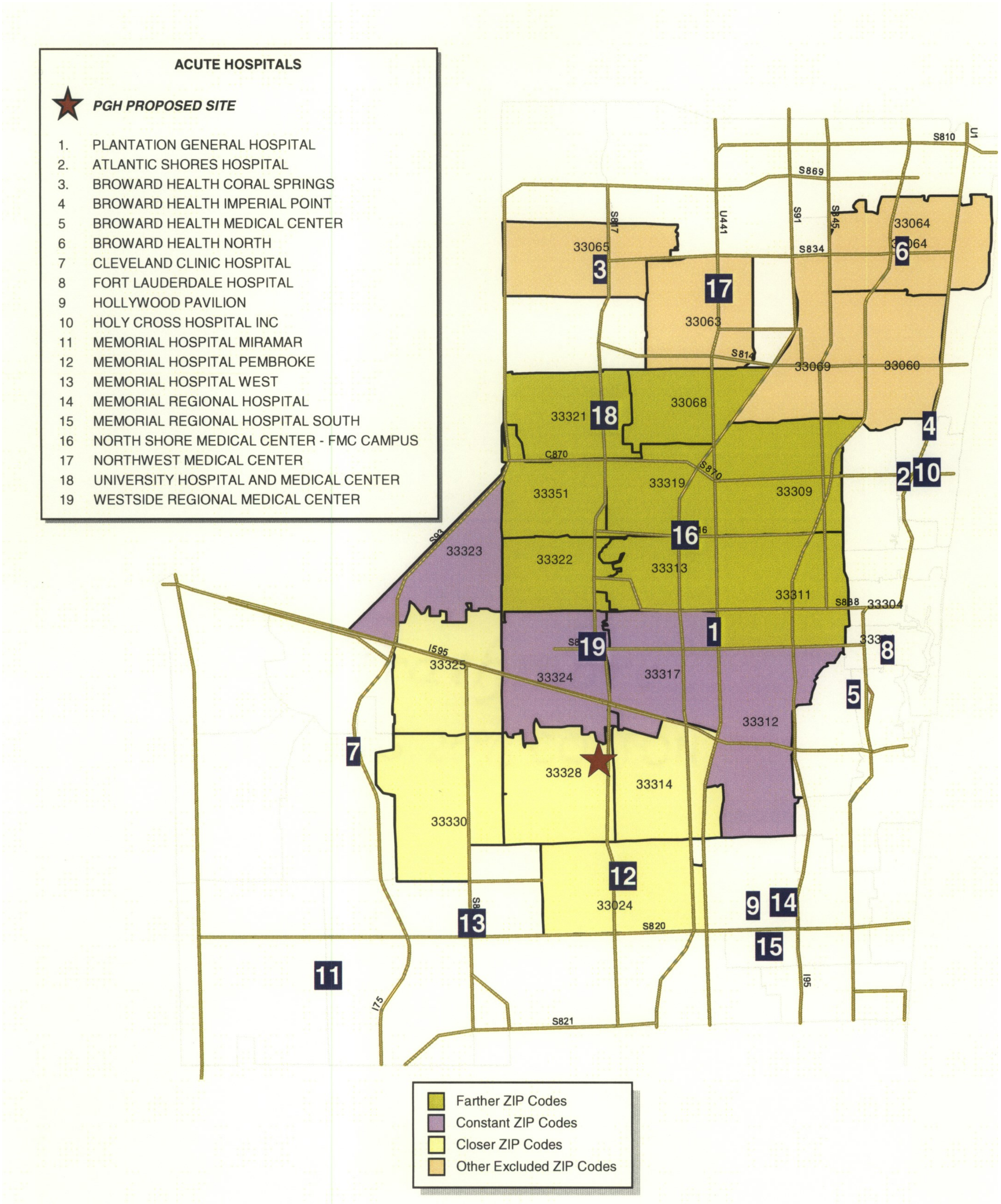


Figure 8 Taxonomy of Service Area in Second CON Application

Note: Hospital Number 1 is the existing site of Plantation General Hospital, while the star is the proposed new site for the hospital. The 17 zip code service area consists of the Closer, Constant, and Farther zip codes. Source is the opposition statement of North Broward Medical District to CON Application No. 10235.

years previous multiplied by predicted total zip code demand.[21] We use data from all births in 2014 to estimate the fraction of new mothers and the fraction of previous birth mothers whose previous birth was one year to eight years previous. For each mother, we randomly select the hospital chosen based on each hospital's utility, where we use the distance and switching cost estimates from either the fixed effects, lagged dependent variable, or standard logit models depending upon the simulation conducted.

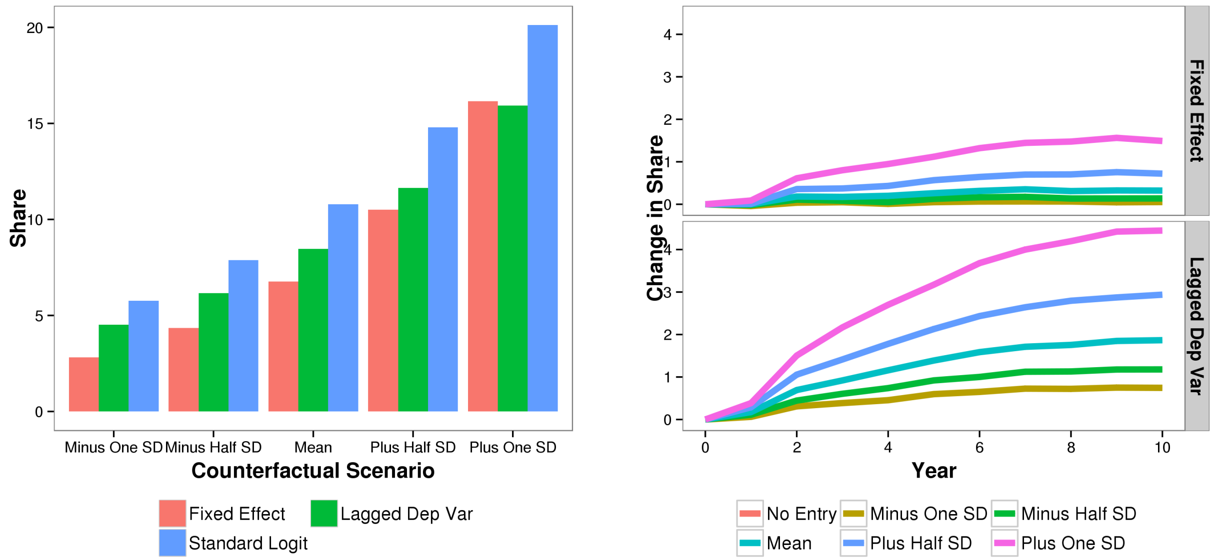
For the network exclusion, we estimate simulations both assuming that the hospital was excluded from patient choice sets, or was not excluded from patient choice sets. For the hospital entry case, we estimate six different counterfactual simulations. We first estimate the mean value of unobserved heterogeneity ξ across hospitals under each demand model for each zip code, as well as the standard deviation across ξ across all zip codes and hospitals. Each scenario sets the entering hospital at the proposed relocation site for PGH. The first scenario gives the entering hospital the mean value of ξ for each zip code, the second scenario the mean value minus one half the standard deviation, the third scenario the mean value minus the standard deviation, the fourth scenario the mean value plus one half the standard deviation, and the fifth scenario the mean value plus the standard deviation. The last scenario examines the case without any entry.

D.3 Simulation Results Using Market Shares

The left figure of Figure 9 depicts the share of the entrant hospital across all models and counterfactual scenarios in 2015, the first year of entry. The share of the entering hospital is higher for the lagged dependent variable model compared to the fixed effects model for each counterfactual scenario, and higher for the standard logit model compared to the lagged dependent variable model, for almost all scenarios. For example, for the mean scenario, the share of the entrant hospital is 6.8 percent under the fixed effects model, 8.5 percent under the lagged dependent variable model, and 10.7 percent under the standard logit model. The greater value of distance means that the lagged dependent variable and standard logit models have a higher share of the entering hospital than the fixed effects model, while switching costs suppress the share of the entering hospital for the lagged dependent variable model compared to the standard logit model.

However, as the quality of the entrant rises, the gap in predicted entrant share between the lagged dependent variable and fixed effects models shrinks, while the gap between the lagged dependent variable and standard logit models rises. In addition, changes in the share of the entering hospital over time are much more dramatic under the lagged dependent variable model. The right figure in Figure 9 depicts changes over time in the share of the entering hospital. For the plus one standard deviation entry scenario, the share of the entering hospital goes up by 4.5 percentage points over a decade under the lagged dependent variable model, compared to only 1.5 percentage

[21] For example, if total predicted demand in the zip code was 1000, the fraction of new mothers 50%, and 30%, 15%, and 5% the fraction of births from women with a previous birth 1, 2 and 3 years previously, we would randomly draw 500 new mothers, 300 mothers from women with a previous birth one year prior, 150 mothers from women with a previous birth two years prior, and 50 mothers from women with a previous birth three years prior.



(a) At Entry

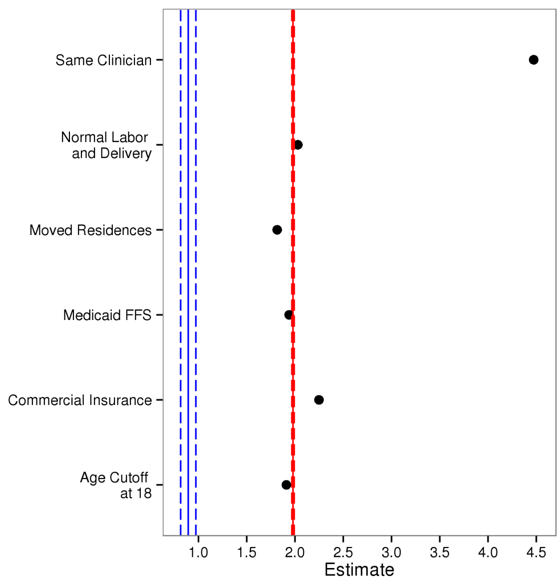
(b) Over Time

Figure 9 Percent Change in Entering Hospital Share for Different Counterfactual Entry Scenarios at Initial Entry and Over Time

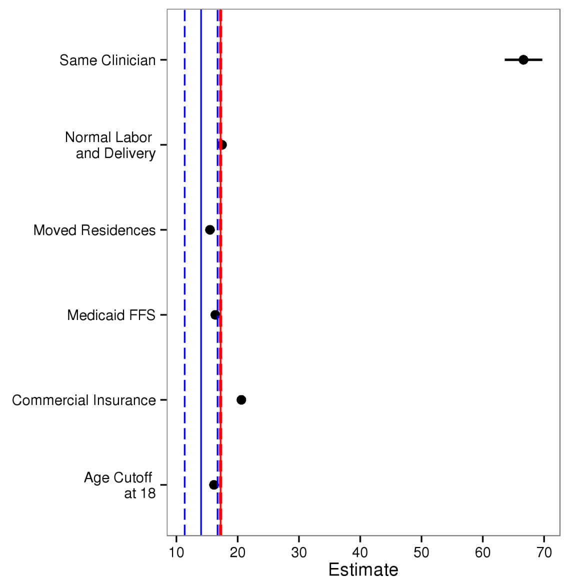
Note: The left figure depicts the market share of the entering hospital in 2015, the year of entry, for the merged hospitals under six different counterfactual entry scenarios. The right figure depicts changes over time in the share of the entering hospital. All estimates average across 100 simulation draws.

points under the fixed effects model. Again, the dynamic effects of entry are much more muted under the fixed effects model because switching costs are lower.

E Additional Graphs and Tables



(a) Switching Costs



(b) Switching Costs (min)

Figure 10 Robustness Checks for Estimates of Switching Costs under Lagged Dependent Variable Model

Note: The red lines are the parameter estimates from the lagged dependent variable model while the blue lines are the parameter estimate from the fixed effects model. The solid lines are the point estimates while the dashed lines are the 95% confidence interval, all of which are based upon the lagged dependent variable model. The black horizontal lines are the parameter estimates from our robustness checks. The dot is the point estimate and the lines are the 95% confidence interval. See Table V for a table of the estimates and standard errors used to generate this figure.

Table III Parameter Estimates: Main Results

	Lagged Dep Var (Full)	Lagged Dep Var (Restricted)	Fixed Effect
Travel Time	-0.12	-0.10	-0.06
	(0.00)	(0.00)	(0.01)
Previous Choice	1.98	1.61	0.89
	(0.01)	(0.04)	(0.04)
Previous Choice / Travel Time	17.20	15.98	14.03
	(0.06)	(0.66)	(1.37)
Hospital/Patient Fixed Effects			X
N	626,738	3,883	3,883

Note: All specifications include time invariant hospital indicator variables. N gives the units of observation used to compute the asymptotics. For the models that include hospital/patient fixed effects, this is the number of women, while for the other discrete choice models this is the number of admissions. The age restriction sample includes all women in Florida that were of age 21 or less in 2006. The Honore sample only includes women in the age restriction sample who switch hospitals between their first and second birth, have a third birth, and do not move residences between second and third birth.

Table IV Parameter Estimates: Main Results Robustness

	Travel Time	Previous Choice	Previous Choice (min)	N
Normal Labor and Delivery	-0.07	0.97	13.09	2,147
	(0.01)	(0.05)	(1.58)	
Moved Residences	-0.06	1.08	16.79	1,581
	(0.01)	(0.07)	(1.79)	
Age Cutoff at 18	-0.07	0.83	12.44	2,065
	(0.01)	(0.05)	(1.65)	
Same Clinician	-0.03	0.70	24.98	154
	(0.04)	(0.18)	(31.53)	
Commercial Insurance	-0.09	1.10	12.36	177
	(0.04)	(0.20)	(6.53)	
Medicaid FFS	-0.06	0.83	14.87	910
	(0.01)	(0.08)	(3.28)	

Note: All specifications include time invariant hospital/patient fixed effects. N gives the units of observation used to compute the asymptotics - the number of women. The samples for each regression are restricted as follows. "Normal Labor and Delivery" further restricts to patients with three normal deliveries. "Same Clinician" only includes patients with the same operating physician for all three births. "Moved Residences" only includes women who moved residences between their first and second births. "Age Cutoff at 18" only includes women who were at most 18 in 2006. Finally, "Commercial Insurance" only includes women using commercial insurance for all three births, and "Medicaid FFS" only includes women on Medicaid Fee for Service for all three births.

Table V Parameter Estimates: Lagged Dependent Variable Results Robustness

	Travel Time	Previous Choice	Previous Choice (min)	N
Age Cutoff at 18	-0.12	1.91	16.13	359,416
	(0.00)	(0.01)	(0.08)	
Normal Labor and Delivery	-0.12	2.03	17.45	499,332
	(0.00)	(0.01)	(0.07)	
Commercial Insurance	-0.11	2.25	20.60	142,141
	(0.00)	(0.02)	(0.17)	
Medicaid FFS	-0.12	1.94	16.35	353,559
	(0.00)	(0.01)	(0.09)	
Moved Residences	-0.12	1.82	15.48	530,635
	(0.00)	(0.01)	(0.08)	
Same Clinician	-0.07	4.47	66.65	52,261
	(0.00)	(0.02)	(1.57)	

Note: All specifications include time invariant hospital indicator variables. N gives the units of observation used to compute the asymptotics - the number of admissions. The samples for each regression are restricted as follows. “Normal Labor and Delivery” further restricts to patients with a normal delivery. “Same Clinician” only includes patients with the same operating physician as their previous birth. “Moved Residences” only includes women who moved residences before their last birth. “Age Cutoff at 18” only includes women who were at most 18 in 2006. Finally, “Commercial Insurance” only includes women using commercial insurance for a given birth, and “Medicaid FFS” only includes women on Medicaid Fee for Service for a given birth.