Contents

1 Introduction 1
1.1 Market definition details 1

2 Descriptive evidence 2

3 Discrete choice model 3

4 Reflection 4
4.1 Assignment reflection 4
4.2 Course reflection 5

1 Introduction

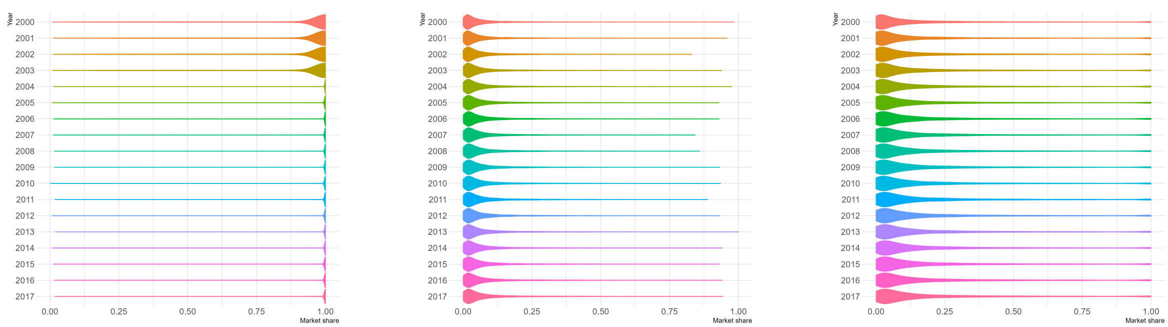
This exercise explores the empirical challenge of defining markets for health care. In particular, I look at the differences between defining markets for hospitals using ZIP codes, hospital referral regions (HRRs), and community detection algorithms, as in John Grave's recent work. All code for this project is available online at https://github.com/scai5/e771-exercise. Data used primarily comes from the Hospital Cost Report Information System for the years 2000 to 2017. More detailed data documentation is available at the Github link.

1.1 Market definition details

ZIP codes are postal codes that describe mail routes, which results in geographic areas likely unable to capture markets for patient hospital choice. ZIP codes can also change over time at the discretion of the USPS and are, in general, arbitrary areas for any kind of statistical analysis. In order to get a better sense of where patients go for their care, the Dartmouth Atlas project has defined both hospital service areas (HSAs) and hospital referral regions (HRRs). HSAs aggregate ZIP codes based on Medicare hospitalization data from 1992-1993, yielding a measure of local hospital markets. HRRs aggregate HSAs using Medicare patient referral patterns for major cardiovascular surgery and for neurosurgery, resulting in 306 HRRs. By design, HRRs are well-defined markets for tertiary care among Medicare beneficiaries, but they may not be good market definitions for other types of care, especially services that are not used by Medicare beneficiaries. This motivates the usage of community detection algorithms, which can compute markets based on data for different types of services or different patient groups. For the purposes of this exercise, the algorithm uses overall patient flows obtained from the Hospital Services Area Files (HSAF).

2 Descriptive evidence

First, I look at the distribution of market shares over time for each market definition (ZIP code, HRR, community detection algorithm). Market share is measured using hospital discharges for all three definitions of market. Discharges are a measure of demand for inpatient care. Distributions of market share over time are shown in Figure 1.



(a) ZIP code (b) HRR (c) Community detection

Figure 1: Distribution of hospital market shares of discharges over time from 2000 to 2017. Market is defined using ZIP code, HRR, and community detection algorithms. For additional enlarged graphs, see Figure A1 in the appendix.

Defining hospital markets at the ZIP code-level results in a large number of markets where there is only one hospital. Compare this to the distribution of market share when defining markets as HRR, which shows that there are almost no hospital monopolies. This is suggestive evidence that ZIP codes are geographically too small to properly capture hospital market dynamics, while HRRs may be geographically too large to capture markets for non-tertiary care. The distribution of market shares using the community detection algorithm are quite similar to that when using HRR, but the mean market share is somewhat higher. This suggests that the community detected algorithm generates markets that are larger than ZIP codes but smaller than HRRs, which indicates that the community detected markets likely pick up local market dynamics the best.

Next, I will provide some associative evidence on the relationship between price and market concentration. In particular, for each definition of market m, I estimate the equation

$$p_{ht} = \beta HHI_{m(h)t} + \lambda x_{ht} + \theta z_{m(h)t} + \gamma_h + \gamma_t + \epsilon_{ht}, \quad (1)$$

where p_{ht} is the price for hospital h at time t, $HHI_{m(h)t}$ measures market concentration, x_{ht} are time-varying hospital characteristics, $z_{m(h)t}$ are time-varying market characteristics, and γ_h and γ_t are hospital and year fixed effects, respectively. I include number of beds, estimated cost of uncompensated care, and bad debt as hospital characteristics, and market size as a market characteristic. The main results from estimating equation 1 using OLS are shown in Table 1, and full results are available in Table A1 in the appendix.

Table 1: Hospital Price and Market Concentration Association

	Price		
	ZIP	HRR	Community
Market HHI	323.447	-4213.027*	-3324.915**
	(1161.838)	(2427.175)	(1479.640)
Observations	12011	12011	12011
R^2	0.865	0.865	0.865

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
All models include number of beds, cost of uncompensated care, bad debt, number of hospitals in the market, and hospital and year FE.

Controlling for hospital and market characteristics, I find that ZIP code HHI is positively associated with price, although insignificant at the 0.10 level, while HRR and community-detected market HHI are both negatively associated with price. Taken at face value, this means that increasing ZIP code market concentration will increase price, and increasing HRR and community-detected market HHI will decrease price. These associations can be explained by the average market share coming from these different market definitions. Most ZIP codes are heavily consolidated, so increasing market HHI further generates more monopoly power. HRRs and community-detected markets have very small average market share, so some consolidation may result in hospital efficiencies, which may result in lower consumer-facing prices. This explanation, of course, abstracts away from the details of health insurance.

3 Discrete choice model

I will now estimate a logit discrete choice model using market-level data. In this context, patients choose to go to a hospital within their market. The utility to patient i when choosing hospital h is

$$U_{ih} = \alpha p_h + x_h' \beta + \xi_j + \epsilon_{ih}, \tag{2}$$

where p_h is hospital price, w_h consists of hospital characteristics, ξ_j denotes unobserved plan characteristics, and ϵ_{ih} are i.i.d. type I extreme value random variables. This yields the usual logit choice probabilities. Following the inversion approach presented in Berry (1994), the multinomial logit can be estimated using market-level data through the following equation:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{ht} + x_{ht}' \beta + \gamma_h + \gamma_t + \xi_{jt}, \tag{3}$$

where s_j denotes hospital j's market share and γ_h and γ_t are hospital and year fixed effects, respectively. Hospital characteristics again include beds, cost of uncompensated care, and bad debt. For the purposes of this exercise, there is no outside option, so I normalize $\ln(s_{0t})$ to

0. There is an obvious endogeneity concern about price, but for the purposes of this exercise, I ignore it and do not instrument for price. Note that the own-price elasticity is calculated as $-\alpha p_{jt}(1 - s_{jt})$, which can be averaged across all observations to obtain an estimator of price elasticity. The estimators of the price parameter and price elasticities obtained from estimating equation 3 using OLS are shown in Table 2, and full results are available in A2 in the appendix.

Table 2: Multinomial Logit Estimation

	ln(Market share)		
	ZIP	HRR	Community
Price	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Price elasticity	0.000	0.054	0.049
Observations	12011	12011	12011
R^2	0.960	0.988	0.970

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
All models include number of beds, cost of uncompensated care, bad debt, and hospital and year FE.

Demand is consider inelastic under all three market definitions. Price elasticity is near zero when markets at defined at the ZIP code-level. Combined with the fact that most ZIP code markets are monopolies, this further supports the descriptive evidence that zip code markets are too small to pick up on market dynamics. Demand for hospitals may be inelastic due to institutional factors like risk-sharing and insurance networks, but it seems unrealistic for it to be entirely inelastic. Among the three market definitions, demand is most elastic for the broadest market definition, HRR, and price elasticity for community-detected markets is most similar to that of HRR-defined markets. As with the market share distributions and descriptive evidence, HRRs and community-detected markets are more similar to each other in terms of price elasticity than to ZIP code markets. Overall, the price elasticities estimated here corroborate descriptive evidence and show that the broadest market definition is HRR, and the narrowest definition is ZIP code, which is unsuitable for examining hospital markets. Community-detected markets, which are the theoretically most suitable market definition, are quite similar to HRRs.

4 Reflection

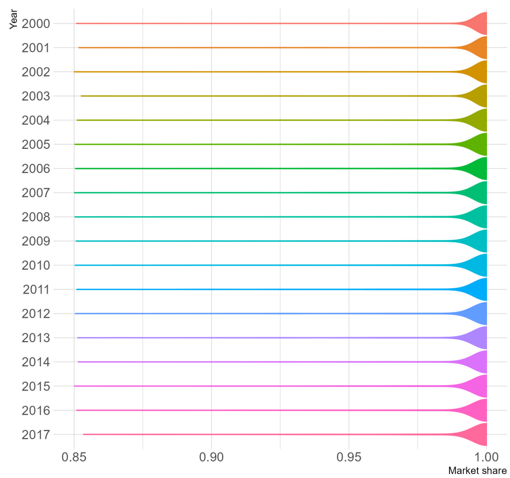
4.1 Assignment reflection

The most challenging part of this assignment was cleaning the data and choosing hospital covariates that do not not go into the estimated price. Had the community-detected markets not already been provided through the class OneDrive folder, constructing the community-detected markets would have been the hardest part of the assignment by far. The most surprising part of the assignment is how little the distribution of hospital market shares changed over time, as I expected more variation from 2000 to 2017. I was also surprised that HRRs and community-detected markets ended up being so similar in terms of price elasticity and average market share.

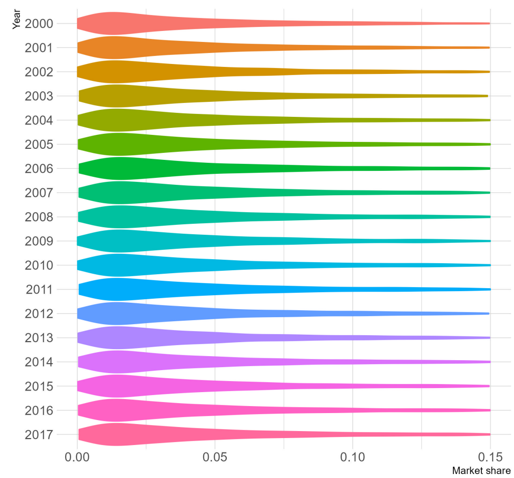
4.2 Course reflection

I thought the class was really great and was a succinct overview of the supply-side literature. I especially appreciated the inclusion of newer research papers alongside older, more classic papers. Purely based on my own research interests, I wish that more time could have been spent on health care variation.

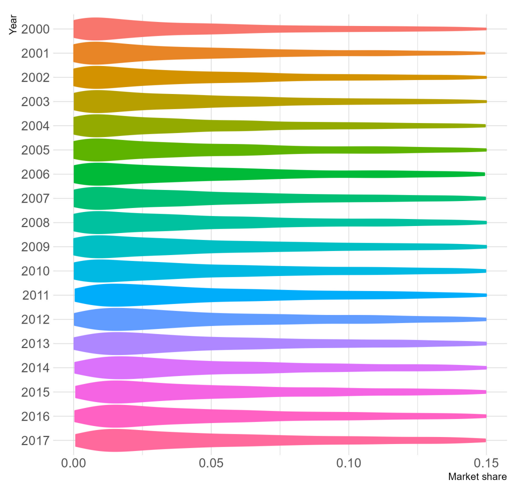
Appendix



(a) ZIP code



(b) HRR



(c) Community detection

Figure A1: Distribution of hospital market shares of discharges over time from 2000 to 2017, enlarged exclude the long tails of the distributions. Market is defined using ZIP code, HRR, and community detection algorithms.

Table A1: Hospital Price and Market Concentration Association

	Price		
	ZIP	HRR	Community
Market HHI	323.447	-4213.027*	-3324.915**
	(1161.838)	(2427.175)	(1479.640)
Market size	104.132	-7.258	-3.247
	(213.272)	(9.353)	(3.170)
Beds	0.107***	0.100**	0.107***
	(0.041)	(0.043)	(0.041)
Uncompensated care	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Bad debt	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Observations	12011	12011	12011
R^2	0.865	0.865	0.865

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Market size is measured by number of hospitals in the market.
All models include hospital and year FE.

Table A2: Multinomial Logit Estimation

	ln(Market share)		
	ZIP	HRR	Community
Price	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Beds	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Uncompensated care	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Bad debt	0.000	0.000	0.000*
	(0.000)	(0.000)	(0.000)
Price elasticity	0.000	0.054	0.049
Observations	12011	12011	12011
R^2	0.960	0.988	0.970

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
All models include hospital and year FE.