Online Appendix to "Using Disaster Induced Closures to Evaluate Discrete Choice Models of Hospital Demand"

Devesh Raval — Federal Trade Commission — draval@ftc.gov

Ted Rosenbaum — Federal Trade Commission — trosenbaum@ftc.gov

Nathan E. Wilson — Federal Trade Commission — nwilson@ftc.gov

September 22, 2021

1 Details on the *Semipar* Estimator

This section provides details on the *Semipar* estimator described in the text. This estimator is nearly identical to that described in Raval et al. (2017).

In contrast to the other estimation approaches in the text, *Semipar* models hospitals in product space. In particular, this approach estimates alternative specific constants for each group of patients (indexed by g).

$$s_j^g = \frac{e^{\delta_j^g}}{\sum_{k \in J} e^{\delta_k^g}}.$$

Given group assignments and data on shares for each hospital j recovering $e^{\delta_j^g}$ is straightforward to compute from the equation above by using the inversion from Berry (1994).

Therefore, the key choices for the researcher in this model involve how to assign patients to groups. In our application, all patients are assigned a group based upon their zip code, disease

acuity (DRG weight), age group, and area of diagnosis (MDC). Anyone in a group above the minimum group size is assigned choice probabilities based upon the share of patients in that group that go to each hospital. Then, we drop a characteristic, reconstruct groups, and again compute group-level shares for the full set of patients, both those previously grouped and those not previously grouped.[1] All patients who have not yet been assigned a choice probability and are in groups above the minimum group size are assigned a choice probability based on that round's group-level shares. We continue until there are no more covariates to group or until all patients are assigned a choice probability.[2] We drop characteristics in the reverse order listed above (i.e., MDC, age group, etc.).

The minimum group size regulates the bias-variance tradeoff; increasing the minimum group size increases the bias but reduces variance. For this article, we use a minimum group size of 50. Our approach is equivalent to estimating different multinomial logit models with group-specific alternative specific constants for each level of grouping and assigning choice probabilities to an individual based upon the most refined level of grouping that exceeds the pre-specified minimum group size. Once the researcher has constructed the set of groups, predicted choice probabilities can be estimated as the empirical shares of hospitals within each group.

Because in this approach one estimates many fixed effects in a non-linear model, there are a number of potential concerns. First, one may be concerned that due to an incidental parameters problem, our main parameter estimates will not be consistently estimated. However, in this approach, the researcher directly estimates the group level choice probabilities, which are the main parameters of interest. Therefore, there are no incidental parameters being estimated. Second, one may worry that the estimates of the group level choice probabilities are not consistent or are estimated with significant variance in finite samples. However,

[1] Raval et al. (2017) used previously non-grouped individuals to compute the choice probabilities. The approach in this article will have a higher bias, but relatively smaller variance. The approach in this article is analogous to the one in Raval et al. (2021).

[2] In this last round of grouping, we do not impose a minimum group size restriction. So, for example, if a zip code only has 10 residents, we compute choice probabilities based upon these 10 people.

the combination of an imposed minimum group size and a finite number of possible groups means that the estimates should converge for sufficiently large group sizes. In this article, we are able to test the finite sample performance of this estimator, and should observe the model's performance suffer to the extent that these parameters are imprecisely estimated.

Finally, because the group sizes are relatively small, there is a potential concern that an individual's observed choice has an important effect on their predicted choice. To address this concern, in computing individual i's mean utility for hospital j ($\hat{\delta}_k^{i(g)}$), we leave out individual i by computing s_j^g excluding individual i.

2 Estimation Details

In Table I to Table XII below, we report estimation details for all of the models except *Semipar*, including the number of parameters in the model, the estimated log likelihood of the model, the AIC and BIC criteria, and McFadden's pseudo R^2. We provide these details for both the non-elective sample and the full sample.

The total number of observations for each experiment is the number of admissions in the pre-period multiplied by the number of choices. Thus, the total number of observations for the non-elective sample was 75,492 for Moore, 1,083,342 for NYU, 560,592 for Coney, 657,100 for Bellevue, 77,955 for Sumter, and 1,432,305 for St John's. The total number of observations for the full sample was 117,156 for Moore, 1,519,050 for NYU, 791,996 for Coney, 916,263 for Bellevue, 104,100 for Sumter, and 2,030,091 for St John's.

3 Case Mix

In this section, we examine how the case mix changed from the period before the disaster to the period after the disaster. In Table XIII, we examine changes in case mix across a set of variables including age, fraction aged less than 18, fraction aged above 64, diagnosis acuity (DRG weight), fraction emergency, fraction circulatory diagnosis (MDC 5), fraction

Table I Estimation Details for Moore, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	11.00	-14480.24	28982.49	29084.04	0.07
Char	20.00	-14253.20	28546.40	28731.04	0.09
CDS	51.00	-13631.99	27365.97	27836.79	0.13
Time	13.00	-14273.72	28573.44	28693.45	0.09
Ho	51.00	-13991.55	28085.10	28555.92	0.10
GNT	36.00	-13858.01	27788.02	28120.37	0.11
Inter	122.00	-12974.07	26190.14	27307.18	0.17

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table II Estimation Details for NYU, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	18	-148420.61	296877.23	297091.35	0.12
Char	20	-118967.96	237969.93	238172.15	0.29
CDS	51	-115203.01	230506.02	231100.80	0.31
Time	20	-113485.05	227010.10	227248.02	0.32
Ho	58	-110106.66	220315.33	220922.00	0.34
GNT	50	-111472.56	223043.12	223626.00	0.34
Inter	185	-107852.65	216073.31	218262.09	0.36

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

labor/pregnancy diagnosis (MDC 14), fraction using a commercial payer, fraction using Medicare, and average number of admissions per month. We report the percent change for the variable in the post-period compared to pre-period. Table XIII is based upon the full sample, whereas Table XIV is based upon the non-elective sample. We summarize results for the full sample below.

There are no large changes in age across the hospitals, except that the fraction admitted under 18 falls by 23 percent for Moore and 45 percent for Sumter. Diagnosis acuity does not change much after the disasters. The only large change in type of insurance is for Sumter, where the fraction of commercial insurance falls by about 30 percent after the disaster. We examined this change; the fraction of patients reporting "Unspecified Other" payer rises

Table III Estimation Details for Coney, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	16	-75843.46	151718.93	151898.72	0.19
Char	20	-80677.82	161389.64	161580.67	0.14
CDS	51	-72383.05	144868.10	145441.18	0.23
Time	18	-73253.35	146542.70	146744.96	0.22
Ho	56	-70937.26	141972.52	142523.12	0.24
GNT	46	-72750.39	145590.78	146096.44	0.22
Inter	167	-70237.82	140807.64	142672.94	0.25

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table IV Estimation Details for Bellevue, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	19	-82180.46	164398.92	164615.24	0.16
Char	20	-74312.75	148657.51	148839.68	0.24
CDS	51	-72050.89	144201.77	144771.06	0.27
Time	21	-70584.41	141210.82	141449.92	0.28
Ho	59	-67881.36	135866.72	136458.78	0.31
GNT	52	-69888.13	139878.26	140458.93	0.29
Inter	194	-67811.66	136009.32	138206.75	0.31

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

precipitously in the first quarter after the disaster, and then falls back to a small fraction of patients. Our belief is that this reflects improper coding post-disaster.

The fraction of labor/pregnancy diagnosis rises in all service areas, and by more than 10 percent for Bellevue and Coney, which may be because pregnancies cannot be postponed or ignored and so have no extensive margin. Overall, we do not find major changes in case mix after the disaster, except for the rise in pregnancy admissions across the service areas and the fall in the under 18 share for Sumter and Moore.

Table V Estimation Details for Sumter, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	14	-8343.83	16715.66	16845.36	0.41
Char	20	-6480.57	13001.14	13186.42	0.54
CDS	51	-6186.46	12472.91	12936.11	0.56
Time	16	-6261.38	12554.76	12702.98	0.56
Ho	54	-5875.31	11858.61	12358.86	0.58
GNT	42	-5924.61	11931.21	12311.03	0.58
Inter	149	-5370.74	11037.47	12408.53	0.62

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table VI Estimation Details for St John's, Non-Elective Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	20	-168172.34	336384.67	336628.10	0.19
Char	20	-155353.88	310747.76	310991.18	0.25
CDS	51	-146277.59	292657.17	293277.90	0.29
Time	22	-144297.12	288638.24	288906.01	0.30
Ho	60	-141634.94	283389.88	284120.15	0.32
GNT	54	-141488.72	283085.44	283742.68	0.32
Inter	203	-136753.56	273911.12	276369.69	0.34

Note: All models estimated on the non-elective sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table VII Estimation Details for Moore, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	11	-22331.56	44685.12	44791.51	0.08
Char	20	-22083.47	44206.95	44400.37	0.09
CDS	51	-21075.79	42253.58	42746.81	0.13
Time	13	-21982.78	43991.56	44117.29	0.09
Ho	51	-21548.72	43199.43	43692.67	0.11
GNT	36	-21405.45	42882.90	43231.06	0.12
Inter	122	-20084.84	40411.67	41581.90	0.17

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table VIII Estimation Details for NYU, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	18	-211848.84	423733.68	423953.88	0.10
Char	20	-171159.33	342352.66	342560.63	0.27
CDS	51	-163021.86	326143.72	326755.40	0.31
Time	20	-164552.98	329145.95	329390.62	0.30
Ho	58	-159893.26	319888.53	320512.44	0.32
GNT	50	-162076.84	324251.68	324851.12	0.31
Inter	185	-158610.38	317588.75	319839.73	0.33

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table IX Estimation Details for Coney, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	16	-109623.34	219278.68	219464	0.17
Char	20	-114909.94	229853.89	230050.78	0.13
CDS	51	-103857.15	207814.29	208393.41	0.21
Time	18	-105906.74	211849.48	212057.96	0.20
Ho	56	-102327.91	204753.82	205321.35	0.22
GNT	46	-105114.44	210318.87	210840.08	0.20
Inter	167	-101809.59	203951.17	205873.83	0.23

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table X Estimation Details for Bellevue, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	19	-117919.64	235877.28	236100.11	0.15
Char	20	-104182.25	208396.50	208584.14	0.25
CDS	51	-101134.46	202368.91	202955.32	0.27
Time	21	-100109.69	200261.38	200507.66	0.28
Ho	59	-96684.78	193473.55	194083.41	0.30
GNT	52	-99120.44	198342.88	198941.01	0.28
Inter	194	-96544.97	193475.94	195739.46	0.30

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table XI Estimation Details for Sumter, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	14	-11914.76	23857.52	23991.27	0.37
Char	20	-9676.37	19392.74	19583.81	0.49
CDS	51	-9029.20	18158.40	18636.06	0.52
Time	16	-9386.21	18804.43	18957.28	0.50
Ho	49	-8876.75	17851.50	18319.61	0.53
GNT	42	-8699.91	17481.82	17873.50	0.54
Inter	149	-8094.26	16484.51	17898.37	0.57

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table XII Estimation Details for St John's, Full Sample

Model	No. Parameters	Log Likelihood	AIC	BIC	Pseudo R^2
AggShare	20	-239495.27	479030.54	479281.01	0.19
Char	20	-225640.76	451321.52	451571.99	0.24
CDS	51	-215470.24	431042.48	431681.18	0.27
Time	22	-209454.45	418952.89	419228.41	0.29
Ho	60	-206835.28	413790.55	414541.97	0.30
GNT	54	-206498.56	413105.13	413781.40	0.30
Inter	203	-199907.05	400218.11	402747.87	0.32

Note: All models estimated on the full sample. The second column is the number of parameters in the model, the third column the estimated log likelihood of the model, the fourth and fifth columns the AIC and BIC criteria, and the sixth column McFadden's pseudo R^2.

Table XIII Post-Disaster Changes in Case-Mix, in Percent Change from Pre-Period, Full Sample

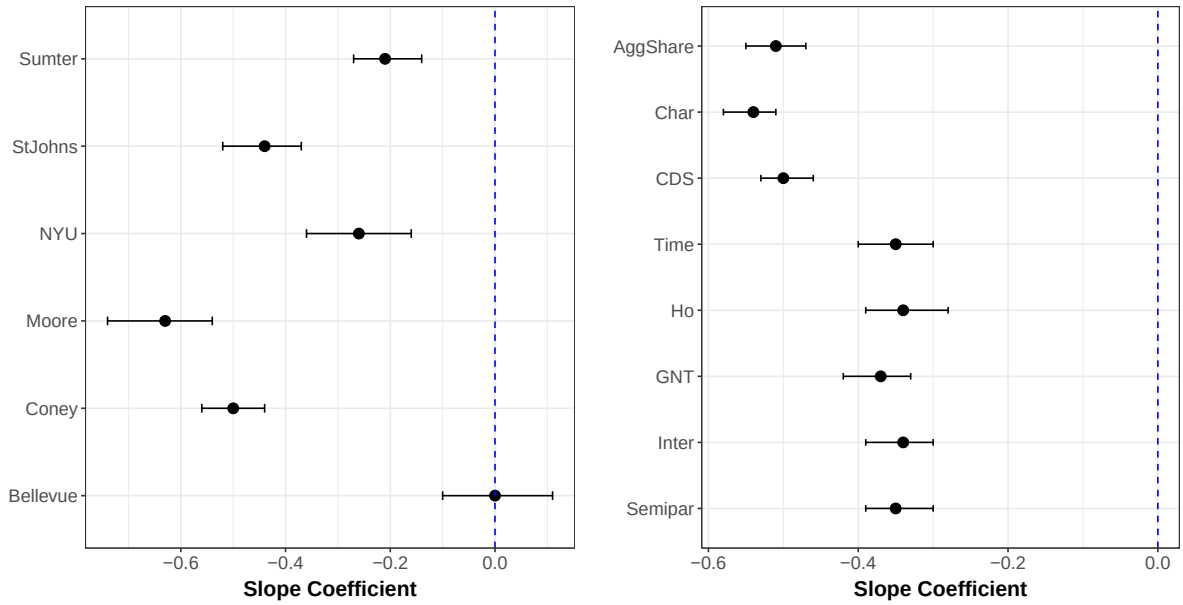
Variable	Sumter	Moore	NYU	Bellevue	Coney	StJohns
Age	1.09	0.12	0.94	2.36	0.11	-0.98
	(-0.6,2.52)	(-1.39,1.56)	(0.35,1.63)	(1.28,3.28)	(-0.82,1.15)	(-1.67,-0.31)
Age < 18	-45.43	-23.73	2.84	-12.54	2.41	11.72
	(-55.59,-37.22)	(-34.92,-12.86)	(-4.22,10.39)	(-22.07,-4.16)	(-7.43,11.06)	(3.39,19.34)
Age > 64	-4.28	-2.88	4.74	9.01	3	-2.07
	(-8.79,0.23)	(-6.93,1.33)	(3.11,6.34)	(5.76,11.86)	(0.68,5.5)	(-3.84,-0.17)
Diagnosis Acuity	3.72	2.12	1.02	3.29	3.29	3.3
	(0.29,7.02)	(-1.35,5.53)	(-0.51,2.61)	(0.67,6.14)	(1.12,6.07)	(1.68,5.1)
Emergency	-11.34	18.02	2.47	0.44	-7.51	-1.22
	(-15.88,-7.3)	(11.29,24.57)	(1.04,3.74)	(-1.04,1.88)	(-8.88,-6.01)	(-3.46,0.7)
Circulatory Diagnosis	11.28	-12.1	-7.82	-7.04	-5.24	5.49
	(3.68,19.63)	(-22.4,-2.78)	(-11.57,-4.5)	(-12.47,-1.95)	(-9.15,-0.87)	(1.61,9.28)
Labor/Pregnancy Diagnosis	7.36	7.05	7.03	10.51	13.62	5.62
	(-2.03,15.61)	(0.38,13.78)	(3.89,10.24)	(5.16,15.25)	(7.78,19.11)	(2.43,9.22)
Commercial Payer	-28.41	3.88	-2.84	-2.06	-6.08	6.19
	(-33.45,-24.22)	(-1.13,7.92)	(-5.07,-0.27)	(-5.73,1.55)	(-10.12,-2.13)	(4.52,8.02)
Medicare Payer	-4.94	-1.91	4.91	9.24	2.18	-8.79
	(-8.32,-1.35)	(-5.19,2.24)	(3.07,6.75)	(6.19,11.98)	(-0.05,4.49)	(-10.8,-6.58)

Note: Estimates based upon the Full Sample. Each cell reports the percent change for a given variable in the post-period compared to pre-period for a given experiment, along with the 95% confidence interval.

Table XIV Post-Disaster Changes in Case-Mix, in Percent Change from Pre-Period, Non-Elective Sample

Variable	Sumter	Moore	NYU	Bellevue	Coney	StJohns
Age	0.28	-1.66	1.34	1.49	-0.07	-1.99
	(-1.45,2.42)	(-3.72,0.07)	(0.64,2.03)	(0.22,2.71)	(-1.16,0.93)	(-2.9,-1.13)
Age < 18	-33.26	-21.66	8.38	-8.94	4.06	10.91
	(-45.07,-21.49)	(-38.38,-5.75)	(-2.45,18.26)	(-18.9,1.28)	(-6.87,15.45)	(2.44,20.17)
Age > 64	-4.79	-5.74	6.37	7.29	2.5	-5.31
	(-10.4,-0.29)	(-12.58,-0.23)	(4.39,8.77)	(2.79,11.21)	(-0.17,5.35)	(-7.65,-3.33)
Diagnosis Acuity	4.68	-1.29	2.84	3.08	4.03	2.69
	(1.07,8.64)	(-4.92,2.3)	(0.88,4.87)	(0.13,6.14)	(1.28,6.67)	(0.22,5.01)
Emergency	-14.68	18.38	2.06	0.45	-8.5	-2.2
	(-18,-11.08)	(8.53,26.89)	(0.76,3.39)	(-1.71,2.37)	(-10.25,-7.08)	(-3.88,-0.5)
Circulatory Diagnosis	6.15	-7.54	-7.96	-8.53	-7.41	1.42
	(-3.9,15.08)	(-26.05,11.01)	(-12.34,-3.67)	(-14.62,-3.57)	(-13,-2.1)	(-2.02,5.09)
Labor/Pregnancy Diagnosis	3.76	11.86	2.94	10.81	9.77	4.55
	(-6.05,12)	(3.13,18.19)	(-0.27,5.93)	(5.79,17.17)	(4.67,15.12)	(1.19,8.11)
Commercial Payer	-27.46	3.89	1.33	-0.72	-7.33	6.19
	(-33.2,-22.56)	(-2.77,9.57)	(-1.28,3.95)	(-5.13,3.24)	(-12.05,-1.8)	(3.85,8.28)
Medicare Payer	-8.4	-6.2	5.33	7.25	1.77	-11.48
	(-12.83,-3.53)	(-11.64,-1.35)	(3.13,7.39)	(3.55,10.73)	(-0.72,4.41)	(-13.72,-9.47)

Note: Estimates based upon the non-elective sample. Each cell reports the percent change for a given variable in the post-period compared to pre-period for a given experiment, along with the 95% confidence interval.



(a) By Experiment (for *Semipar*) (b) By Model

Figure 1 Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, Non-Elective Sample

Note: The left figure depicts the slope of the observed diversion ratio on the prediction error by model and the right figure depicts the same by experiment for the *Semipar* model. Bars represent 95% confidence intervals computed from 200 bootstrap replications; we also apply a bootstrap bias correction. See Table XXIV and Table XXV for tables of the estimates and confidence intervals used to generate these figures.

4 Supplemental Figures: Full Sample

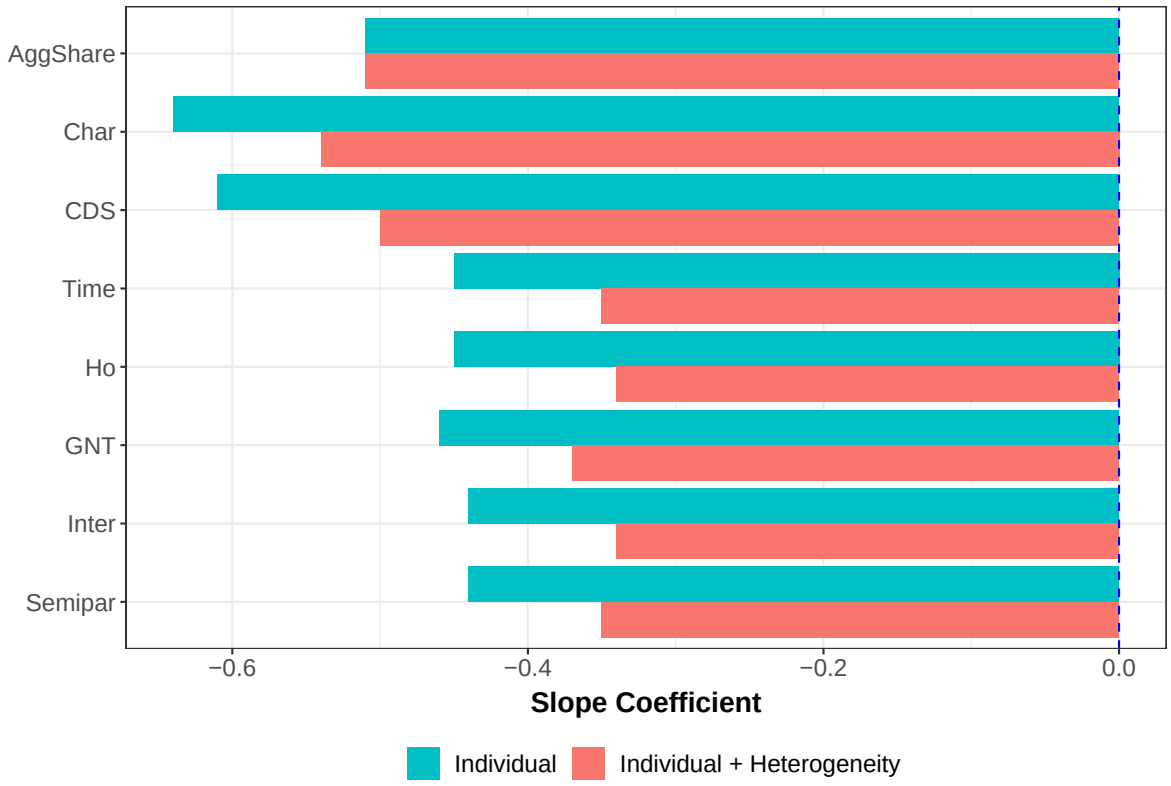
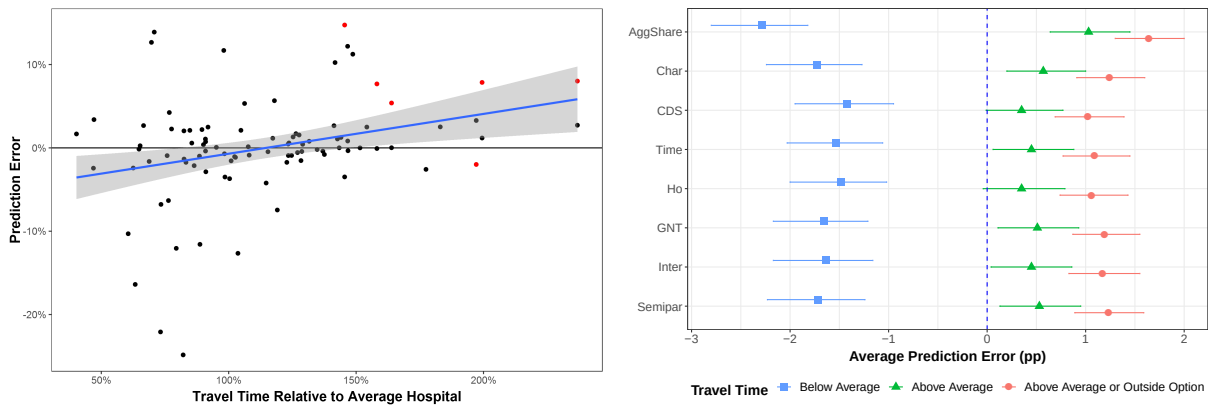


Figure 2 Decomposition of Average Predicted Diversion, Full Sample

Note: We report the slope coefficient of the observed diversion ratio on the prediction error based upon the average individual diversion ratio in blue, and based upon the individual diversion ratio plus the heterogeneity factor (i.e. the total predicted diversion) in red, for each model. Each term is as defined in the text. See Table XXVI for a table of the estimates and confidence intervals used to generate this figure.



(a) By Model

(b) Average Prediction Error by Distance Group

Figure 3 Prediction as a Function of Distance, Full Sample

Note: First panel shows the prediction error as a function of the average travel time to the hospital expressed as a percent of the average travel time in the market. Second panel presents the average prediction error, differentiating between hospitals whose travel time is below average for their market, above average for their market, or above average plus the Outside Option. Bars represent 95% confidence intervals computed from 200 bootstrap replications; we also apply a bootstrap bias correction. See Table XXVII for tables of the estimates and confidence intervals used to generate these figures.

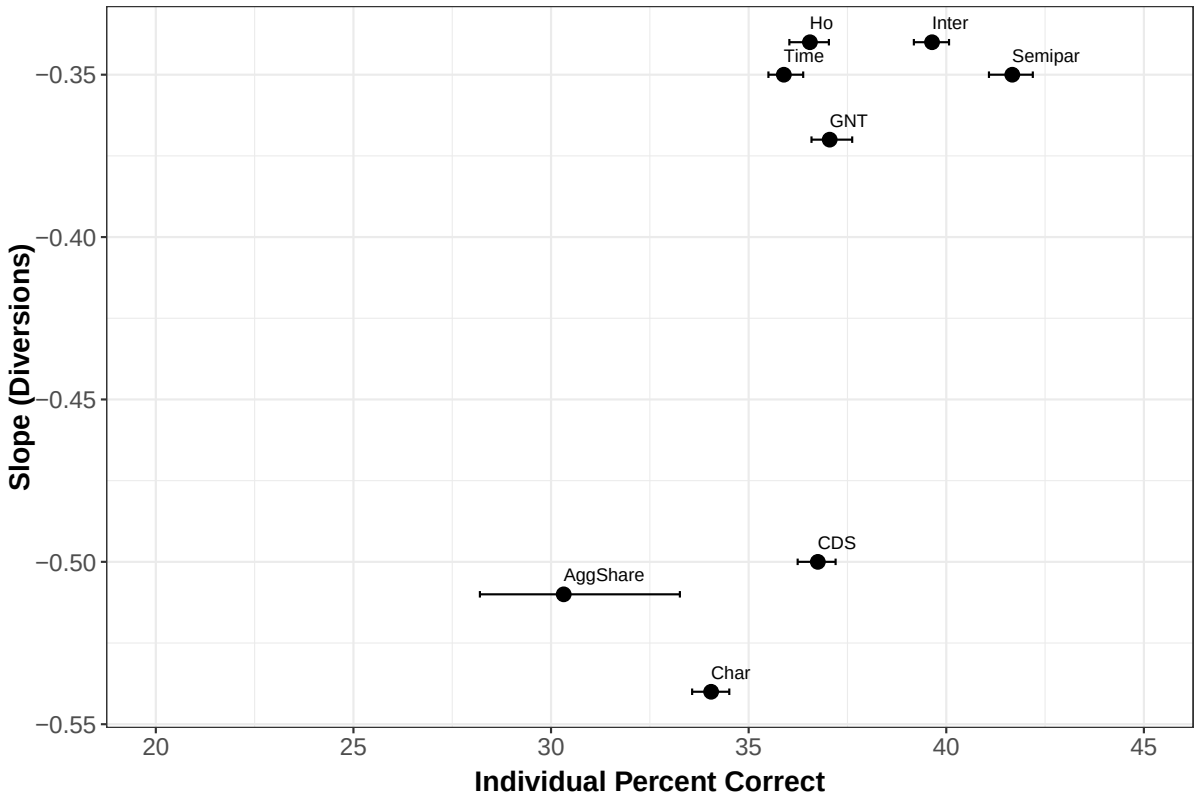


Figure 4 Average Percent Correct of Individual Predictions vs. Slope Coefficient of Observed Diversion Ratio and Prediction Error, Full Sample

Note: The Figure compares the slope coefficient of the observed diversion ratio on the prediction error to the average percentage of individual choices correctly predicted. Bars represent 95% confidence intervals computed from 200 bootstrap replications; we also apply a bootstrap bias correction. See Table XXIV and Table XXVIII for estimates and confidence intervals used to generate the figure.

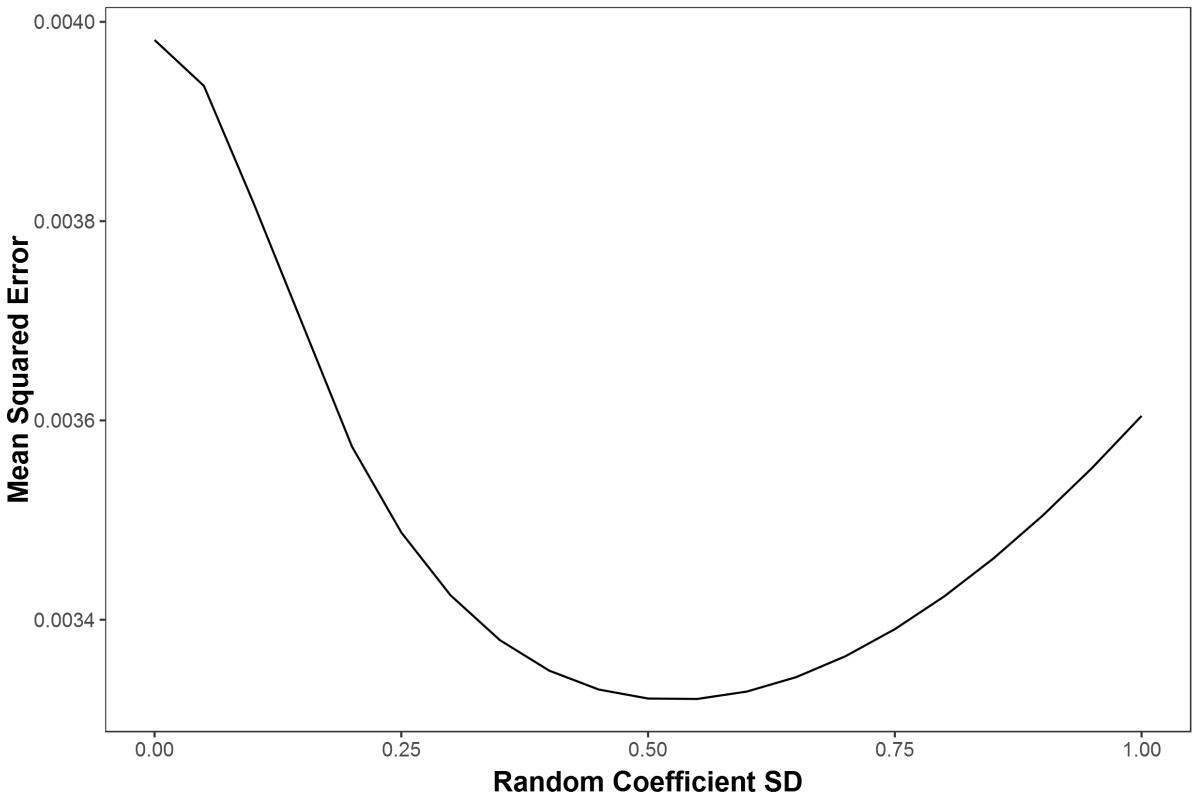
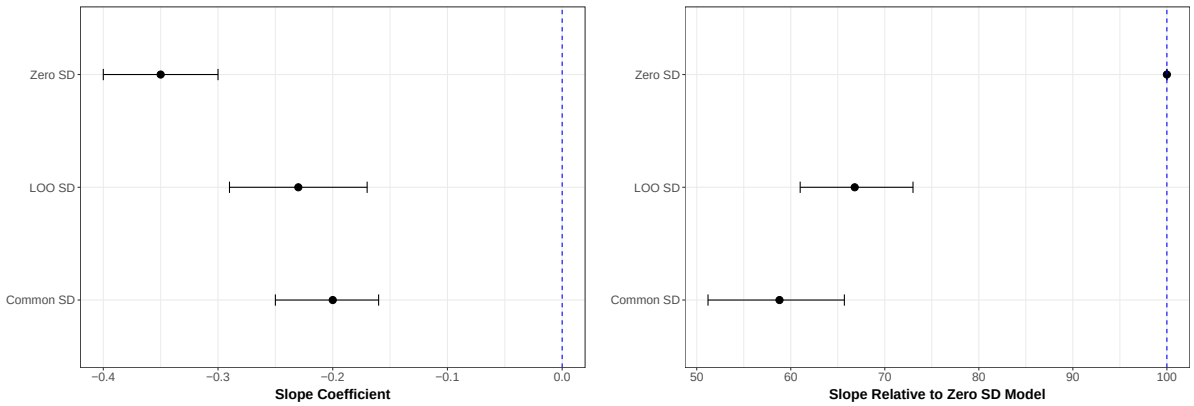


Figure 5 MSE by Standard Deviation of Random Coefficient, Full Sample

Note: This figure shows the mean squared prediction error (MSE) of the *Common SD* class of models described in the text, where the models vary by the standard deviation of the random coefficient.



(a) By Model (b) Correlation by Distance

Figure 6 Random Coefficient Relative Performance, Full Sample

Note: The left panel presents the slope of the observed diversion ratio on the prediction error. The right panel depicts the slope for a model relative to that for the Zero SD model. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction. See Table XXXI for estimates and confidence intervals used to generate the figure.

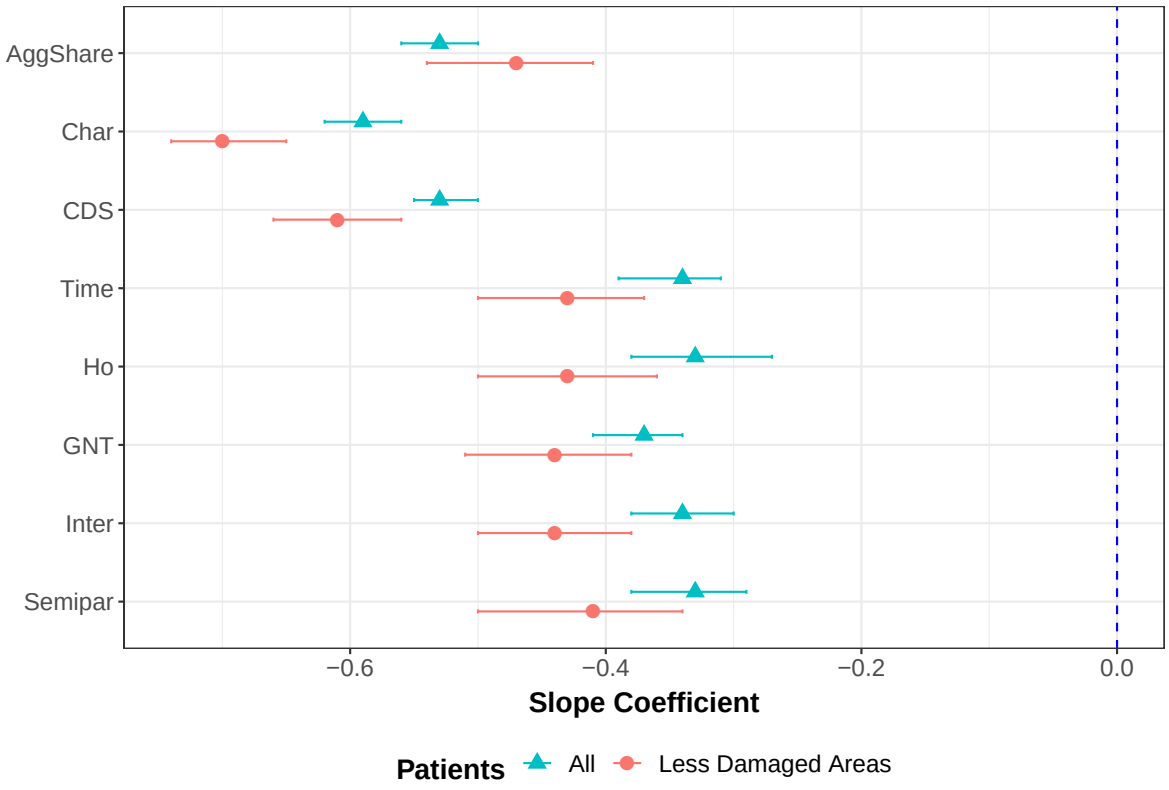


Figure 7 Slope Coefficient of Observed Diversion Ratio on Prediction Error By Disaster Damage, Full Sample

Note: The figure depicts the slope of the observed diversion ratio on the prediction error. The figure uses only data from the Coney, St. John's, or Sumter experiments, and includes either all patients or patients in zip codes with less disaster damage. Bars represent 95% confidence intervals computed from 200 bootstrap replications; we also apply a bootstrap bias correction. See Table XXIX, and Table XXX for tables of the estimates and confidence intervals used to generate this figure.

5 Supplemental Tables

5.1 Non-Elective Sample

Table XV Descriptive Statistics fo Affected Hospital Service Areas, Non-Elective Sample

	Pre-Period Admissions	Post-Period Admissions	Zip Codes	Choice Set Size	Outside Option Share	Destroyed Share	Destroyed Acuity
Sumter	5,197	3,951	11	15	3.1	51.2	0.97
StJohns	68,205	12,946	29	21	9.0	16.8	1.20
NYU	57,018	12,714	38	19	12.3	3.6	0.96
Moore	6,291	2,451	5	12	1.8	12.5	0.91
Coney	32,976	7,120	8	17	7.0	16.2	1.12
Bellevue	32,855	6,715	19	20	7.3	10.5	1.18

Note: The first column indicates the number of admissions in the pre-period data, the second column the number of admissions in the post-period data, the third column the number of zip codes in the service area, the fourth column the number of choices (including the outside option), the fifth column the share of admissions in the pre-period from the 90% service area that went to the outside option, the sixth column the share of admissions in the pre-period from the 90% service area that went to the destroyed hospital, and the seventh column the average DRG weight of admissions to the destroyed hospital in the pre-period data.

Table XVI Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, By Model, Non-Elective Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.58	100
	(-0.62,-0.54)	(100,100)
CDS	-0.58	99.3
	(-0.62,-0.54)	(96.8,101.6)
Char	-0.61	104.8
	(-0.65,-0.57)	(101.9,107.8)
Time	-0.43	75.1
	(-0.49,-0.38)	(70.8,79.8)
Ho	-0.42	72.2
	(-0.48,-0.36)	(67,77.3)
GNT	-0.44	76.6
	(-0.5,-0.39)	(72.8,80.2)
Inter	-0.42	72.2
	(-0.48,-0.36)	(67.4,77.2)
Semipar	-0.43	73.8
	(-0.49,-0.38)	(69.6,79.3)

Note: All results use data from all of the experiments. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model. The lower and upper bounds of the 95% confidence intervals, which are computed from 200 bootstrap replications, are shown below the average slope coefficients; we also apply a bootstrap bias correction.

Table XVII Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error for *Semipar*, by Experiment, Non-Elective Sample

Experiment	Slope
Sumter	-0.11
	(-0.19,-0.03)
StJohns	-0.56
	(-0.65,-0.47)
NYU	-0.6
	(-0.72,-0.51)
Moore	-0.67
	(-0.78,-0.59)
Coney	-0.6
	(-0.65,-0.55)
Bellevue	0.07
	(-0.05,0.19)

Note: All results use data from all of the experiments for the *Semipar* model. The second column presents the slope of the observed diversion ratio on the prediction error. The lower and upper bounds of the 95% confidence intervals, which are computed from 200 bootstrap replications, are shown below the average slope coefficients; we also apply a bootstrap bias correction.

Table XVIII Decomposition of Average Predicted Diversion, Non-Elective Sample

Model	Individual	Individual + Heterogeneity Factor
AggShare	-0.58	-0.58
	(-0.62,-0.54)	(-0.62,-0.54)
CDS	-0.67	-0.58
	(-0.7,-0.64)	(-0.62,-0.54)
Char	-0.69	-0.61
	(-0.73,-0.66)	(-0.65,-0.57)
Time	-0.52	-0.43
	(-0.57,-0.47)	(-0.49,-0.38)
Ho	-0.51	-0.42
	(-0.57,-0.47)	(-0.48,-0.36)
GNT	-0.53	-0.44
	(-0.58,-0.48)	(-0.5,-0.39)
Inter	-0.51	-0.42
	(-0.56,-0.47)	(-0.48,-0.36)
Semipar	-0.51	-0.43
	(-0.56,-0.47)	(-0.49,-0.38)

Note: The second column depicts the slope of the observed diversion ratio on the prediction error based on the predicted individual diversion ratio and the third column depicts the slope of the observed diversion ratio on the prediction error based on the predicted individual diversion ratio plus the heterogeneity factor. Each term is as defined in the text. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XIX Average Prediction Error by Average Hospital Distance, Non-Elective Sample

Model	Below Avg	Above Avg	Above Avg or Outside Option
AggShare	-1.79	0.49	1.28
	(-2.4,-1.25)	(0.04,0.92)	(0.9,1.71)
CDS	-1.04	-0.1	0.75
	(-1.64,-0.48)	(-0.57,0.35)	(0.35,1.17)
Char	-1.34	0.13	0.96
	(-1.94,-0.8)	(-0.34,0.56)	(0.58,1.38)
Time	-1.05	-0.07	0.75
	(-1.66,-0.51)	(-0.52,0.36)	(0.37,1.19)
Ho	-1.01	-0.16	0.73
	(-1.64,-0.43)	(-0.66,0.28)	(0.32,1.17)
GNT	-1.16	-0.02	0.83
	(-1.75,-0.61)	(-0.47,0.42)	(0.44,1.25)
Inter	-1.2	-0.04	0.86
	(-1.8,-0.65)	(-0.52,0.42)	(0.47,1.29)
Semipar	-1.2	0.04	0.86
	(-1.82,-0.63)	(-0.43,0.49)	(0.46,1.3)

Note: We report estimates of average prediction error, examining separately hospitals whose travel time is below average for their market, above average for their market, or above average plus the Outside Option. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XX Predictive Accuracy – Averaged over all Experiments, Non-Elective Sample

Model	Percent Correct	RMSE
AggShare	25.28	0.2298
	(24.96,25.59)	(0.2295,0.23)
CDS	37.88	0.2179
	(37.55,38.25)	(0.2175,0.2184)
Char	34.14	0.2223
	(33.73,34.54)	(0.2219,0.2227)
Time	35.79	0.219
	(35.44,36.11)	(0.2186,0.2193)
Ho	38.5	0.2159
	(37.2,43.17)	(0.2112,0.2172)
GNT	37.55	0.218
	(37.21,37.94)	(0.2177,0.2184)
Inter	40.08	0.2142
	(39.7,41.09)	(0.2124,0.2147)
Semipar	41.62	0.2126
	(41.18,42.09)	(0.2121,0.2132)

Note: The second column is the average percent of individual predictions predicted correctly for each model; we define a prediction as correctly predicting if the patient goes to the hospital with the highest predicted probability, and average across experiments. The third column is the average RMSE for individual predictions, averaged across the different experiments. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXI Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Model, For Areas With Less Disaster Damage, Non-Elective Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.51	100
	(-0.58,-0.45)	(100,100)
CDS	-0.63	123
	(-0.68,-0.58)	(115.2,128.9)
Char	-0.72	140.2
	(-0.77,-0.67)	(129.1,149.1)
Time	-0.47	91.1
	(-0.54,-0.4)	(88.3,94.4)
Ho	-0.46	90.1
	(-0.54,-0.39)	(85.1,95.3)
GNT	-0.47	92.7
	(-0.55,-0.41)	(90.1,95.6)
Inter	-0.48	93.3
	(-0.55,-0.41)	(88.2,99.1)
Semipar	-0.46	89.7
	(-0.53,-0.38)	(81.7,97.2)

Note: Results based upon the Coney, St John's, and Sumter experiments for areas with less disaster damage, as defined in the text. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model; this value is thus 100 for *AggShare*. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXII Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Model, For Coney, St John's, and Sumter Experiments, Non-Elective Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.55	100
	(-0.58,-0.51)	(100,100)
CDS	-0.55	100
	(-0.58,-0.51)	(96.1,103.9)
Char	-0.6	109.9
	(-0.64,-0.56)	(104.4,114.5)
Time	-0.36	66.6
	(-0.4,-0.32)	(61.5,71.6)
Ho	-0.34	62.8
	(-0.39,-0.29)	(56,68.5)
GNT	-0.39	71.1
	(-0.43,-0.34)	(66.5,76.2)
Inter	-0.36	66.5
	(-0.41,-0.31)	(58.7,72.2)
Semipar	-0.37	67.5
	(-0.41,-0.31)	(59.9,74.6)

Note: Results based upon the Coney, St John's, and Sumter experiments using all zip codes in the relevant service area. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXIII Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Random Coefficient Model, Non-Elective Sample

Model	Slope	Slope Rel to Zero SD
Zero SD	-0.43	100
	(-0.49,-0.38)	(100,100)
LOO Common SD	-0.35	80.2
	(-0.41,-0.28)	(75.6,85.4)
Common SD	-0.32	74
	(-0.37,-0.27)	(67.4,79.6)

Note: The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the Zero SD model. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

5.2 Full Sample

Table XXIV Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, By Model, By Model, Full Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.51	100
	(-0.55,-0.47)	(100,100)
CDS	-0.5	97.5
	(-0.53,-0.46)	(95.3,99.6)
Char	-0.54	105.8
	(-0.58,-0.51)	(102.9,108.4)
Time	-0.35	68.2
	(-0.4,-0.3)	(63.9,72.9)
Ho	-0.34	65.7
	(-0.39,-0.28)	(60.3,70.7)
GNT	-0.37	73
	(-0.42,-0.33)	(68.9,76.9)
Inter	-0.34	67.2
	(-0.39,-0.3)	(62.9,71.9)
Semipar	-0.35	67.9
	(-0.39,-0.3)	(63,73.3)

Note: All results use data from all of the experiments. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model. The lower and upper bounds of the 95% confidence intervals, which are computed from 200 bootstrap replications, are shown below the average slope coefficients; we also apply a bootstrap bias correction.

Table XXV Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error for *Semipar*, by Experiment, Full Sample

Experiment	Correlation
Sumter	-0.21
	(-0.27,-0.14)
StJohns	-0.44
	(-0.52,-0.37)
NYU	-0.26
	(-0.36,-0.16)
Moore	-0.63
	(-0.74,-0.54)
Coney	-0.5
	(-0.56,-0.44)
Bellevue	0
	(-0.1,0.11)

Note: All results use data from all of the experiments for the *Semipar* model. The second column presents the slope of the observed diversion ratio on the prediction error. The lower and upper bounds of the 95% confidence intervals, which are computed from 200 bootstrap replications, are shown below the average slope coefficients; we also apply a bootstrap bias correction.

Table XXVI Decomposition of Average Predicted Diversion, Full Sample

Model	Individual	Individual + Heterogeneity Factor
AggShare	-0.51	-0.51
	(-0.55,-0.47)	(-0.55,-0.47)
CDS	-0.61	-0.5
	(-0.64,-0.58)	(-0.53,-0.46)
Char	-0.64	-0.54
	(-0.67,-0.61)	(-0.58,-0.51)
Time	-0.45	-0.35
	(-0.49,-0.41)	(-0.4,-0.3)
Ho	-0.45	-0.34
	(-0.49,-0.4)	(-0.39,-0.28)
GNT	-0.46	-0.37
	(-0.5,-0.42)	(-0.42,-0.33)
Inter	-0.44	-0.34
	(-0.49,-0.4)	(-0.39,-0.3)
Semipar	-0.44	-0.35
	(-0.48,-0.4)	(-0.39,-0.3)

Note: The second column depicts the slope of the observed diversion ratio on the prediction error based on the predicted individual diversion ratio and the third column depicts the slope of the observed diversion ratio on the prediction error based on the predicted individual diversion ratio plus the heterogeneity factor. Each term is as defined in the text. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXVII Average Prediction Error by Average Hospital Distance, Full Sample

Model	Below Avg	Above Avg	Above Avg or Outside Option
AggShare	-2.29	1.03	1.64
	(-2.8,-1.82)	(0.64,1.45)	(1.3,2)
CDS	-1.42	0.35	1.02
	(-1.95,-0.95)	(-0.01,0.77)	(0.69,1.39)
Char	-1.73	0.57	1.24
	(-2.24,-1.27)	(0.2,1)	(0.91,1.6)
Time	-1.53	0.45	1.09
	(-2.03,-1.06)	(0.06,0.88)	(0.77,1.45)
Ho	-1.48	0.35	1.06
	(-2,-1.02)	(-0.04,0.79)	(0.74,1.43)
GNT	-1.66	0.51	1.19
	(-2.17,-1.21)	(0.11,0.93)	(0.87,1.55)
Inter	-1.64	0.45	1.17
	(-2.17,-1.16)	(0.04,0.86)	(0.83,1.55)
Semipar	-1.72	0.53	1.23
	(-2.23,-1.24)	(0.13,0.95)	(0.89,1.59)

Note: We report estimates of average prediction error, examining separately hospitals whose travel time is below average for their market, above average for their market, or above average plus the Outside Option. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXVIII Predictive Accuracy – Averaged over all Experiments, Full Sample

Model	Percent Correct	RMSE
AggShare	30.25	0.236
	(28.16,33.25)	(0.2358,0.2363)
CDS	36.73	0.2259
	(36.18,37.24)	(0.2254,0.2264)
Char	34.06	0.2302
	(33.54,34.58)	(0.2298,0.2307)
Time	35.86	0.2256
	(35.38,36.33)	(0.2251,0.226)
Ho	36.54	0.2242
	(36.02,37.08)	(0.2236,0.2247)
GNT	37	0.2254
	(36.56,37.5)	(0.2249,0.226)
Inter	39.6	0.2211
	(39.14,40.04)	(0.2206,0.2216)
Semipar	42.26	0.2188
	(41.76,42.83)	(0.2179,0.2196)

Note: The second column is the average percent of individual predictions predicted correctly for each model; we define a prediction as correctly predicting if the patient goes to the hospital with the highest predicted probability, and average across experiments. The third column is the average RMSE for individual predictions, averaged across the different experiments. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXIX Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Model, For Areas With Less Disaster Damage, Full Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.47	100
	(-0.54,-0.41)	(100,100)
CDS	-0.61	129.5
	(-0.66,-0.56)	(120.9,134.9)
Char	-0.7	148
	(-0.74,-0.65)	(136.1,157.6)
Time	-0.43	92
	(-0.5,-0.37)	(90.1,94.5)
Ho	-0.43	90.9
	(-0.5,-0.36)	(85.5,94.9)
GNT	-0.44	94.3
	(-0.51,-0.38)	(92.2,96.8)
Inter	-0.44	93.3
	(-0.5,-0.38)	(89.3,98.1)
Semipar	-0.41	87.5
	(-0.5,-0.34)	(79.3,95.7)

Note: Results based upon the Coney, St John's, and Sumter experiments for areas with less disaster damage, as defined in the text. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model; this value is thus 100 for *AggShare*. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXX Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Model, For Coney, St John's, and Sumter Experiments, Full Sample

Model	Slope	Slope Rel to AggShare
AggShare	-0.53	100
	(-0.56,-0.5)	(100,100)
CDS	-0.53	99.9
	(-0.55,-0.5)	(97,102.9)
Char	-0.59	111.6
	(-0.62,-0.56)	(107.8,115.5)
Time	-0.34	64.9
	(-0.39,-0.31)	(60.8,70)
Ho	-0.33	62
	(-0.38,-0.27)	(53.4,67.8)
GNT	-0.37	70.6
	(-0.41,-0.34)	(66.4,75.1)
Inter	-0.34	63.9
	(-0.38,-0.3)	(58.8,69)
Semipar	-0.33	62.9
	(-0.38,-0.29)	(57,69.4)

Note: Results based upon the Coney, St John's, and Sumter experiments using all zip codes in the relevant service area. The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the *AggShare* model. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

Table XXXI Slope Coefficient of Observed Choice Removal Diversion Ratios on Prediction Error, by Random Coefficient Model, Full Sample

Model	Slope	Slope Rel to Zero SD
Zero SD	-0.35	100
	(-0.4,-0.3)	(100,100)
LOO Common SD	-0.23	66.8
	(-0.29,-0.17)	(61,73)
Common SD	-0.2	58.8
	(-0.25,-0.16)	(51.2,65.7)

Note: The second column presents the slope of the observed diversion ratio on the prediction error. The third column depicts the slope for a model relative to that for the Zero SD model. 95% confidence intervals are computed from 200 bootstrap replications; we also apply a bootstrap bias correction.

6 Table of Variables Used

	AggShare	Char	CDS	Time	Ho	GNT	Inter
Hospital Indicators	X			X	X	X	X
× Weight						X	X
× Time							X
× Obstetrics							X
× Circulatory							X
× Digest							X
× Muscular							X
× Respiratory							X
× Female Repro							X
Inside		X	X				
× Cardiac Surg Diag			X				
Same County			X				
Time		X	X	X	X	X	X

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Median Income			X			X	
× LOS			X				
× nPX			X				
× nDX			X				
× Emergency			X		X		
× Medical			X				
× Obstetrics			X				X
× Weight		X	X			X	X
× Age						X	
× Under18			X				X
× Over64		X	X			X	X
× Female		X	X			X	X
× Black							X
× Cardiac Surg Diag			X				

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Circulatory			X				X
× Digest			X				X
× Muscular			X				X
× Respiratory			X				X
× Female Repro			X				X
× RN Share			X				
× Teach			X				
× RN Intense			X				
× For Profit						X	
× Beds						X	
× Residents Per Bed						X	
× Teach						X	
Squared Time		X	X	X	X	X	X
× Weight		X					X

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Over64		X					X
× Under18							X
× Female		X					X
× Black							X
× Obstetrics							X
× Circulatory							X
× Digest							X
× Muscular							X
× Respiratory							X
× Female Repro							X
Closest						X	
Cardiac Surg Hosp							
× Cardiac Surg Diag × Adult		X					

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Weight × Adult		X					
Obstetrics Hosp							
× Obstetrics Diag		X					
× Female		X					
NICU Hosp							
× Female		X					
× Obstetrics Diag						X	
Residents/Bed							
× Weight		X					
× Over64		X					
× Female		X					
RN Share							
× Female			X				
× Over64			X				

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Median Income			X				
× LOS			X				
× nPX			X				
× nDX			X				
× Under18			X				
RN Int							
× Female			X				
× Over64			X				
× Median Income			X				
× LOS			X				
× nPX			X				
× nDX			X				
× Under18			X				
RN/Bed							

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Commercial					X		
× Cardiac					X		
× Oncology Alt					X		
× Neurology					X		
× Digest Alt					X		
× Labor and Delivery					X		
× Median Income					X		
For Profit							
× Weight		X					
× Over64		X					
× Female		X					
× Commercial					X		
× Cardiac					X		
× Oncology Alt					X		

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Neurology					X		
× Digest Alt					X		
× Labor and Delivery					X		
× Median Income					X		
Imaging Complexity							
× Commercial					X		
× Cardiac					X		
× Oncology Alt					X		
× Neurology					X		
× Digest Alt					X		
× Labor and Delivery					X		
× Median Income					X		
Teach							
× Female			X				

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Old			X				
× LOS			X				
× nPX			X				
× nDX			X				
× Under18			X				
× Commercial					X		
× Cardiac					X		
× Oncology Alt					X		
× Neurology					X		
× Digest Alt					X		
× Labor and Delivery					X		
× Median Income			X		X		
Cardiac Complexity							
× Commercial					X		

	AggShare	Char	CDS	Time	Ho	GNT	Inter
× Cardiac					X		
× Median Income					X		
× Commercial					X		
× Oncology Alt					X		
× Median Income					X		
Birth Complexity							
× Commercial					X		
× Labor and Delivery					X		
× Median Income					X		
Oncology Diag × Cancer Center			X				
Delivery × Birth Room			X				
Circulatory × Cardiac ICU			X				
Circulatory × Cath Lab						X	

	AggShare	Char	CDS	Time	Ho	GNT	Inter
Under18 × Ped Beds			X				
Trauma × CTC			X				
Imaging Diag × MRI						X	

Variable	Source	Description
Adult	Disch	Age greater than 17
Age	Disch	Patient age
Beds	AHA	Number of beds in hospital
Black	Disch	Patient is racially black
Birth Room	AHA	Whether hospital has birthing (LDR, LDRP) room
Birth Complexity	AHA	We apply the Ho services intensity algorithm (see below) to the obstetrics and birth room flags from the AHA data
Cardiac Surg Diag	Disch	For V24 DRG coding: DRGs between 215 and 236; Between V24 and V12: DRGs in this list (104,105,106,108,515,525,535,536,547,548,549,550); Below V12: DRGs between 103 and 107, 207[3]
Cardiac Surg Hosp	AHA	Whether hospital has an cardiac surgery program
Cardiac Diagnosis	Disch	3 digit ICD9 diagnosis codes between (and including) 393 and 398, 401 and 405, 410-417, 420-429

[3]Across the models, three different underlying variables are based on the patient's diagnosis. First, the discharge data include ICD9 diagnosis codes for patients; these diagnosis codes, along with other variables such as procedures, age, sex, discharge status, and the presence of complications or comorbidities, are used to assign a Diagnosis Research Group or DRG. The DRGs themselves are grouped into 25 different Major Diagnosis Categories or MDCs. For example, a patient presenting signs of "maple syrup urine disease" would have ICD9 diagnosis code 270.3, DRG 642 (Inborn and other disorders of metabolism), and MDC 10 (Diseases and Disorders of the Endocrine, Nutritional And Metabolic System).

Cardiac ICU	AHA	Whether hospital has cardiac ICU
Cardiac Complexity	AHA	We apply the Ho services intensity algorithm (see below) to adult diagnostic catheterization, cardiac intensive care, adult interventional cardiac catheterization, and adult cardiac surgery flags from the AHA data
Cath Lab	AHA	Whether a hospital has both a diagnostic and interventional catheterization lab
Cancer Center	AHA	Whether hospital has oncology services
Cancer Complexity	AHA	We apply the Ho services intensity algorithm (see below) to the cancer and maximum of the image-guided radiation and intensity-modulated radiation flags from the AHA data
Closest	Disch	Whether hospital is closest facility to patient
Circulatory	Disch	MDC equals 5
Commercial	Disch	Patient has a commercial insurer
CTC	AHA	Certified Trauma Center
Delivery	Disch	For DRG coding above V24: DRGs in this list (765, 766, 774, 775, 767, 768, 776, 769, 777, 780, 781, or 782). For DRG coding below V12: DRGs from 370-378 and 382-384
Digest	Disch	MDC equals 6
Digest Alt	Disch	3 digit ICD9 diagnosis codes between (inclusive) 520 and 579
Emergency	Disch	Patient admitted through emergency room
Female	Disch	Patient is female
Female Repro	Disch	MDC equals 13
For Profit	AHA	Whether hospital is a for profit facility

Imaging Complexity	AHA	We apply the Ho services intensity algorithm (see below) to SPECT, MRI, CT, ultrasound, and PET scan flags from the AHA survey
Imaging Diag	Disch	MDC code is 1, 5, or 8
Labor and Delivery	Disch	ICD9 diagnosis codes between (inclusive) 650 and 657, 644, 647, 648, V22, V23, V24, V27
Inside	NA	Hospital is not the outside option
Median Income	ACS	Median income of zip code
Medical	Disch	Medical DRG
MRI	AHA	Hospital has an MRI
Muscular	Disch	MDC equals 8
nDX	Disch	Number of diagnoses
nPX	Disch	Number of procedures
LOS	Disch	Length of stay
Neurology	Disch	3 digit ICD9 diagnosis codes between 320 and 326, 330 and 337, or 340 and 359 (inclusive)
NICU Diag	Disch	For V24 DRG Coding: DRG 790 or 791; Pre V24: DRG 386 or 387
NICU Hosp	AHA	Hospital has a NICU
Obstetrics Diag	Disch	MDC equals 14
Obstetrics Hosp	AHA	Hospital has an obstetrics program

Oncology Diag	Disch	MDC equals 17 or for DRG later than V24 in this list (54, 55, 146, 147, 148, 180, 181, 182, 374, 375, 376, 420, 421, 422, 435, 436, 437, 542, 543, 544, 582, 583, 584, 585, 597, 598, 599, 656, 657, 658, 686, 687, 688, 711, 712, 715, 716, 722, 723, 724, 739, 740, 741, 736, 737, 738, 744, 745, 754, 755, 756, 843, 844, 835, 836, 837, 838, or 839). For DRG pre V12 in this list (10, 11, 64, 82, 172, 173, 199, 203, 239, 257, 258, 259, 260, 274, 275, 303, 318, 319, 338, 344, 346, 347, 354, 355, 357, 363, 366, 367, 406, 407, 408, 413, 414).
Oncology Alt	Disch	3 digit ICD9 diagnosis codes between 140 and 239 (inclusive)
Over64	Disch	Patient is over 64 years old
PatCounty	Disch	Patient's County of Residence
Ped Beds	Disch	Hospital has pediatric beds
Respiratory	Disch	MDC equals 4
Residents Per Bed	MCR	Residents per bed from Medicare Cost Reports
RN Share	AHA	Nurses regularly working as a share of licensed nurses
RN Intense	AHA	Nurses regularly working as share of inpatient days
RN/Bed	AHA	Nurses per bed
Same County	Disch	Hospital and patient in same county
Teach	AHA	Teaching hospital
Trauma	Disch	MDC equals 24
Time	Disch/Compute	Travel time from centroid of patients zip code to hospital
Under18	Disch	Patient is under 18
Weight	Disch	DRG weight

Description of Ho Services Intensity Algorithm: Hospitals were rated on a scale of zero to one, reflecting the sophistication of their services in different categories. Zero indicates low sophistication and one indicates a high level of sophistication. The four categories are cardiac, imaging, cancer, births.

The intensity variable for category c in hospital h is given by:

$$\max\{\max_{x \in X_c} 1_{xh} * (1 - \bar{x})(1 - \bar{y}_c), 1_{yh_c}\}$$

where

- x indexes the services in each category
- 1_{xh} is 1 if hospital h offers service x and 0 if not
- $\bar{x}$ is the state share of hospitals offering that service
- y is the service with the smallest $\bar{x}$
- 1_{yh} is 1 if hospital h offers service y and 0 if not
- $\bar{y}$ is the percent of hospitals offering service y

For more details see Table IX in Ho (2006).

References

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

Ho, K. "The Welfare Effects of Restricted Hospital Choice in the US Medical Care Market." *Journal of Applied Econometrics*, vol. 21 (2006), pp. 1039–1079.

Raval, D., Rosenbaum, T., and Tenn, S.A. "A Semiparametric Discrete Choice Model: An Application to Hospital Mergers." *Economic Inquiry*, vol. 55 (2017), pp. 1919–1944.

Raval, D., Rosenbaum, T., and Wilson, N.E. "How do machine learning algorithms perform in predicting hospital choices? evidence from changing environments." *Journal of Health Economics*, vol. 78 (2021), p. 102481. URL https://www.sciencedirect.com/science/article/pii/S0167629621000667.