Free to Choose? Reform and Demand Response in the British National Health Service

Martin Gaynor* Carol Propper◇ Stephan Seiler⋆

*Carnegie Mellon University, University of Bristol and NBER

◇Imperial College, University of Bristol and CEPR

⋆Stanford University and CEPR

London
May 2013

Big Question

- What is impact of competition on quality in health care markets?
- US relies on marketsand reforms outside the US seek to expand choice using markets.
  - Big deal - quality of care can have a big impact on welfare.
  - We know something about the impact of market structure on quality (e.g., Kessler and McClellan, 2000; Cutler, Huckman, Kolstad, 2010; Cooper, Gibbons, Jones, McGuire, 2010; Gaynor, Moreno- Serra, Propper, 2010).
- But don't know much about the mechanisms by which this occurs.

Big Question cont

- In economic models of competition it is standard that firms' demands become more elastic with the # of firms, i.e. with more consumer choice.
- A necessary condition for competition to get tougher with the # of firms.
- Many policies rest on assumption that patient choice will provide benefits by forcing providers to compete more for patients and so respond to consumer preferences.
- However little direct testing of this.

Our Question

- Directly examine whether demand becomes more elastic in response to increased choice.
- Exploit a policy change in English NHS that mandated expanded choices for consumers.
  - Prices are regulated so choice is driven by non-price attributes, in particular, quality of care.
- Not obvious that increased choice will make demand more responsive to quality.
  - Switching costs.
  - Non-rational decisions (e.g. status quo bias).
  - Lack of information.
  - Poor information.
- Previous evidence on demand responsiveness to quality.
  - Luft, Garnick, Mark, Peltzman, Phibbs, Lichtenberg, McPhee, (1990), Tay (2003), Howard (2005), Sivey (2008).

What We Do

- Use individual data on patient choice of hospital for CABG surgery in the English NHS before and after the reform to estimate responsiveness of choice to quality (mortality, waiting time).
  - Hospital discharge data from the NHS.
  - Multinomial logit model with heterogeneity.
  - Risk adjust mortality; instrument for waiting time.
  - No supply side.

We Find

- Patients value lower mortality rates, lower waiting times, less distance.
- Allowing for heterogeneity
  - No difference in effect of mortality rates on demand pre- and post-reform for the average patient, but sicker patients are more responsive post-reform (no difference for poorer patients).
  - The effect of waiting time on demand is smaller post reform. Sicker patients are less responsive to waiting times post-reform, poorer patients are more responsive.
- Hospitals are spatially differentiated. Distance negatively affects demand and cross-elasticities with respect to mortality rates fall with distance between hospitals.

The Reform

- Reforms in English NHS in 2006 designed to promote patient choice and competition among hospitals.
  - "Choose and Book": Mandated that patients offered choice of 5 hospitals.
    - Pre-reform patients referred to hospital with which their local HA had a contract.
    - 30% of patients aware of choice in 2006; 50% in 2009.
  - "PbR": Hospitals paid fixed, regulated prices (analogous to PPS).
    - Pre-reform local health authorities negotiated contracts with hospitals (price, volume, waiting time).
  - Hospitals classified as Foundation Trusts allowed to keep surpluses. Hospitals designated FTs on performance (financial, waiting times).

The Reform summary

- Patients provided with more choices, both via Choose and Book and end of selective contracting.
- Hospitals moved from negotiated price environment to regulated price environment.
- Hospitals have incentives for financial performance.

CABG Surgery

- Coronary artery bypass graft surgery.
  - Treatment for coronary artery disease
- Frequently performed elective treatment ($\sim$13K patients per year).
- Risky, so mortality rate good, observable measure of quality of care.
- Referral
  - Patients in NHS with coronary artery disease or angina are referred to a cardiologist.
  - Referral for CABG surgery typically made by cardiologist (possibly by GP).
- Market
  - 28, 29 hospitals offer CABG (out of $\sim$160).
  - Market is national (England).
  - Patients pay nothing.

Data

- Inpatient discharge data set for all NHS hospitals (Hospital Episode Statistics, HES).
  - All CABG patients treated in NHS hospitals (vs. private providers).
  - Diagnoses, procedures, patient post code, hospital location, patient characteristics (age, sex, co-morbidities), time between referral and treatment (waiting time), patient mortality within 30 days of treatment, elective treatment.
  - Merge area characteristics.
- Select elective CABGs (75% of total).
- Pre-reform is fiscal years 2003/04, 2004/05.
- Post-reform is fiscal years 2006/07, 2007/08 (fiscal years begin in April).

Descriptive Statistics – Patients

	Mean	Median	Standard Deviation	10th Percentile	90th Percentile
Age	65.76	66	55.04	53	76
Index of Multiple Deprivation	0.14	0.1	0.11	0.04	0.31
Comorbidity Count	5.42	5	2.81	2	9
Capped Comorbidity Count (Cap = 6)	4.57	5	1.61	2	6
Probability of Informedness About Choice	0.53	0.53	0.07	0.45	0.63
Fraction Male	81.18%				

Descriptive Statistics – Hospitals

	Total Admissions CABGs		Waiting Times (Days)	
	Mean	Std	Mean	Std
2003	502.9	189.4	109.1	32.1
2004	507.5	200.0	100.5	20.7
2005	449.1	170.8	67.8	15.2
2006	425.4	172.7	65.6	17.3
2007	459.9	169.9	64.9	21.4

Descriptes Statistics – Hospitals Continued

	Mortality Rate CABGs		Adjusted Mortality Rate, CABGs	
	Mean	Std	Mean	Std
2003	1.88	0.82	1.67	1.39
2004	1.93	0.78	1.46	1.45
2005	1.90	0.57	1.19	1.14
2006	1.95	0.79	1.40	1.18
2007	1.51	0.69	0.73	0.90

Descriptive Statistics – Distance

		Mean	Median	Standard Deviation	10%	90%	95%
Distance	Pre	34.93	22.34	44.97	4.77	71.40	98.15
	Post	32.24	22.91	32.94	4.93	70.58	92.36

Fraction of Patients Visiting Closest Hospital	Pre	68.14
	Post	68.67

Patterns in the Data

- Are patterns consistent with reforms having an impact?
- Do better hospitals have higher market shares after the introduction of choice?
  - Regress case-mix adjusted hospital mortality rates and hospital fixed effects on hospital market shares pre- and post-reform, for elective and for emergency CABG (placebo test).
  - Mortality rates are negatively related to market share post-reform for elective CABG. No significant relationship pre-reform for elective CABG or at all for emergency CABG.

Mortality Rates and Market Shares – Elective CABGs

	(1)	(2)	(3)	(4)
Dependent variable	Elective CABGs		Market-share	
Time period	Pre	Post	Pre	Post
Coefficient on Case-Mix Adjusted Mortality Rate	0.0042	-0.1652**	0.0416	-0.0692*
	(0.0552)	(0.0622)	(0.0267)	(0.0321)
Hospital Fixed Effects	No	No	Yes	Yes
Number of Observations	142	143	142	143
Hospitals	29	29	29	29
Quarters	5	5	5	5

Mortality Rates and Market Shares – Emergency CABGs

	(5)	(6)	(7)	(8)
Dependent variable	Emergency CABGs		Market-share	
Time period	Pre	Post	Pre	Post
Coefficient on Case-Mix Adjusted Mortality Rate	0.0488	-0.0667	-0.0373	-0.0127
	(0.0531)	(0.0744)	(0.0379)	(0.0466)
Hospital Fixed Effects	No	No	Yes	Yes
Number of Observations	142	143	142	143
Hospitals	29	29	29	29
Quarters	5	5	5	5

Patterns in the Data 2

- Did the mortality rate drop after the introduction of choice?
  - Yes
- Did the mortality rate drop more for patients who bypass the nearest hospital?
  - Yes – consistent with patients exercising more choice.

Changes in the Expected Mortality Rate

Sample	Mean Pre	Mean Post	Difference
Mortality Rate – Raw Rate			
All Patients	1.344	0.948	-0.396
Patients Visiting Neareast Hospital	1.287	1.022	-0.265
Patients Not Visiting Nearest Hosp.	1.462	0.779	-0.683
Mortality Rate – Case-Mix Adjusted			
All Patients	1.471	0.748	-0.723
Patients Visiting Nearest Hospital	1.352	0.809	-0.543
Patients Not Visiting Nearest Hosp.	1.716	0.606	-1.110

Modelling Approach

- 2 parts to econometric model
  - Hospital choice as a function of hospital characteristics/dimensions of service
  - Hospital production function for quality of clinical care
- Focus is demand model – pin down how sensitive hospital choice decisions are to quality of care
  - Need to find appropriate measure for hospital quality of care – production model helps with that
- Agency
  - Physician plays a major role in hospital choice – can't separately identify patient and physician preferences
  - Not an issue for identifying impact of reform on hospital choice
  - Physicians in NHS salaried – no financial incentive wrt referrals

Choice Model

- Patient i chooses hospital j to visit in time period t which provides the maximum utility according to:

$$V_{ijt} = \beta_{w,it} W_{jt} + \beta_{z,it} Z_{jt} + f(D_{ij}) + \xi_{jt} + \varepsilon_{ijt}$$

- where:
  - W is waiting time, Z is the hospital's mortality rate, D is the distance from patient i to hospital j.
  - ξ_{jt} is unobserved hospital quality, ε_{ijt} is a random iid shock distributed Type I Extreme Value.
  - We allow for heterogeneity across patients in preferences for observables and unobservables.

Choice Model, cont'd.

- Allow impacts of quality of care and waiting times to differ across patients and before and after the reform

$$\begin{aligned}\beta_{z,it} &= [\overline{\beta}_{z,0} + \beta_{z,0} X_i + \sigma_{z,0} v_{z,i}] \cdot \mathbf{1}(t=0) \\ &\quad + [\overline{\beta}_{z,1} + \beta_{z,1} X_{it} + \sigma_{z,1} v_{z,i}] \cdot \mathbf{1}(t=1).\end{aligned}$$

$$\begin{aligned}\beta_{w,it} &= [\overline{\beta}_{w,0} + \beta_{w,0} X_i + \sigma_{w,0} v_{w,i}] \cdot \mathbf{1}(t=0) \\ &\quad + [\overline{\beta}_{w,1} + \beta_{w,1} X_i + \sigma_{w,1} v_{w,i}] \cdot \mathbf{1}(t=1)\end{aligned}$$

- where $(t=0)$ is pre-reform time period, $(t=1)$ is post-reform, $\overline{\beta}_{z,t}$ is average effect across consumers in period t, X_i is observable patient demographics, $\beta_{z,t}$ captures differences from the average effect across consumers, $\sigma_{z,t}$ captures unobserved heterogeneity

Choice Model, cont'd.

- Utility function is thus:

$$\begin{aligned} u_{ijt} &= \delta_{jt} \\ &+[\beta_{w,0}X_i \cdot \mathbf{1}(t=0) + \beta_{w,1}X_i \cdot \mathbf{1}(t=1)] \cdot W_{jt} \\ &+[\sigma_{w,0}v_{w,i} \cdot \mathbf{1}(t=0) + \sigma_{w,1}v_{w,i} \cdot \mathbf{1}(t=1)] \cdot W_{jt} \\ &+[\beta_{z,0}X_i \cdot \mathbf{1}(t=0) + \beta_{z,1}X_i \cdot \mathbf{1}(t=1)] \cdot Z_{jt} \\ &+[\sigma_{z,0}v_{z,i} \cdot \mathbf{1}(t=0) + \sigma_{z,1}v_{z,i} \cdot \mathbf{1}(t=1)] \cdot Z_{jt} \\ &+f(D_{ij}) + \varepsilon_{ijt} \end{aligned}$$

- where δ captures average effects and unobserved quality

$$\begin{aligned} \delta_{jt} &= [\overline{\beta}_{w,0} \cdot \mathbf{1}(t=0) + \overline{\beta}_{w,1} \cdot \mathbf{1}(t=1)] \cdot W_{jt} \\ &+[\overline{\beta}_{z,0} \cdot \mathbf{1}(t=0) + \overline{\beta}_{z,1} \cdot \mathbf{1}(t=1)] \cdot Z_{jt} \\ &+\xi_{jt} \end{aligned}$$

Correlation

- May be correlations between waiting times or mortality rates and unobserved hospital quality.
  - Unobservably better hospitals may attract more patients and thus have longer waiting times.

$$Cov(w_{jt}, \xi_{jt}) \neq 0$$

  - Unobservably better hospitals may attract sicker patients and thus have higher mortality rates.

$$Cov(z_{jt}, \xi_{jt}) \neq 0$$

- We assume distance (d_{ij}) is uncorrelated with unobserved hospital quality (ξ_j), i.e. patients don't choose where to live based on hospitals' quality of care for CABG surgery.

Waiting Times

- Hospitals with higher unobserved quality will have greater demand, and therefore higher waiting times.
- Could use IV or control for unobserved heterogeneity via fixed effects to absorb the variation in ξ_{jt}.
- No obvious good instrument so we use hospital-quarter fixed effects i.e. time varying fixed effects.
- Some costs to this: have to use two-step approach to identify average effects – almost 300 fixed effects.

Production Function for Clinical Quality of Care

- Use hospital's mortality rate for CABG patients as measure of the quality of care.
- Use the production function to deliver an appropriate measure of quality of care for demand model.
- Linear probability model of mortality:

$$M = JT\psi + H^{obs}\gamma_{obs} + H^{unobs}\gamma_{unobs} + \eta$$

- where M is mortality, JT is a matrix of hospital-time period dummy variables, H^{obs} is a matrix of patient characteristics that capture observable health status. H^{unobs} is unobserved health status.

- Issue - better hospitals attract sicker patients leading to a correlation with the error.
- Very similar to problem of hospital mortality rates being "contaminated" by differences in patient case-mix.
- Implies raw mortality rate not accurate measure – need to control for the patient's health status.
- Use approach of Gowrisankaran and Town (1999): instrument for the hospital dummies using distance.
  - Use distance and dummies for closest hospital: 2 $\times$ as many instruments as hospital dummies.
  - Estimate as linear probability model.
  - Use the IV estimates of $\hat{\psi}_{jt}$ as our risk-adjusted measure of the hospital mortality rate.

First Step Estimates

- Sicker patients are sensitive to hospital mortality rates and more sensitive after choice reform.
- "Low income" patients more sensitive to waiting times after choice reform.
- "Informed" patients more sensitive to waiting times.
- Distance matters.

First Step Estimates

			Coefficient	Stand. Error	
Income	Waiting	Pre	0.01	0.68	
Deprivation	Times	Post	-3.85	0.65	**
Index	Mortality	Pre	0.12	0.86	
	Rate	Post	-0.02	1.13	
Co-	Waiting	Pre	5.67	0.43	**
Morbidity	Times	Post	4.10	0.58	**
Count	Mortality	Pre	-10.11	0.56	**
	Rate	Post	-13.18	0.94	**
Patient	Waiting	Pre	1.25	0.66	*
Informedness	Times	Post	-4.41	0.65	**
	Mortality	Pre	5.17	0.83	**
	Rate	Post	0.01	1.06	
Unobserved	Waiting	Pre	-0.22	75.36	
Preference	Times	Post	-0.26	76.70	
Heterogeneity	Mortality	Pre	35.02	0.87	**
	Rate	Post	39.04	1.81	**

First Step cont'd.

		Coefficient	S. E.	
Distance	Linear	-14.86	0.21	**
	Square	4.91	0.11	**
	Cube	-0.57	0.02	**
	Closest Dummy	1.07	0.02	**
	Closest "Plus 10" Dummy	-0.01	0.00	*
	Closest "Plus 20" Dummy	0.01	0.05	

2nd Step Estimates – Average Effects

- Patients more responsive to mortality rates post-reform. Large impact.
- Waiting times not significant either pre- or post-reform.(Major policy aimed at reducing waiting times before the choice reform.)

2nd Step Estimates – Av Effects (cont)

		Baseline Specification		Sensitivity Check	
Average Effect		Coeff.	S.E.	Coeff.	S.E.
Waiting Times	Pre	-4.24	3.15	0.43	4.09
	Post	6.25	4.74	13.45	7.53
Quality	Pre	-4.85	3.70	-1.63	3.81
	Post	-12.40	4.00**	-11.39	3.96**
Hospital Fixed Effects		Constant Fixed Effects Pre- and Post-Reform		Separate Fixed Effects Pre- and Post-Reform	

Elasticities

- Calculate patient and hospital level elasticities using the parameter estimates.
- Calculate impact of a 1 standard deviation change in mortality rate.
- Patients
  - Large increases in responsiveness, elasticities small.
- Hospitals
  - Large increases in responsiveness.
  - Elasticities small (but larger than patient).

Patient–Level Elasticities of Demand

		Impact on Patient's Purchase Probability From 1 S.D. Shift in Adjusted Mortality	Elasticity
Pre-Reform	Average Patient	-2.69	-0.021
	Lower Income	-2.63	-0.021
	Higher Comorb	-8.30	-0.066
	More Informed	0.18	0.001
Post-Reform	Average Patient	-7.08	-0.056
	Lower Income	-7.09	-0.056
	Higher Comorb	-14.40	-0.114
	More Informed	-7.08	-0.056

Hospital Elasticities

1-S.D. Shift	Mean	S.D.	25th Perc.	Median	75th Perc.
Pre-Reform	-0.36	5.11	-1.73	0.11	0.56
Post-Reform	-4.83	4.73	-5.66	-3.14	-2.34
Change	-5.38	5.81	-6.25	-2.88	-2.28

Elasticities	Mean	S.D.	25th Perc.	Median	75th Perc.
Pre-Reform	0.02	0.16	-0.03	0.00	0.01
Post-Reform	-0.12	0.07	-0.16	-0.10	-0.05
Change	-0.14	0.19	-0.15	-0.07	-0.05

Policy Evaluations

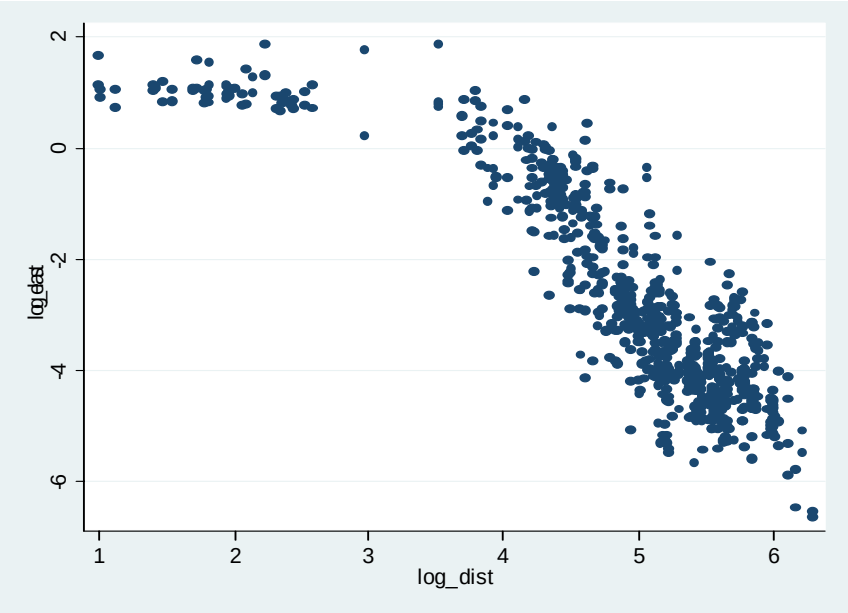
- Impact of choice on patient survival.
  - How many more people would have died absent the reform?
  - If patients in the post-reform period were subject to pre-reform constraints and chose with pre-reform parameters.
  - 12 more people would have died absent the reform.
- Change in patient welfare
  - What is impact on utility of freedom of choice?
  - Compare utility post-reform patients received with utility they would have received if they'd chosen with "restricted" pre-reform choice parameters.
  - Freeing choice led to 7.68% increase in patient welfare.
- Supply Side
  - Regress change in adjusted mortality on change in demand elasticity with respect to quality.
  - Large response of mortality rate to elasticity – consistent with a supply side response.

Supply Side Response

Dependent Variable	Change in Case-Mix Adjusted Mortality Rate
Change in the Elasticity of Demand with Respect to the Mortality Rate	-0.1296*** (0.0209)
Observations	27

Spatial Differentiation

Calculate cross-elasticities with respect to mortality rate for every pair of hospitals (both directions), graphed against distance between the two hospitals.



Hospitals nearby are close substitutes, little competition with far away hospitals.

Summary and Conclusions

- Tested one of basic premises of competitive models – demand elasticity increases with choice.
- Find that it does, and there is substantial heterogeneity in consumer response.
  - More severely ill patients become more sensitive to quality of care post-reform – reform had a stronger impact on patients for whom quality of care is most important.
  - No evidence that lower income patients chose lower quality hospitals post reform though became more sensitive to waiting times.
- Implies that the NHS reforms had an impact on demand, a necessary condition for competition.
- Sheds some light on the results from previous studies.
- Choice and competition can play an important role in health care.

Thank you

Demand with Heterogeneity and Hospital Fixed Effects

Time Period	Pre-Reform	Post-Reform
Waiting Time	-0.022	-0.020
	(0.014)	(0.015)
Mortality Rate (Adjusted)	-0.003	-0.007
	(0.004)	(0.006)
Error From First Stage	0.027*	
(Control Function)	(0.014)	
Waiting Time	0.028***	0.025***
* Comorbidity Count	(0.002)	(0.003)
Mortality Rate (Adjusted)	-0.021***	-0.060***
* Comorbidity Count	(0.002)	(0.004)
Waiting Time	-0.040	-0.336***
* IMD Index	(0.033)	(0.044)
Mortality Rate (Adjusted)	0.035	0.060
* IMD Index	(0.034)	(0.049)
Observations	51,793	

IV Mortality Rate Model

In order to obtain adjusted mortality rates we run a linear probability model, regressing a dummy for death on a set of quarter-specific hospital dummies. The hospital dummies are stacked in a block-diagonal matrix, each block representing one quarter out of 20 quarter for the time period 2001 to 2005. The case-mix is restricted to enter in the same way in all quarters. Specifically, the data are arranged as follows:

$$X = \begin{bmatrix} X_1 & & & & CM_1 \\ & X_2 & & & CM_2 \\ & & \ddots & & \vdots \\ & & & X_{20} & CM_{20} \end{bmatrix}$$

IV Mortality Rate Model Cont'd.

Where CM_t denotes a matrix with various variables capturing the health status of patients within a particular quarter t. All elements in the matrix other then the matrices X_1 to X_{20} and CM_1 to CM_{20} are equal to zero. The block-diagonal elements are given by:

$$X_t = \begin{bmatrix} x^t_{11} & \cdots & x^t_{1k_t} \\ \vdots & \ddots & \vdots \\ x^t_{n_t 1} & \cdots & x^t_{n_t k_t} \end{bmatrix}$$

Where n_t denotes the number of patients in a particular quarter t, i.e. the number of observations in the data. k_t denotes the number of hospital dummies in each quarter. This number varies across quarters because of hospital entry, exit and mergers.

IV Mortality Rate Model Cont'd.

The matrix of instruments is arranged in a similar fashion:

$$Z = \begin{bmatrix} Z_1 & & & CM_1 \\ & Z_2 & & CM_2 \\ & & \ddots & \vdots \\ & & Z_{20} & CM_{20} \end{bmatrix}$$

with

$$Z_t = \begin{bmatrix} z^t_{11} & \cdots & z^t_{1l_t} \\ \vdots & \ddots & \vdots \\ z^t_{n_t 1} & \cdots & z^t_{n_t l_t} \end{bmatrix}$$

IV Mortality Rate Model Cont'd.

Where n_t denotes the number of patients in a particular quarter t (as in the X-matrix above). l_t denotes the number of quarter-specific instruments. In general we need the condition $l_t > k_t - 1$ to be fulfilled in all quarters (We do not need an instrument for the constant in each quarter, i.e. the average quarterly death rate over all hospitals). In practice we use the distance to each hospital available in the quarter and a dummy for whether this is the closest hospital for the individual patient as instruments. This yields $l_t = 2 * k_t$ instruments for each quarter.

Changes in Survival Probability due to the Reform

Change in Survival when Choices Post-Reform are Made with Pre-Reform Parameters		-12.17
Post-Reform (5 Quarters)	Admissions	20338
	Deaths	393
	Mortality Rate	1.93
	Recomputed Mortality Rate	1.87

Hospital Selection Based on Severity

Dependent Variable	Adjusted Mortality Rate	Adjusted Mortality Rate
Co-morbidity Count	-0.220**	-0.180**
	(0.006)	(0.006)
Quarter Fixed Effects (Flexible Time Trend)	No	Yes
Number of Observations	32,715	32,715