Appendix To:

Adoption and Learning Across Hospitals: The Case of a Revenue-Generating Practice

Adam Sacarny

June 2018

A Appendix

A.1 Revenue at Stake

To determine how a hospital would have been paid had it coded HF differently, I use a computer program called a grouper that translates an inpatient claim into its Medicare payment diagnosis-related group (DRG). I use the DRGGroupers.net Perl grouper software. For each patient i with a HF diagnosis, I process her claim as-is, then reprocess it replacing her secondary HF codes with a low-severity/non-CC code (428.0 – congestive HF, unspecified), medium-severity/CC code (428.22 – HF, systolic, chronic), and high-severity/MCC code (428.21 – HF, systolic, acute) using the Medicare DRG rules in year t^*. The result is a set of DRG weights $\left(w_i^{asis,t^*}, w_i^{noncc,t^*}, w_i^{cc,t^*}, w_i^{mcc,t^*}\right)$ – a measure of the expected cost of treatment for patients in each DRG that is uniform across hospitals. These weights are then used in the calculations for revenue at stake in a given year from higher intensity HF coding and to produce the *ex ante* revenue at stake for hospitals, described in the following sections.

A.1.1 Contemporaneous

To calculate the revenue at stake from HF coding in a given year t (as shown in Figure 2), I start with the set of all patients with HF in the grand sample in year t, P_t. I let C_t be the average conversion factor from DRG weights to dollars in year t (calculated by taking, for all patients in the MEDPAR file with FFS Medicare Part A & B coverage in year t, the average ratio of the "drgprice" variable to the DRG weight). The economy-wide potential gain per patient from chronic HF codes (expressed in constant 2009 dollars) is calculated as:

$$gainpp_t^{cc} = C_{2009} \times \frac{\sum_{i \in P_t} \left(w_i^{cc,t} - w_i^{noncc,t}\right)}{|P_t|}$$

The potential gain from acute HF codes is calculated as:

$$gainpp_t^{mcc} = C_{2009} \times \frac{\sum_{i \in P_t} \left(w_i^{mcc,t} - w_i^{noncc,t}\right)}{|P_t|}$$

These gains are visualized in Figure 2 for years $t = 2007 \ldots 2010$.

A.1.2 Predictor of Revenue at Stake

The revenue at stake from the reform for a particular patient depends on whether she was diagnosed with chronic or acute HF. I therefore construct a predictor of the acuity of the patient's HF. This predictor uses HF patients at hospitals that were relatively detailed coders in 2010 – hospitals that gave at least 85% of their HF patients a detailed code. The sample includes 90,653 patients and 171 hospitals. I regress whether the patient was coded as having high-severity HF on well-measured patient attributes: indicators for age, race, sex, month of admission, admission through the emergency department, 19 chronic conditions, and the 25 major diagnostic categories classifying the underlying cause of admission (the chronic conditions are listed in Appendix Section A.4).

I use the coefficients from this regression to fit the probability that a patient would have received a high-severity HF code under full adoption of the coding practice, $\hat{p}_i^{mcc}$, constraining the fitted value to be between 0 and 1. For patients who were already coded as getting a high severity code, I set $\hat{p}_i^{mcc} = 1$; patients already coded with a medium-severity code get $\hat{p}_i^{mcc} = 0$. Patients who do not receive a high-severity code are assumed to receive a medium-severity code i.e. $\hat{p}_i^{cc} = 1 - \hat{p}_i^{mcc}$. I then re-price these patients under the pricing rules of year t^*. Their expected DRG weight under full coding according to the payment rules of year t^* is defined as:

$$\hat{w}_i^{t^*} = \hat{p}_i^{mcc} w_i^{mcc,t^*} + \hat{p}_i^{cc} w_i^{cc,t^*}$$

The expected gain to using the detailed codes under the payment rules of year t^* equals the expected DRG weight under full coding less the DRG weight with no detailed codes:

$$\hat{gain}_i^{t^*} = \hat{w}_i^{t^*} - w_i^{nocc,t^*}$$

The *ex ante* per-patient gain from full HF coding for hospital h, depicted in Figure 3 and used in the analysis regressions, equals the rise per HF patient in DRG payments when the hospital's 2007 patients are processed under 2009 rules (expressed in 2009 dollars for consistency with the rest of the paper). Let $P_{h,t}$ be the HF patients at hospital h in year t for whom their chronic conditions are observed:

$$exantepp_h = C_{2009} \frac{\sum_{i \in P_{h,2007}} \hat{gain}_i^{2009}}{|P_{h,2007}|}$$

The depiction in Figure A2 follows the same formula but divides by the total number of patients, not just those with HF. To improve precision and reduce the leverage of outliers, when this predictor is used in the main regressions and displayed in the figures, hospitals with fewer than 50 HF patients in 2007 as well as those with an outlying top or bottom 1% of revenue on the table per patient were culled from this measure.

Figure 1 displays hospitals' capture of the HF revenue over time. The plot is at the weekly level and shows the fraction of revenue at stake that was captured according to the contemporaneous payment rules. It uses the aforementioned prediction algorithm to impute the probability that each patient has medium or high severity HF. Let weeks be indexed by k and let $t(k)$ be the year of week k; let P_k be all patients with HF in week k with chronic conditions observed. Since the figure also plots the revenue that would have been captured in 2007 if 2008 payment rules were in effect, let $\tilde{t}(k) = \max(t(k), 2008)$. Define the realized gain from specific coding for the patient according to rules of year t^* as:

$$gain_i^{t^*} = w_i^{asis,t^*} - w_i^{nocc,t^*}$$

Then each point in the figure is defined as:

$$capture_k = \frac{\sum_{i \in P_k} gain_i^{\tilde{t}(k)}}{\sum_{i \in P_k} \hat{gain}_i^{\tilde{t}(k)}}$$

A.2 Quality and Performance Measures

A.2.1 Standards of Care

I construct a composite measure of hospital utilization of standards of care by adding together standardized measures of AMI, HF, pneumonia, and surgery standards of care in 2006.

The AMI measure includes 8 processes (aspirin at arrival, aspirin at discharge, ACE inhibitors, smoking cessation advice, β-blockers at discharge, β-blockers at arrival, thrombolytics at arrival, and PCI at arrival). The heart failure measure includes 4 processes (discharge instructions, evaluation of left ventricular systolic function, ACE inhibitors, and smoking cessation advice). The pneumonia measure includes 7 processes (oxygenation assessment, pneumococcal vaccine, blood culture before antibiotics, smoking cessation advice, timely antibiotics, appropriate antibiotics, and influenza vaccine), and the surgery measure includes 3 measures (preventative antibiotics, appropriate antibiotics, and antibiotics stopped quickly).

For each of the 4 groups of scores, I calculate an overall score by summing together the numerators from all the component measures and dividing it by the sum of the denominators. I standardize this measure, then add together the four standardized measures and standardize the result, yielding one composite Z-score of process of care use.

A.2.2 Adjusted AMI Survival

I construct adjusted AMI survival by starting with a sample of all AMI episodes in FFS Medicare in fiscal years 2000-2006. This sample is generated as described in Chandra et al. (2013) and is a subset of the analysis sample used in that paper. I restrict the analysis to hospitals that treated at least 25 AMI patients during that time frame. I then regress an indicator for a patient's 30-day survival on age-race-sex interactions, logged inputs (real resources), 25 risk-adjusters, and hospital fixed effects.[16] The hospital fixed effects are then extracted and their standard errors estimated under a homoscedasticity assumption; they are then Empirical Bayes adjusted to account for measurement error when they are used in the analysis regressions (note that here I use constant weights across hospitals so that all facilities receive the same weight, whereas the procedure in the previous paper

[16] I use 8 additional risk-adjusters beyond those of Chandra et al. (2013) but constructed in the same way (i.e. on the basis of prior hospitalizations): heart failure, myocardial infarction, unstable angina, chronic atherosclerosis, respiratory failure, hypertensive heart disease, valvular heart disease, and arrhythmia.

uses optimal weights).

One difference between this study and Chandra et al. (2013) is that the latter used log-survival days censored at 1 year as its outcome measure, whereas I use an indicator for 30-day survival. In practice, these measures yield similar results in the main regressions when they are standardized because the two measures have a correlation coefficient of 0.916.

A.3 Coding at the Hospital System and Geographic Region Levels

Variations in coding can also be studied at the level of the hospital system and geographic region. The main text focuses on hospitals instead of systems and regions because the key economic questions concern provider behavior, and these alternative levels aggregate over providers. However, given the large literature on health care variations (which includes studies of coding, c.f. Song et al., 2010 and Finkelstein et al., 2017), and the potential role for hospital systems to drive diffusion, statistics at these levels may be of interest. I match hospitals to hospital systems using American Hospital Association survey data, to Dartmouth Hospital Referral Regions (HRRs) using ZIP code, and to Metropolitan Statistical Areas (MSAs) using the CMS Provider of Services file.

First, to get a sense of the explanatory power of these levels, I regress hospital coding scores on system/region fixed effects and report the R^2 adjusted for degrees of freedom and regressand (coding score) measurement error. The adjustment process is described at the end of this section.

The upper section of Table A4 presents the results. Each row estimates the hospital score with different first-step controls. Of the 2,341 hospitals in the analysis sample, 1,521 were in a hospital system, and there were 321 distinct systems. All hospitals had an HRR (since these regions partition the entire U.S.) and 1,705 were in an MSA. Both hospital system and geographic fixed effects left the majority of hospital coding variations unexplained, though both levels had nontrivial explanatory power: systems and geographies explained one-sixth to one-fourth of variation in coding across hospitals when the physician component was not removed in the first step and one-tenth to one-fifth of variation when the physician component was removed.

The lower section of Table A4 estimates the dispersion in system and region effects directly. I estimate equation 1 replacing the hospital and physician fixed effects with (respectively) system, HRR, and MSA fixed effects. As shown previously for hospitals and physicians, variations attenuate with the addition of patient characteristics observable upon admission, but do not further attenuate with additional controls for chronic conditions. There are meaningful variations across systems and regions, though the magnitudes are smaller than variations at the hospital and physician level. Accounting for patient characteristics, the standard deviation in coding across hospital systems is 12.1 percentage points. The standard deviation is 8.3 percentage points across HRRs and 9.9 percentage points across MSAs.

Then, in Table A5, I show the relationship between HRR coding scores and Medicare spending per enrollee from the 2010 Dartmouth Atlas by regressing the former on the latter. Spending is adjusted for demographics and prices (though not coding intensity – the price adjustment removes geographic factors, using, for example, nationally standardized prices for DRGs). Perhaps surprisingly, columns 1-3 of the table show that higher-spending regions are less likely to code heart failure in a detailed fashion; columns 4-6 show no association between inpatient spending (hospital and SNF) and coding.

However, breaking down spending into its components and looking at conditional associations, columns 7-9 find that inpatient spending (hospital and SNF) is positively associated with coding, while physician spending is negatively associated with it. That is, holding fixed the other components of spending, areas with higher hospital payments per enrollee tended to code more, while areas with higher physician payments per enrollee tended to code less.

A.3.1 Adjustment of R^2

To compute the R^2, I initially regress the hospital coding score on a set of level fixed effects:

$$\hat{\alpha}_h = \alpha_{l(h)} + \xi_h + \epsilon_h$$

Where h indexes hospitals and l indexes the level (system or region). $\hat{\alpha}_h$ is the estimate of the hospital coding score and α_l is the level effect. The regression error term has two components. ξ_h is measurement error in the regressand ($\hat{\alpha}_h = \alpha_h + \xi_h$). ϵ_h is the remaining error – the true regression error term that would have occurred with a precisely measured regressand. Because the overall error term includes measurement error, R^2 will understate the fraction of variation in underlying hospital coding that is explained by the fixed effects:

$$R^2 = 1 - \frac{\text{Var}\left(\xi_h + \epsilon_h\right)}{\text{Var}\left(\hat{\alpha}_h\right)} = 1 - \frac{\text{Var}\left(\xi_h\right) + \text{Var}\left(\epsilon_h\right)}{\text{Var}\left(\alpha_h\right) + \text{Var}\left(\xi_h\right)} < 1 - \frac{\text{Var}\left(\epsilon_h\right)}{\text{Var}\left(\alpha_h\right)}$$

I adjust the R^2 for regressand measurement error by estimating $\text{Var}\left(\xi_h\right)$ as the average squared standard error of the first-step hospital fixed effects. I estimate $\text{Var}\left(\xi_h + \epsilon_h\right)$ as the mean squared error of the regression and $\text{Var}\left(\hat{\alpha}_h\right)$ as the sample variance of the hospital coding scores. These estimates use a degrees of freedom adjustment and so already will yield a degrees of freedom adjusted

R^2 (i.e. what is typically called an adjusted R^2). Plugging in the estimates I arrive at my R^2 adjusted for measurement error and degrees of freedom:

$$\hat{R}^2_* = 1 - \frac{\hat{\text{Var}}\left(\xi_h + \epsilon_h\right) - \hat{\text{Var}}\left(\xi_h\right)}{\hat{\text{Var}}\left(\hat{\alpha}_h\right) - \hat{\text{Var}}\left(\xi_h\right)}$$

Note that the denominator is the formula used to estimate the underlying variance of the hospital effects.

A.4 First-Step Specifications

I use four specifications to estimate the hospital fixed effects using the analysis sample described in Section 4.1 of the main text. The regressions are of the form of equation 1. The first three of these specifications use no physician fixed effects while the final specification adds physician fixed effects. The regressions are run in stata with the commands `felsdvregdm` (Mihaly et al., 2010) and `reghdfe` (Correia, 2016). Standard errors are derived from `felsdvregdm` and a modified version of the command `fese` (Nichols, 2008).

The first specification uses no controls at all. The second specification uses only controls that were observable from the patient's admission and not historical data. These controls are: age-race-sex interactions (age in 5 year categories starting at 65 and with age 90+ treated as one category, race as white/nonwhite, sex as female/not female), month of year indicators, an indicator for being admitted through the emergency department, and indicators for 179 categories of the primary diagnosis code. The 179 categories are constructed from the HCUP Clinical Classifications Software ICD-9 diagnosis code multi-level categories. The aim is to include an indicator for each commonly used category of codes and roll up uncommon categories that are clinically similar. Starting with the most finely grained level (level 4), categories comprising at least 0.1% of the population were included as indicators. Categories comprising less than 0.1% were replaced with their level 3 codes. Then, looking at the level 3 codes, those comprising at least 0.1% were included as indicators, and the rest were replaced with their level 2 codes, and so on for levels 2 and 1.

The third specification adds to these controls additional indicators for the patient's history of chronic conditions. These indicators are based on the Medicare Chronic Conditions segment. This file reports whether patients had received a diagnosis for the conditions in Medicare claims during a reference period ranging from 1-3 years. An indicator is provided for each condition at midyear and at the end of the year. To identify preexisting conditions only, I use the most recent report of chronic conditions that occurred before the patient's admission to the hospital. I include indicators for 19 chronic conditions: acute myocardial infarction, atrial fibrillation, cataract, chronic kidney disease, COPD, HF, diabetes, glaucoma, hip fracture, ischemic heart disease, depression, osteoporosis, rheumatoid arthritis or osteoarthritis, stroke or transient ischemic attack, breast cancer, colorectal cancer, prostate cancer, lung cancer, and endometrial cancer.

The fourth specification amends the third specification to add physician fixed effects.

A.5 Procedure for Bias-Corrected Dispersion of High Dimensional Fixed Effects

Here I describe a feasible approach to estimating the underlying variance of hospital and physician fixed effects. The original formulas for this approach were laid out in Andrews et al. (2008), but involved the construction and inversion of a large matrix. Gaure (2014) then developed a version that bypassed the direct creation of this matrix. My procedure is nearly identical to Gaure (2014), except that where he estimates the observation-level (i.e. patient-level) underlying variance of the effects, the procedure here estimates the hospital-level and physician-level variance of the two sets of effects, respectively.

I assume outcomes are an additively separable function of patient-level covariates and hospital and physician effects:

$$Y = X\beta + D\theta + F\psi + \epsilon \tag{A1}$$

Where Y is an $N \times 1$ vector of outcomes, X is an $N \times K_\beta$ matrix of patient-level covariates, D is an $N \times K_\theta$ matrix of physician effects, F is an $N \times K_\psi$ matrix of hospital effects, and ϵ is an $N \times 1$ vector of disturbances. β is a vector of coefficients on the patient level covariates, θ are the physician fixed effects, and ψ are the hospital fixed effects.

I assume uncorrelated and homoscedastic disturbances:

$$\mathbb{E}\left[\epsilon\epsilon'\right] = \sigma_\epsilon^2 \times I_N$$

Let $M_{X,D}$ be the orthogonal projection matrix with respect to X and D. Premultiplying A1 by $M_{X,D}$ yields:

$$M_{X,D}Y = M_{X,D}F\psi + M_{X,D}\epsilon$$

Then standard OLS regression yields estimates $\hat{\psi}$ for ψ:

$$\hat{\psi} = \left(F'M_{X,D}F\right)^{-1} F'M_{X,D}Y$$

Plugging in A1 to replace Y with the right-hand side, I express $\hat{\psi}$ as a sum of a true component and a sampling noise component:

$$\hat{\psi} = \psi + \left(F' M_{X,D} F\right)^{-1} F' M_{X,D} \epsilon$$

Now define the true and estimated variance of the hospital effects:

$$\sigma_{\psi}^2 = \frac{\psi' M_1 \psi}{K_{\psi}},\ \tilde{\sigma}_{\psi}^2 = \frac{\hat{\psi}' M_1 \hat{\psi}}{K_{\psi}}$$

Where M_1 is the demeaning matrix i.e. $I - 1\left(1'1\right)^{-1} 1$. The expected value of $\tilde{\sigma}_{\psi}^2$ has a true component and bias term due to sampling variance:

$$\begin{aligned}
\mathbb{E}\left[\tilde{\sigma}_{\psi}^2\right] &= \sigma_{\psi}^2 + \delta_{\psi} \\
\delta_{\psi} &= \mathbb{E}\left[\frac{\epsilon' M_{X,D} F \left(F' M_{X,D} F\right)^{-1} M_1 \left(F' M_{X,D} F\right)^{-1} F' M_{X,D} \epsilon}{K_{\psi}}\right]
\end{aligned}$$

Relying on the exchange of the expectation and trace operators as well as the invariance of the trace to cyclic permutations, the bias term simplifies to:

$$\begin{aligned}
\delta_{\psi} &= \operatorname{tr}\left(\mathbb{E}\left[\epsilon\epsilon'\right] \frac{M_{X,D} F \left(F' M_{X,D} F\right)^{-1} M_1 \left(F' M_{X,D} F\right)^{-1} F' M_{X,D}}{K_{\psi}}\right) \\
&= \frac{\sigma_{\epsilon}^2}{K_{\psi}} \times \operatorname{tr}\left(M_1 \left(F' M_{X,D} F\right)^{-1}\right) \qquad \text{(A2)}
\end{aligned}$$

The trace term requires the inversion of a potentially large matrix. To avoid directly computing this matrix, I utilize an approach described in Gaure (2014). I note the following equality for $\mathbb{E}\left[x\right] = 0$ and $\operatorname{Var}\left(x\right) = 1$:

$$\operatorname{tr}\left(A\right) = \mathbb{E}\left[x' A x\right] \qquad \text{(A3)}$$

So I can compute the bias term with a stochastic approximation, i.e.:

$$\operatorname{tr}\left(M_1 \left(F' M_{X,D} F\right)^{-1}\right) = \mathbb{E}\left[x' M_1 \left(F' M_{X,D} F\right)^{-1} x\right]$$

Per Gaure (2014), the variance-minimizing $x'Ax$ for symmetric A is from an x that is m independent draws from a discrete uniform distribution over the sign function i.e. $\{-1, 1\}$. I define x in this way. Now define v:

$$v = \left(F' M_{X,D} F\right)^{-1} x$$

Computing v from x would require inverting a large matrix, but I can also write v as the vector that solves the following linear system:

$$F' M_{X,D} F v = x \tag{A4}$$

Equation A4 can be solved using a linear solver, though this would require computing $F' M_{X,D} F$, which is memory intensive as N and K_ψ grow. Thus for large N and K_ψ I mimic Gaure, 2014 in using a conjugate gradient algorithm. The advantage of this method is that $F' M_{X,D} F$ need never be directly calculated. The vector Fv merely needs to be residualized with respect to X and D, which is readily done by regression fixed effects methods.

Then plugging in x and v:

$$\delta_\psi = \frac{\sigma_\epsilon^2}{K_\psi} \times \mathbb{E}\left[x' M_1 v\right]$$

Let $\hat{\delta}_\psi$ be a feasible estimate of this bias term. One such feasible estimate is to replace σ_ϵ^2 in equation A2 with $\hat{\sigma}_\epsilon^2$:

$$\hat{\delta}_\psi = \frac{\hat{\sigma}_\epsilon^2}{K_\psi} \times \operatorname{tr}\left(M_1 \left(F' M_{X,D} F\right)^{-1}\right) \tag{A5}$$

However, since computing the trace term is computationally intensive, I define the stochastic approximation to the feasible estimate:

$$\hat{\delta}_\psi^s = \frac{\hat{\sigma}_\epsilon^2}{K_\psi} \times \hat{\mathbb{E}}\left[x' M_1 v\right] \tag{A6}$$

Then the feasible estimate of underlying variance is the variance of the estimated fixed effects less the bias estimate:

$$\hat{\sigma}^2_\psi = \tilde{\sigma}^2_\psi - \hat{\delta}^s_\psi$$

This estimate can be computed up to arbitrary levels of precision by repeatedly drawing new x vectors, re-running the algorithm to estimate $\hat{\delta}^s_\psi$, and taking the average of the results.

By identical argument, a feasible estimate of the underlying variance of the physician effects can be constructed as:

$$\begin{aligned}
\hat{\sigma}^2_\theta &= \tilde{\sigma}^2_\theta - \hat{\delta}_\theta \\
\tilde{\sigma}^2_\theta &= \frac{\hat{\theta}' M_1 \hat{\theta}}{K_\theta} \\
\hat{\delta}_\theta &= \frac{\hat{\sigma}^2_\epsilon}{K_\theta} \times \text{tr}\left(M_1 \left(D' M_{X,F} D\right)^{-1}\right) \\
\hat{\delta}^s_\theta &= \frac{\hat{\sigma}^2_\epsilon}{K_\theta} \times \hat{\mathbb{E}}\left[w' M_1 y\right]
\end{aligned}$$

Where w is an iid vector drawn from the sign distribution like x, and y solves $D' M_{X,F} D y = w$.

Appendix Figures

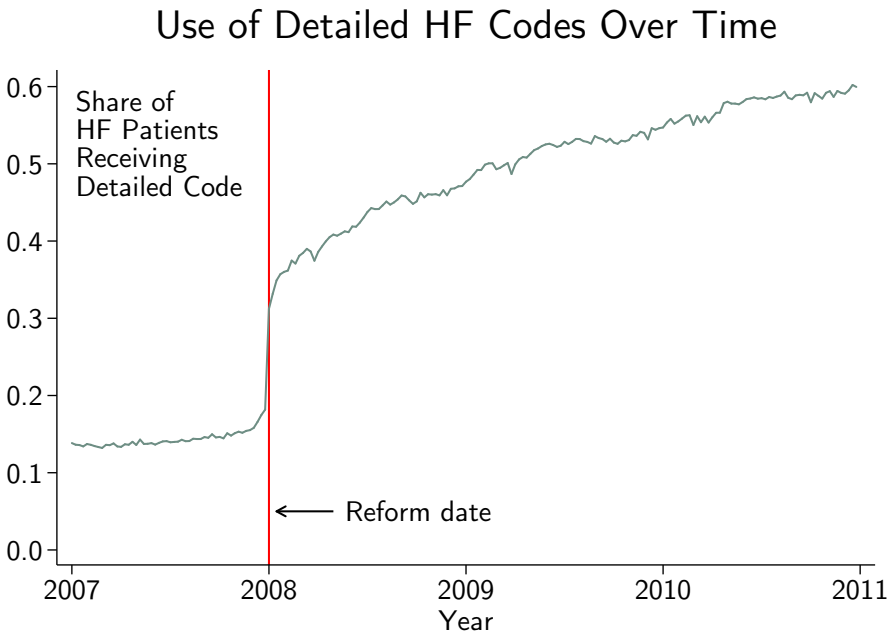
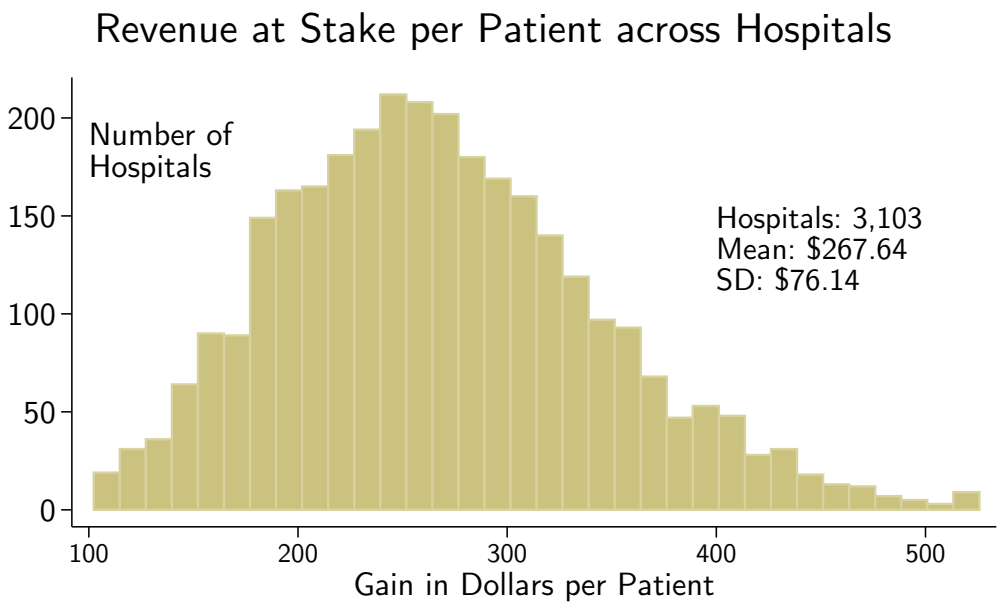


Figure A1



Revenue at stake is calculated using pre-reform (2007) patients processed under post-reform (2009) payment rules. The prediction process is described in the appendix. The 422 hospitals with <50 HF patients are suppressed and the upper and lower 1% in revenue at stake per patient are then removed.

Figure A2

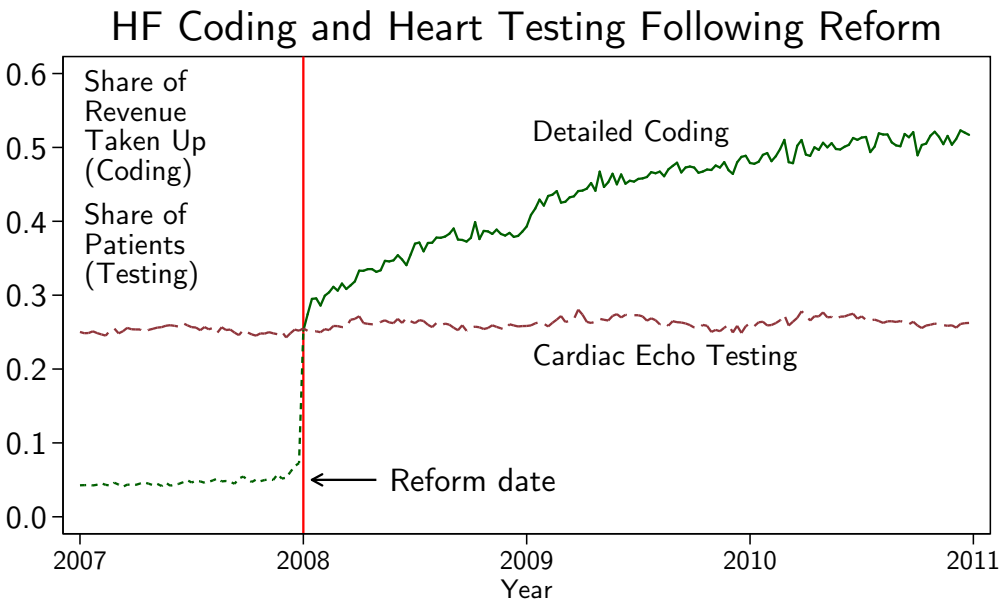
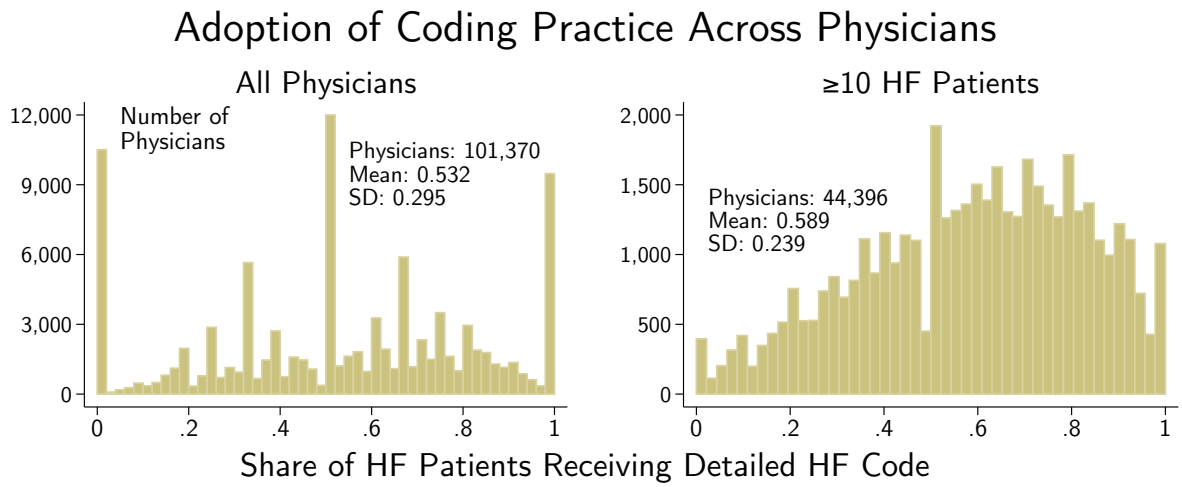


Figure plots the weekly share of revenue available for detailed coding of HF that was captured by hospitals alongside the weekly share of all patients who received a cardiac echo, a heart test. The dotted line shows revenue that would have been captured in 2007 if hospitals had been paid per 2008 rules.

Figure A3



Histogram plots the share of HF patients who received a detailed HF code across physicians. To reduce dispersion due to measurement error, the right panel restricts to physicians who treated ≥10 HF patients. Standard deviations are unadjusted for measurement error.

Figure A4

Appendix Tables

Table A1 - Other Codes for HF

		Severity		Counted as
Code	Description	Before	After	Specific
428.1	Left HF	High	Medium	Yes
398.91	Rheumatic HF	High	Medium	Yes
402.01	Malignant HHD w/ HF	High	Medium	Yes
402.11	Benign HHD w/ HF	High	Low	No
402.91	Unspecified HHD w/ HF	High	Low	No
404.01	Malignant HHCKD, CKD stage 1-4/unspec w/ HF	High	Medium	Yes
404.11	Benign HHCKD, CKD stage 1-4/unspec w/ HF	High	Medium	Yes
404.91	Unspecified HHCKD, CKD stage 1-4/unspec w/ HF	High	Medium	Yes
404.03	Malignant HHCKD, CKD stage 5/ESRD w/ HF	High	Medium	Yes
404.13	Benign HHCKD, CKD stage 5/ESRD w/ HF	High	Medium	Yes
404.93	Unspecified HHCKD, CKD stage 5/ESRD w/ HF	High	Medium	Yes

This table lists ICD-9 codes besides those of Table 1 that indicate heart failure. These codes can be used alongside the codes listed in Table 1. Codes that raise patients to medium or higher severity after the reform are counted as specific codes. Definitions: HHD - hypertensive heart disease, HHCKD - hypertensive heart and chronic kidney disease, CKD - chronic kidney disease, ESRD - end stage renal disease.

Table A2 - Relationship Between Coding and Testing

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Detailed HF Coding			Echocardiogram		
Post ×	0.46***	0.46***	0.45***	0.067***	0.059***	0.059***
Ex Ante HF Rate	(0.024)	(0.023)	(0.022)	(0.014)	(0.013)	(0.013)
Observations	47,355,656	47,355,656	47,355,656	9,460,759	9,460,759	9,460,759
R^2	0.076	0.082	0.117	0.087	0.098	0.101
Patient Controls	None	A/R/S	Full	None	A/R/S	Full

These regressions are at the patient level and are based on the "grand sample" described in the text. The sample is expanded to include patients with and without HF, and is extended to include all such patients from 2005-2010. Post is an indicator for year 2008 and later. Ex ante HF rate is the fraction of patients in the patient's major diagnostic category (MDC) in 2003-2004 with HF. The outcome in columns 1-3 is an indicator for whether the patient received a detailed HF code and in columns 4-6 it is an indicator for whether the patient received an echocardiogram (these columns use the 20% of observations for which echocardiograms are observed). All regressions include year and MDC fixed effects. Columns 1 and 4 include no patient controls. Columns 2 and 5 control for age, race, and sex interactions. Columns 3 and 6 further add controls for histories of chronic conditions. Standard errors clustered at the MDC level in parentheses.

*** significant at 1% level; ** significant at 5% level; * significant at 10% level

Table A3 - Coding Dispersion and Hospital Characteristics by Study Sample

Sample Subset	(1) Grand Sample All Hospitals	(2) Grand Sample ≥50 HF	(3) Analysis Sample Step 1	(4) Analysis Sample Step 2
A. Heart Failure Coding Rates and Dispersion				
Average No. of HF Patients	552.7	610.8	533.7	601.8
Average Share Specific Code	0.494	0.524	0.526	0.546
Raw SD of Share Specific Code	0.244	0.220	0.220	0.201
Adjusted SD of Share Specific Code	0.230	0.218	0.212	0.199
B. Hospital Characteristics				
Beds	233.2	248.2	260.4	287.9
Ownership				
Government	0.189	0.181	0.165	0.167
Non-Profit	0.605	0.631	0.630	0.671
For-Profit	0.207	0.188	0.205	0.161
Location				
Rural Area	0.284	0.291	0.216	0.224
Large Urban Area	0.384	0.387	0.426	0.422
Other Urban Area	0.316	0.317	0.347	0.354
Teaching Status				
Non-Teaching	0.678	0.677	0.642	0.623
Major Teaching Hospital	0.082	0.086	0.094	0.101
Minor Teaching Hospital	0.222	0.231	0.251	0.276
Hospitals	3,414	3,081	2,831	2,341

This table shows how HF coding and hospital characteristics vary across the different samples discussed in the text. The grand sample refers to all HF patients and is described in Sections 2.3 and 4.1; I present statistics including all hospitals (column 1) and dropping hospitals with fewer than 50 HF patients (column 2). The step 1 analysis sample (column 3) includes only patients for whom attending physician and chronic condition histories are observed and then restricts to the largest mobility group of hospitals; it is described in Section 4.1. The step 2 analysis sample (column 4) is the subset of these hospitals for which all characteristics are observed, and is the focus of Sections 4.3-4.5.

Panel A presents the average and standard deviation of HF coding across hospitals in 2010. The Adjusted SD statistic accounts for measurement error (see Section 3.1.2). Panel B shows hospital characteristics as averages and shares. In columns 1-3, at most 2% of hospitals could not be matched to the data on characteristics (Impact file for location and Provider of Services for the remaining covariates). In these cases the hospitals were omitted from the relevant statistics. All hospitals matched in column 4 as matching was a sample restriction.

Table A4 - Coding at the Hospital System and Region Levels

Level	(1) Hospital System	(2) Region (HRR)	(3) Region (MSA)
Share of Variation in Hospital Effects Explained by System/Region Fixed Effects			
No First-Step Controls	0.249	0.163	0.166
First-Step Admission Patient Controls	0.261	0.234	0.175
First-Step Full Patient Controls	0.260	0.233	0.176
First-Step Full Patient Controls and Physician FE	0.093	0.168	0.207
Observations (Number of Hospitals)	1,521	2,341	1,705
Fixed Effects (Number of Systems/Regions)	321	299	311
Standard Deviation of Coding Across Systems/Regions			
No Controls	0.167	0.109	0.129
Admission Patient Controls	0.122	0.084	0.100
Full Patient Controls	0.121	0.083	0.099
Number of Systems/Regions	341	303	314

This table analyzes variations in coding at the hospital system and geographic region levels. The upper section shows the share of variation in coding across hospitals that can be explained by regressing hospital coding on system (column 1) and region (columns 2 and 3) fixed effects. Shares are adjusted for regressand measurement error. The set of hospitals is restricted to the 2,341 in the previous dispersion and regression analyses.

The lower section estimates the dispersion in coding across systems/regions by running equation 1 with system or region fixed effects in place of hospital and physician fixed effects. Dispersion is adjusted for measurement error. The sample is all patients treated at the 2,831 "step one" analysis sample hospitals – the same sample previously used to estimate hospital coding scores.

Hospital systems are derived from AHA data. Regions are measured using hospital referral regions (HRRs) and metropolitan statistical areas (MSAs) based on the location of the hospital.

Table A5 - Association Between Region Characteristics and Coding

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome	Score	Score	Score	Score	Score	Score	Score	Score	Score
Total spending per enrollee	-0.015***	-0.013***	-0.013***						
	(0.005)	(0.004)	(0.004)						
Hospital and SNF spending per				-0.006	-0.007	-0.007	0.046***	0.032***	0.031***
enrollee				(0.010)	(0.008)	(0.008)	(0.013)	(0.010)	(0.010)
Physician spending per enrollee							-0.082***	-0.063***	-0.065***
							(0.016)	(0.012)	(0.012)
Outpatient facility spending per							-0.062**	-0.046*	-0.047**
enrollee							(0.030)	(0.023)	(0.023)
Home health spending per enrollee							-0.058**	-0.037**	-0.036**
							(0.026)	(0.018)	(0.018)
Hospice spending per enrollee							0.018	-0.007	-0.005
							(0.053)	(0.041)	(0.041)
Durable medical equipment spending							-0.220	-0.183	-0.168
per enrollee							(0.156)	(0.123)	(0.122)
Observations	303	303	303	303	303	303	303	303	303
R^2	0.030	0.036	0.036	-0.002	-0.000	-0.000	0.110	0.108	0.109
Patient Controls	None	Admission	Full	None	Admission	Full	None	Admission	Full
Physician Controls	None	None	None	None	None	None	None	None	None

This table presents the results of regressing Hospital Referral Region (HRR) coding scores on HRR characteristics, taken from Dartmouth Atlas data on Medicare spending in 2010 across HRRs. All spending is measured in $1,000s and adjusted for price, age, sex, and race. Columns 1, 4, and 7 use no controls to calculate the HRR scores; columns 2, 5, and 8 add controls for patient characteristics observable upon admission; and columns 3, 6, and 9 add histories of chronic conditions. Robust standard errors in parentheses.

*** significant at 1% level; ** significant at 5% level; * significant at 10% level