How Do Hospitals Respond to Price Changes?

By Leemore S. Dafny*

This paper examines hospital responses to changes in diagnosis-specific prices by exploiting a 1988 policy reform that generated large price changes for 43 percent of Medicare admissions. I find hospitals responded primarily by "upcoding" patients to diagnosis codes with the largest price increases. This response was particularly strong among for-profit hospitals. I find little evidence hospitals increased the volume of admissions differentially for diagnoses subject to the largest price increases, despite the financial incentive to do so. Neither did they increase intensity or quality of care in these diagnoses, suggesting hospitals do not compete for patients at the diagnosis level. (*JEL* H0, I0, L0)

The vast majority of U.S. healthcare is privately provided. Yet until the 1980s, the sector was largely immune from standard market forces promoting efficiency in production. The canonical healthcare market imperfections—informational asymmetries between providers and consumers, and an insurance-induced wedge between marginal out-of-pocket costs and patient benefits—were exacerbated by a cost-plus-reimbursement system and primarily not-for-profit providers. So long as providers could always earn nonnegative profits, there was little supply-side incentive to cut costs, and consumers' incentives via co-payments and deductibles were weak.

In 1984, the federal government injected market discipline into the system by implementing the Prospective Payment System (PPS). Under PPS, hospitals receive a fixed payment for each Medicare patient in a given diagnosis-related group (DRG), regardless of the actual expenses the hospital incurs in caring for the patient.[1] Other public and private insurers subsequently implemented similar payment schemes, wresting price-setting control from providers and imposing "yardstick competition" in which providers are rewarded based on performance relative to the average (Andrei Shleifer, 1985).

A large literature documents hospitals' responses to the introduction of PPS, but few studies have explored reactions to *changes* in these prices. Yet once the transition to a fixed-price regime is complete, price levels constitute the sole lever in the system, and there remain several unanswered empirical questions regarding their effect. In the face of a price increase for a particular diagnosis or treatment, will hospitals find ways to attract more such patients? Will they compete more vigorously for these patients by improving the quality of their care, thereby dissipating some of the rents from the price increase? Will they game the system and shift patients to higher-paying diagnoses without altering any aspects of their care? The answers to these questions are critical to ongoing policy decisions, and can also provide valuable insights into hospital industry conduct and the effectiveness of fixed-price regulation more generally.

The main challenge in studying these questions is locating an exogenous source of varia-

* Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, National Bureau of Economic Research, and Institute for Policy Research at Northwestern University (e-mail: l-dafny@kellogg.northwestern.edu). I am grateful to Jonathan Gruber, James Poterba, David M. Cutler, and Glenn Ellison for excellent guidance. I thank two anonymous referees, Josh Angrist, William Collins, Joseph Doyle, David Dranove, Mark Duggan, David Levine, Joseph Newhouse, Scott Stern, Nancy Rose, and seminar participants at several universities for helpful comments. Financial support from the National Science Foundation, the National Bureau of Economic Research, and the National Institute on Aging is gratefully acknowledged. The MEDPAR data are confidential and cannot be released.

[1] Hospitals receive additional payments for admissions whose costs exceed the annual "outlier threshold" set by the Centers for Medicare and Medicaid Services. Total outlier payments comprise roughly 8 percent of payments under PPS (68 *Federal Register* 34122).

tion in DRG prices, which are typically adjusted to reflect changes in hospital costs. As a result, positive associations between changes in price and changes in spending or intensity of treatment likely reflect bilateral causality, and do not constitute a priori evidence that hospitals alter treatment patterns in response to price changes. To obtain unbiased estimates of hospital responses to price changes, this study exploits a 1988 policy change that generated large price changes for 43 percent of Medicare admissions.[2] The policy change was the elimination of "age over 69" and "age under 70" in the descriptions of the DRGs to which patients may be assigned. Qualifiers that formerly read "with complications or age over 69" and "without complications and age under 69" now read "with complications" or "without complications." This seemingly innocuous modification, which is described in greater detail in Section I, actually led to substantial changes in prices for DRGs with these qualifiers.

I consider both "nominal" and "real" responses to these price changes, where nominal refers to hospital coding practices and real refers to admissions volumes and intensity of care actually provided. Because hospitals are responsible for coding patients to the appropriate DRGs, raising prices for certain DRGs may entice hospitals to "upcode," or switch patients from lower-paying to higher-paying DRGs. While upcoding does not affect real elements of patient care, it inflates hospital reimbursements. This was the primary response of hospitals to the 1988 policy change. Hospitals also demonstrated a keen awareness of risk-reward trade-offs in their upcoding, which can lead to fines and criminal charges if detected. Although the policy shock created a blanket incentive to increase upcoding in dozens of diagnoses, hospitals upcoded most in those diagnoses where the incentive to do so was greatest. The upcoding response was also strongest among for-profit hospitals, a finding that is consistent with prior research.

The real responses to the policy change were less finely tuned. I find little evidence that hospitals adjusted the intensity or quality of their care in response to changes in diagnosis-specific prices, where intensity is measured by total costs, length of stay, number of surgical procedures, number of intensive-care-unit (ICU) days, and quality by the in-hospital death rate. Rather, hospitals spread the additional funds across *all* admissions. I also find no conclusive evidence that hospitals attracted more patients in diagnoses subject to larger price increases.

The remainder of the paper is organized into six sections. Section I describes PPS and prior related research, and lays out the hypotheses I test. Section II offers a detailed explanation of the 1988 reclassification, followed by a discussion of the data in Section III. Sections IV and V examine the nominal and real responses to price changes, respectively, and Section VI concludes.

I. Background

A. *A PPS Primer*

The defining element of PPS is a reimbursement amount that is fixed regardless of a hospital's actual expenditures on a patient. This payment does vary, however, by the patient's medical diagnosis. Diagnoses are grouped into approximately 500 DRGs. Each DRG is assigned a weight (called a "DRG weight") that reflects the relative resource intensity of admissions within that group. Reimbursement to hospital h for an admission in DRG d is given by

$$P_{hd} = P_h \cdot (1 + IME_h) \cdot (1 + DSH_h) \cdot DRG\ weight_d \tag{1}$$

where P_h is a hospital-specific amount (inflated annually by a congressionally approved "update factor"), IME_h represents an adjustment for indirect medical education (teaching), and DSH_h adjusts payment levels to compensate hospitals with a disproportionate share of indigent patients.[3]

Most of the variation in P_{hd} is due to the DRG weights, which range between 0.09 (DRG 448 for allergic reactions) to 22.8 (DRG 480 for liver transplants).[4] The Health Care Financing

[2] In 2000, Medicare beneficiaries accounted for 37 percent of hospital discharges and 31 percent of total revenues (Data Compendium, *Centers for Medicare and Medicaid Services,* and author's tabulations from the 2000 *Survey of Hospitals* by the American Hospital Association (AHA)).

[3] This simplified formula appears in David M. Cutler (1995).

[4] The range for DRG weights is given for 1985–1996.

Administration (HCFA) uses hospital cost data (obtained by deflating charges using annual cost:charge ratios) to recalibrate the weights annually, raising weights for DRGs that experience relative increases in average costs, and reducing weights for DRGs with relative decreases in average costs.[5] The average DRG weight per hospital admission has risen substantially over time, from 1.13 in 1984 to 1.36 in 1996.[6] This phenomenon has been termed "DRG creep," as patients are increasingly coded into DRGs with higher weights. A 1-percent increase in the average case weight is associated with an additional $930 million in *annual* Medicare payments to hospitals.[7] To reduce the cost of DRG creep, HCFA often sets the annual update factor below the average growth in costs.

Although the implementation of PPS eliminated *marginal* reimbursement for services rendered (within a given DRG, hospitals are not compensated more when they spend more on a patient), economists have noted that *average* payment incentives remain. If P_{hd} is low relative to actual costs in DRG d, hospitals have an incentive to reduce the intensity of care and the number of admissions in that DRG. Due to the regular recalibrations described above, it is difficult to identify hospital responses to changes in average payment incentives (hereafter DRG *prices* or *weights*). When costs increase, DRG prices increase. Thus, the coefficient on DRG price in a regression of costs (or some other measure of intensity of care) on DRG price would suffer from a strong upward bias. To obtain an unbiased estimate of this coefficient, exogenous variation in payment levels is required. This variation is provided by the natural experiment described in Section II.

B. *Hypotheses and Prior Research*

Because more than 80 percent of hospitals are not-for-profit or government-owned, economists typically assume that hospitals maximize an objective function with nonnegative weights on patient care (often called "intensity" or quality) and profits (see, for example, David Dranove, 1988; Burton A. Weisbrod, 2005). Although not-for-profit and government hospitals are also subject to a nondistribution constraint, limiting their ability to distribute profits to stakeholders, under most scenarios they, too, will pursue profit-increasing activities in order to finance their missions. Therefore, a price increase for a particular diagnosis constitutes an incentive to "produce" more such diagnoses.

Under PPS, there are two principal means for producing more: coding existing patients into the diagnosis subject to the price increase, and attracting additional patients in that diagnosis through quality or intensity improvements (loosely defined to include better care, patient amenities, stronger referral networks, etc.).[8] I refer to these as "nominal" and "real" responses, respectively, as the former are essentially accounting tricks and the latter are real aspects of treatment.

Nominal Responses.—The coding of patient conditions is performed by administrative staff, who use specialized software, hospital charts, and diagnosis codes provided by physicians to map patient conditions into DRGs. Hospitals can pursue a variety of approaches to facilitate upcoding into higher-paying DRGs. First, physicians often rely on administrative or nursing staff to identify diagnosis codes, and these staff members can be trained to err on the side of more lucrative diagnoses. Second, hospitals may try to persuade physicians to alter their diagnoses in order to increase hospital revenues.[9] Finally, administrative staff may falsify patient records and/or assign DRGs incorrectly

[5] HCFA is now known as The Centers for Medicare and Medicaid Services. However, I refer to HCFA throughout this paper, as this was the agency's acronym during the study period.

[6] Bruce Steinwald and Laura A. Dummit (1989), author's calculations. The original 1984 weights were constructed so that the average DRG weight for hospitals, called the *case-mix index,* would equal 1.

[7] "Program Information," Centers for Medicare and Medicaid Services, June 2002.

[8] Note Medicare beneficiaries face the same out-of-pocket costs at all hospitals for covered stays, hence hospitals cannot reduce prices to attract these patients. The exception is Medicare HMO enrollees, who accounted for fewer than 3 percent of Medicare beneficiaries during the study period (Lauren A. Murray and Franklin J. Eppig, 2002). Reimbursement for these patients is not determined by PPS, and co-payments may differ across hospitals.

[9] For example, one resident I interviewed described an interaction with hospital coding personnel in which she was asked to reconsider her diagnosis of "urinary tract infection" and replace it with "septicemia" (which occurs when the bacteria enter the bloodstream), as the hospital is

to inflate reimbursements. Hospitals engaging in these practices face substantial penalties if audits reveal systematic efforts to defraud the Medicare program.[10]

When deciding whether to upcode a particular patient or group of patients, hospitals trade off the added revenue (less any change in treatment costs) against the increased risk of detection plus the cost of upcoding. In its purest form, upcoding implies no effect whatsoever on the amount of care received by patients, so treatment costs are unchanged. *Ceteris paribus,* a price increase for a given DRG therefore increases the incentive to upcode patients into that DRG. The policy change I study involves DRGs with particularly low upcoding costs and risks of detection. In these DRGs, simply coding complications such as hypotension (low blood pressure) onto a patient's chart results in a substantially higher price, although the primary diagnosis remains the same. (Compare this scenario to upcoding a patient suffering from bronchitis to the heart transplant DRG.) Section IV examines how increases in the prices paid for complications affected the reported incidence of complications.

The subject of upcoding has generated a substantial literature, not least because the rapid rise in the average case weight is the single largest source of increased hospital spending by Medicare. Another concern is that the strategy used to reduce the resulting financial burden (smaller increases in the annual update factor) punishes all providers uniformly, regardless of the extent to which they upcode. In addition, upcoding may result in adverse health consequences due to corrupted medical records. Robert F. Coulam and Gary L. Gaumer (1991) review the upcoding literature through 1990, concluding that there is evidence of upcoding during the first few years of PPS, but the amount of the case-mix increase attributable to this practice is unknown. There are two general empirical approaches to estimating the magnitude of upcoding: detailed chart review, and comparisons of case-mix trends over time and across hospitals.

Grace M. Carter et al. (1990) created the "gold standard" in chart review to estimate the role of upcoding in the case-mix increase between 1986 and 1987: they sent a nationally representative sample of discharge records from 1986 and 1987 to an expert coding group (called the "SuperPRO") which regularly reviews samples of discharges to enforce coding accuracy. They find that one-third of the case-mix increase was due to upcoding, although the standard error of this estimate is large. More recently, Bruce M. Psaty et al. (1999) used detailed chart review to estimate that upcoding is responsible for over one-third of admissions assigned to the heart failure DRG (DRG 127).

Most of the nonmedical analyses of case-mix increases (e.g., Steinwald and Dummit, 1989) are descriptive, focusing on which types of hospitals exhibit faster case-mix growth (large, urban, and teaching hospitals), and when these increases occur (there is a big jump in the first year a hospital is paid under PPS). Because these studies use data from the transition to PPS, the results are difficult to interpret; patient severity changed dramatically due to changes in patient composition following the implementation of PPS.

A recent study by Elaine M. Silverman and Jonathan S. Skinner (2004) presents strong evidence of post-transition-era upcoding for pneumonia and respiratory infections between 1989 and 1996. Focusing on the share of patients with these diagnoses who are assigned to the most expensive DRG possible, Silverman and Skinner document large increases in upcoding, despite a downward trend in mortality rates. The authors find that for-profit hospitals upcode the most, and not-for-profit hospitals are more likely to engage in upcoding when area market share of for-profit hospitals is higher. This finding suggests that practices of competitors may affect upcoding indirectly through pressure on hospital profits, or directly via the dissemination of upcoding practices.[11] Silverman and Skinner

"underpaid" and "needs the funds to provide care for the uninsured."

[10] Agencies known as "Peer Review Organizations" regularly audit DRG assignments. CMS works with the Office of the Inspector General, the Federal Bureau of Investigations, and the U.S. Attorney's Office to levy fines, recover funds, and prosecute providers who defraud the Medicare program. The *qui tam* provision of the Federal Civil False Claims Act protects and rewards whistle-blowers.

[11] Several recent studies document this indirect channel, e.g., Mark G. Duggan (2002), who finds that not-for-profit hospitals respond more strongly to financial incentives to

also find that hospitals under financial distress upcode *less* than financially sound institutions.

The upcoding analysis I perform improves upon prior research in three ways. First, I exploit a discrete change in upcoding incentives, which enables me to cleanly separate upcoding from time trends that may be associated with omitted factors such as the severity of patients' conditions. Second, I exploit differences across diagnoses in the magnitude of the change in upcoding incentives to see whether hospitals respond to upcoding incentives *on the margin,* upcoding even more when the payoff is greater. Third, I expand the scale of previous work by analyzing nearly 7 million Medicare admissions to 186 DRGs and 5,352 hospitals nationwide.

In addition to estimating upcoding responses for all hospitals financed under PPS, I estimate responses separately for different subsamples of hospitals. Due to heterogeneity in objective functions and endowments, hospitals may respond differently to the same price changes. For example, hospitals placing a higher weight on profits should upcode more, while hospitals facing a greater penalty (real or perceived, monetary or otherwise) or a higher probability of detection should upcode less. All things equal, hospitals experiencing financial distress should be more willing to risk detection. Size may also enter into upcoding decisions, as larger hospitals can spread the costs of training their coding personnel across more admissions. Finally, there are important regional differences in hospital behavior, although there are few theoretical explanations for this phenomenon apart from "cultural norms." In Section IV B, I stratify the hospital sample by a variety of these characteristics in order to explore potential differences in responses.

Real Responses.—Because Medicare patients face the same out-of-pocket costs for covered stays at all hospitals, they are free to select their preferred provider, and may consider such factors as location, quality of care, and physician recommendations.[12] In order to attract patients in diagnoses subject to price increases, hospitals may increase the intensity or quality of treatment (hereafter "intensity") provided to such patients. In Section V, I estimate the elasticity of intensity with respect to price. I instrument for price using price changes that were exogenously or "mechanically" imposed by the new policy, independently of hospitals' upcoding responses.

The few studies that investigate the effect of DRG-level price changes on intensity levels all find a positive relationship, where intensity is measured by length of stay, number of surgical procedures, and/or death rates (Cutler, 1990, 1995; Boyd H. Gilman, 2000).[13] Thus, all evidence to date suggests that a "flypaper effect" operates in the hospital industry: additional income is allocated to the clinical area in which it is earned, rather than spread across a broad range of activities. However, because all of these studies utilize data from a transition to a prospective payment system, they face the formidable challenge of separating two simultaneous changes in incentives: the elimination of marginal reimbursement, and changes in the average *level* of payments for each DRG.

Cutler (1990) examines the transition to PPS in Massachusetts, finding that length of stay and number of procedures per patient declined the most in DRGs subject to the largest price reductions. Despite finding an elasticity of intensity with respect to price of 0.2, Cutler does not find corresponding increases in volume. Cutler (1995) studies the impact of PPS on adverse medical outcomes, again finding an intensity response: reductions in average price levels are associated with a compression of mortality rates into the immediate post-discharge period, although there is no change in mortality at one year post-discharge. Both papers assume that eliminating the *marginal* reimbursement incentive affected all DRGs equally, when in fact there was substantial variation in the degree of "fat to trim." To the extent that price reductions

treat indigent patients in markets with greater for-profit penetration.

[12] Co-payments may vary across hospitals for Medicare beneficiaries enrolled in HMOs. However, fewer than 3 percent of Medicare beneficiaries were enrolled in HMOs during the study period.

[13] Most prior research on real responses examines the effect of hospital-wide financial shocks on overall quality, as measured by the average length of stay and inpatient mortality rates. These studies, which generally exploit Medicare shocks to hospital finances, find a positive effect of reimbursement on quality (e.g., Jack Hadley et al., 1989; Douglas Staiger and Gaumer, 1992).

were correlated with this excess (the very goal of the price-setting process), the intensity responses to price changes will be overstated. More generally, the elasticity estimate will be biased by any omitted factor influencing both price and intensity changes during the transition to PPS.[14] The same critique pertains to Gilman (2000), who investigates the impact of a 1994 reform to Medicaid DRGs for HIV diagnoses in New York.

By investigating responses to changes in average payment levels (i.e., prices) in the post-implementation period, I circumvent both this challenge and the concern that transitory responses are driving previous results. I also explore the effect of price on DRG volume. If hospitals increase intensity and patient demand is elastic with respect to this intensity, admissions volumes should rise in DRGs with price increases. For example, if the price for prostatectomy rises and hospitals invest in nerve-sparing surgical techniques, the number of patients electing to undergo this procedure is likely to increase. Alternatively, hospitals may offer free screening tests for prostate cancer and increase demand through the detection of new cases (a common practice).[15]

As with upcoding responses, there are many reasons that real responses may differ across hospitals. For example, for-profit hospitals may respond more to the financial incentive to attract patients in lucrative DRGs. Differences across DRGs are another possible source of variation in responses. Patient demand for planned or elective admissions may be more sensitive to changes in intensity than demand for urgent care. When a hospitalization is anticipated, a patient can "shop around," soliciting advice and information directly from the hospital, as well as from physicians and friends. The elasticity of demand with respect to quality might therefore be larger for such admissions, raising hospitals' incentives to increase quality in the face of price increases. I explore differences in intensity and volume responses across hospitals and admission types in Section V B.

II. A Price Shock: The Elimination of the Age Criterion

Although there were 473 individual DRG codes in 1987, 40 percent of these codes belonged to a "pair" of codes that shared the same main diagnosis. Within each pair, the codes were distinguished by age restrictions and presence of complications (CC). For example, the description for DRG 138 was "cardiac arrhythmia and conduction disorders age > 69 and/or CC," while that for DRG 139 was "cardiac arrhythmia and conduction disorders without CC." Accordingly, the DRG weight for the top code in each pair exceeded that for the bottom code. There were 95 such pairs of codes, and 283 "single" codes.

In 1987, separate analyses by HCFA and the Prospective Payment Assessment Commission (ProPAC) revealed that "in all but a few cases, grouping patients who are over 69 with the CC patients is inappropriate" (52 *Federal Register* 18877).[16] The ProPAC analysis found that hospital charges for uncomplicated patients over 69 were only 4 percent higher than for uncomplicated patients under 70, while average charges for patients with a CC were 30 percent higher than for patients without a CC. In order to minimize the variation in resource intensity within DRGs and to reimburse hospitals more accurately for these diagnoses, HCFA elimi-

[14] Cutler's methodology for calculating the change in average payment incentives following the implementation of PPS likely yields upward-biased elasticity estimates. Cutler defines the change in average price as the difference between the 1988 PPS price and the price that Medicare would have paid in 1988 were cost-plus reimbursement still in effect. To estimate this latter figure, he inflates 1984 costs for each DRG by the overall cost-growth rate for 55- to 64-year-olds. However, DRGs with disproportionately stronger cost growth between 1984 and 1988 received weight increases, yielding higher 1988 PPS prices and generating the concern that the positive relationship between price changes and intensity levels may be spurious. The possibility that these estimated price changes are not exogenous is reinforced by the use of hospital-specific prices in the specifications. The average price changes are therefore related to hospitals' pre-PPS DRG-specific costs; hospitals with high costs faced price reductions when transitioning to national payment standards. These hospitals may have had "more fat to trim" in terms of intensity provision.

[15] To my knowledge, there are no prior studies that examine the effect of DRG prices on volume. There is compelling evidence, however, that inpatient utilization in general increases with insurance coverage, which affects both out-of-pocket expenses and hospital reimbursements (see Frank R. Lichtenberg, 2002, on the effects of Medicare and Dafny and Jonathan Gruber, 2005, on Medicaid).

[16] ProPAC, now incorporated into MedPAC (Medicare Payment Advisory Commission), was an independent federal agency that reported to Congress on all PPS matters.

TABLE 1—EXAMPLES OF POLICY CHANGE

DRG code	Description in 1987 (Description in 1988)	1987 weight	1988 weight	Percent change in weight	1987 volume (20-percent sample)	1988 volume (20-percent sample)	Percent change in volume
96	Bronchitis and asthma age > 69 and/or CC (bronchitis and asthma age > 17 with CC)	0.8446	0.9804	16%	44,989	42,314	−6%
97	Bronchitis and asthma age 18–69 without CC (bronchitis and asthma age > 17 without CC)	0.7091	0.7151	1%	4,611	10,512	128%
138	Cardiac arrhythmia and conduction disorders age > 69 and/or CC (cardiac arrhythmia and conduction disorders with CC)	0.8136	0.8535	5%	45,080	35,233	−22%
139	Cardiac arrhythmia and conduction disorders age < 70 without CC (cardiac arrhythmia and conduction disorders without CC)	0.6514	0.5912	−9%	4,182	16,829	302%
296	Nutritional and misc. metabolic disorders age > 69 and/or CC (nutritional and misc. metabolic disorders age > 17 with CC)	0.8271	0.9259	12%	45,903	38,805	−15%
297	Nutritional and misc. metabolic disorders age 18–69 without CC (nutritional and misc. metabolic disorders age > 17 without CC)	0.6984	0.5791	−17%	2,033	12,363	508%

Notes: Of the 95 DRG pairs, these three occur most frequently in the 1987 20-percent MedPAR sample. DRG weights are from the *Federal Register.*

nated the age over 69/under 70 criterion beginning in fiscal year (FY) 1988.[17] The agency recalibrated the weights for all DRGs to reflect the new classification system. This recalibration resulted in large increases in the weights for top codes within DRG pairs, and moderate declines for bottom codes.

Table 1 lists the three most commonly coded pairs and their DRG weights before and after the policy change.[18] These examples are fairly representative of the change overall. Using 1987 admissions from a 20-percent sample of Medicare discharge data as weights, the weighted average increase in the top code for all DRG pairs was 11.3 percent, while the weighted average decrease in the bottom code was 6.2 percent. This increase in the "spread" between the weights for the top and bottom codes in each pair strengthened the incentive for hospitals to code complications on patients' records. In Section IV, I exploit variation across DRG pairs in the magnitude of these increases to examine whether hospitals responded to differences in upcoding incentives.

To estimate the elasticity of intensity with respect to price, I exploit recalibration errors made by HCFA when adjusting the weights to reflect the new classifications. As the analysis in Section V reveals, even if hospitals had not reported higher complication rates following the policy change, the average price for admissions to DRG pairs would have risen by at least 2.6 percent on average. This "mechanical effect" varied across DRG pairs as well, ranging from a reduction of $1,232 to an increase of $3,090 *per admission.* I investigate whether hospitals adjusted their intensity of care and admissions volumes in response to these exogenous price changes.

HCFA subsequently acknowledged that mistakes in their recalibrations had led to higher DRG weights. In 1989, the agency published an (unfortunately flawed) estimate of the contribution of recalibration mistakes to the large increase in average DRG weight between 1986 and 1988 (54 *Federal Register* 169). HCFA concluded that 0.93 percentage points could be attributed to faulty recalibration of DRG weights for 1988, and an additional 0.29 percentage points to errors in 1987.[19] These

[17] HCFA's fiscal year begins in October of the preceding calendar year.

[18] The large volume increase for the bottom code in each pair is due to the new requirement that uncomplicated patients over 69 be switched from the top to the bottom code.

[19] The original notice for the policy change states that the goal of the recalibration is to ensure no overall

estimates motivated an across-the-board reduction of 1.22 percent in all DRG weights beginning in 1990. Because this reduction applied uniformly to all DRGs, the relative effects across DRG pairs were unabated.

III. Data

My primary data sources are the 20-percent Medicare Provider Analysis and Review (MedPAR) files (FY85–FY91), the annual tables of DRG weights published in the *Federal Register* (FY85–FY91), the Medicare Cost Reports (FY85–FY91), and the Annual Survey of Hospitals by the American Hospital Association (1987). The MedPAR files contain data on all hospitalizations of Medicare enrollees, including select patient demographics, DRG code, measures of intensity of care (e.g., length of stay and number of surgeries), and hospital identification number.[20] The data span the three years before and after the policy change.

The MedPAR discharge records are matched to DRG weights from the *Federal Register* and hospital characteristics from the Annual Survey of Hospitals and the Medicare Cost Reports for 1987, the year preceding the policy change.[21] Due to the poor quality of hospital financial data, the debt:asset ratio from the Medicare Cost Reports is among the best measures of financial distress. I also construct two additional financial distress measures, Medicare "bite" (the fraction of a hospital's discharges reimbursed by Medicare) and Medicaid "bite" (similarly defined). Table A1 in the Appendix presents descriptive statistics for these measures, together with other hospital characteristics that may be associated with responses to the shock (ownership status, region, teaching status, number of beds, and service offerings). Because price varies at the hospital and DRG level, the individual discharge records are aggregated to form DRG-year or hospital-year cells. Descriptive statistics for these cells are reported in Table 2.

IV. The Nominal Response: More DRG Creep

Although DRG creep was known to be a pervasive problem by 1987, HCFA's policy change nevertheless increased the reward for upcoding. The increase in prices for the top codes in DRG pairs, together with the decrease in prices for the bottom codes, provided a strong incentive to continue using the top code for all older patients (not just those with CC), and to use it more frequently for younger patients. Figure 1 charts the overall fraction of patients in DRG pairs assigned to top codes between 1985 and 1991, broken down by age group. The share of older patients in top codes falls sharply, from nearly 100 percent in 1985–1987 to 70 percent in 1988, and creeps steadily upward thereafter. The trend in complications between 1988 and 1991 exactly matches that for young patients, whose complication rate increases steadily throughout the study period, with the exception of a plateau between 1987 and 1988.

Figure 1 suggests that time-series identification is insufficient for examining the upcoding response to the policy change. While it is possible that the increase in the share of young patients coded with complications between 1988 and 1991 is a lagged response to the policy change, it could also be a continuation of a preexisting trend. For older patients, the data do not reveal what share of admissions were coded with complications prior to the policy change, so it is impossible to discern whether there was an abrupt increase in the reported complication rate. Furthermore, upcoding among the old is likely to be even greater than upcoding among the young, as hospitals with sharp increases in the complication rate of older patients can argue that they failed to code these complications in the pre-1988 era because there was no payoff for doing so.[22] Fortunately, differences across

change in reimbursement to hospitals; that is, the average national DRG weight should have been constant whether the 1987 or the 1988 classification system (called the GROUPER program) was employed on a given set of discharge records.

[20] The data include all categories of eligibles, but exclude admissions to enrollees in Medicare HMOs (see footnote 8).

[21] The Cost Reports also contain an indicator for whether a hospital is paid under the PPS system. I omit exempt hospitals from my sample. For hospitals with missing AHA data in 1987, I use data from the nearest year possible.

[22] HCFA's audit agencies should have had access to the complication rates for older patients throughout the study period; simple manipulations of the GROUPER programs from 1985–1987 could produce these rates in the pre-period. Unfortunately, these programs have reportedly been erased from HCFA's records.

TABLE 2—DESCRIPTIVE STATISTICS

	DRG pair-year			Hospital-year		
Unit of observation	*N*	Mean	*SD*	*N*	Mean	*SD*
Price (DRG weight)	650	1.12	0.62	36651	1.27	(0.19)
Admissions per cell	650	10624	(15013)	36651	373	(389)
Nominal responses						
Fraction(young) in top code	650	0.66	(0.14)			
Fraction(old) in top code	650	0.85	(0.15)			
Real responses						
Mean cost ($)	650	9489	(6230)	36169	12272	(5692)
Mean LOS (days)	650	9.37	(3.32)	36651	8.81	(2.21)
Mean surgeries	650	1.15	(0.69)	35897	1.21	(0.55)
Mean ICU days	650	0.51	(0.65)	28226	0.81	(0.59)
Death rate	650	0.06	(0.06)	34992	0.06	(0.02)
Mean admissions	650	31806	(25822)	36651	778	(538)
Instruments						
Δspread	650	0.20	(0.16)			
Δspread · post	650	0.12	(0.16)			
Δ ln(Laspeyres price)	650	0.03	(0.06)			
Δ ln(Laspeyres price) · post	650	0.01	(0.05)			
Share CC				36651	0.09	(0.03)
Share CC · post				36651	0.05	(0.05)

Notes: The table reports weighted means and standard deviations for DRG pair-year cells and hospital-year cells, which are constructed by aggregating individual-level data in the 20-percent MedPAR sample. Cells are weighted by the number of individual admissions in the 20-percent MedPAR sample, with the exception of admissions per cell. Costs for all years are converted to 2001 dollars using the CPI for hospital services.

DRG pairs in the policy-induced incentive to upcode, as measured by changes in "spread," enable me to calculate lower-bound estimates of the upcoding response in both age groups.

A. *Aggregate Analysis*

The dependent variable for this analysis is $fraction_{pt}$, the share of admissions to pair p in year t that is assigned to the top code in that pair. The independent variable of interest, $spread_{pt}$, is defined as

$$(2)\quad spread_{pt} = DRG\ weight\ in\ top\ code_{pt} - DRG\ weight\ in\ bottom\ code_{pt},$$

e.g., $\text{spread}_{\text{DRG } 138/139,1988} = \text{weight}_{\text{DRG } 138,1988} - \text{weight}_{\text{DRG } 139,1988} = 0.8535 - 0.5912 = 0.2623$. $spread_{pt}$ is simply a measure of the upcoding incentive in pair p at time t. Between 1987 and 1988, mean spread increased by 0.20, approximately $875.[23] However, there was substantial variation in spread changes across pairs, as demonstrated by the frequency distribution of $\Delta spread_{p,88\text{-}87}$ in Figure 2. Due to the recalibration procedure, the largest spread increases occurred in DRG pairs in which complications are particularly costly to treat and/or in which the share of older, uncomplicated patients is high.

To examine the effect of the change in spread on *fraction*, I estimate the specification

$$(3)\quad fraction_{pt} = \alpha + \boldsymbol{\varsigma} pair_p + \boldsymbol{\delta} year_t + \psi \Delta spread_{p,88\text{-}87} \cdot post + \varepsilon_{pt}$$

where p indexes DRG pairs, t indexes years, *post* is an indicator for the years following the policy change (1988–1991), and the dimensions of the coefficient vectors are $\boldsymbol{\varsigma}$ (1×93), $\boldsymbol{\delta}$ (1×6), and ψ (1×1).[24] $\boldsymbol{\delta}$ captures the *average* impact of the policy reform on all pairs, while ψ captures the *marginal* effect of differential upcoding incentives. $\hat{\psi} > 0$ signifies that hospitals

[23] This dollar amount is based on P_h for urban hospitals in 2001, which was $4,376.

[24] Of the 95 DRG pairs in 1991, two pairs are dropped because the age criterion was eliminated one year early for these pairs.

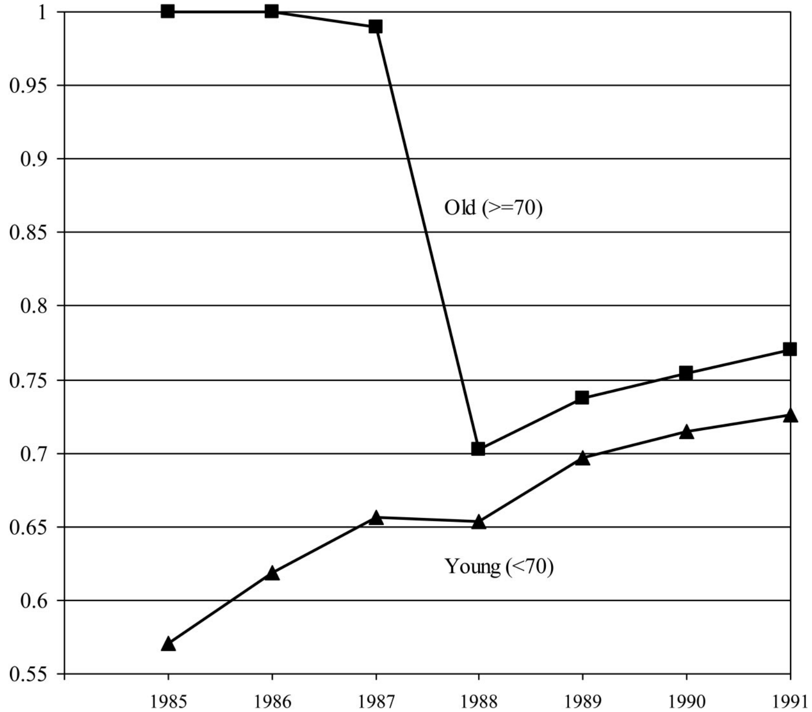

FIGURE 1. FRACTION OF ADMISSIONS IN TOP CODES BY AGE GROUP

Notes: Sample includes all admissions to DRG pairs in hospitals financed under PPS. Author's tabulations from the 20-percent MedPAR sample.

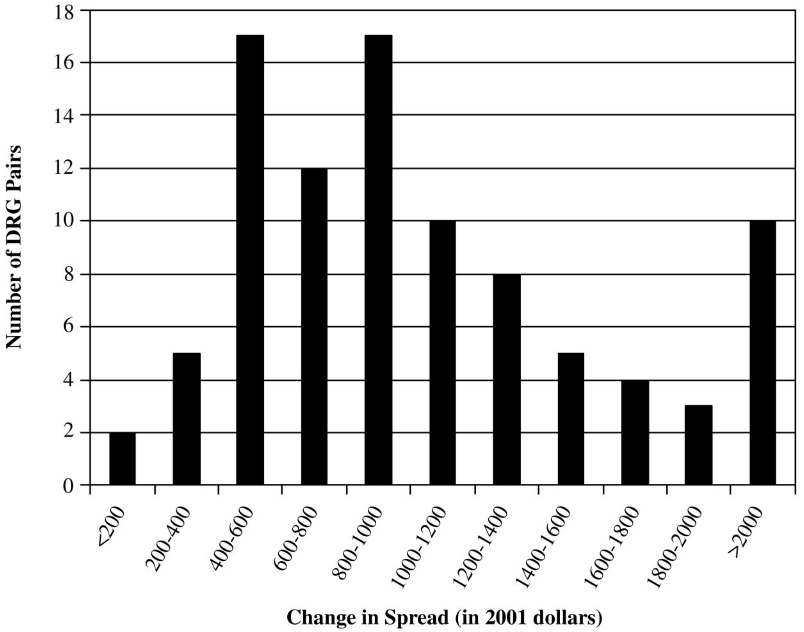

FIGURE 2. DISTRIBUTION OF CHANGES IN SPREAD, 1987–1988

Notes: Spread is the difference in DRG weights for the top and bottom codes in a DRG pair. Changes in spread are converted to 2001 dollars using the base price for urban hospitals in 2001 (=$4,376).

upcoded more in pairs where the incentive to do so increased more.[25]

The estimation results for equation (3) are reported separately by age group in Table 3. In all specifications, observations are weighted by the number of admissions in their respective pair-year cells. Using grouped data produces conservative standard errors, which I further correct for heteroskedasticity. For older patients, I include $fraction(young)_{p,87} \cdot post$ as an estimate of the underlying complication rate in each DRG pair, where $fraction(young)_{pt}$ refers to the fraction of young patients assigned to the top code in pair p and year t. This seems a reasonable baseline assumption, given a correlation coefficient of 0.94 for young and old complication rates in the post period.

The coefficient estimates reveal that upcoding is sensitive to changes in spread, even after controlling (via the year dummies) for the large average increase in spread between 1987 and 1988. Thus, the decline in *fraction* for old patients was least in those pairs subject to the largest increase in spread, while the increase in *fraction* for young patients was greatest in those pairs subject to the largest increase in spread. As hypothesized, the upcoding response was greater for older patients: the coefficient estimates imply a spread-induced increase of 0.022 in the fraction of old patients coded with CC, as compared to 0.015 for younger patients.

Figures 3A and 3B illustrate these responses for young and old patients, respectively, by graphing *fraction* for pairs in the top and bottom quartiles of spread changes. Interestingly, the aggregate plateau in fraction for young patients between 1987 and 1988 appears to be comprised of an increase in *fraction* for those codes with large spread increases and a decrease in *fraction* for those codes with small spread increases. A plausible explanation is that hospitals were concerned that a rise in complications for *all* young patients would attract the attention of regulators, so they chose to upcode in those diagnoses with the greatest payoff and downcode in those with the least payoff. Figure 3B shows a similar response for older patients: the post-shock fraction was higher in the top quartile of spread increases, even though the pre-shock fraction for young patients in these DRGs (the "baseline") was lower.

TABLE 3—EFFECT OF POLICY CHANGE ON UPCODING (N = 650)

	Fraction(young) mean = 0.66	Fraction(old) mean = 0.85
Δspread$_{88\text{-}87}$ · post	0.077***	0.108***
	(0.016)	(0.015)
Fraction(young)$_{87}$ · post		0.731***
		(0.020)
Year dummies		
1986	0.044***	0.000
	(0.008)	(0.005)
1987	0.077***	−0.011*
	(0.008)	(0.005)
1988	0.058***	−0.813***
	(0.011)	(0.014)
1989	0.097***	−0.780***
	(0.009)	(0.014)
1990	0.115***	−0.764***
	(0.009)	(0.014)
1991	0.128***	−0.748***
	(0.010)	(0.014)
Adj. R-squared	0.948	0.960

Notes: The table reports estimates of the effect of changes in spread between 1987 and 1988 on the fraction of patients coded with complications (equation 3 in the text). "Post" is an indicator for the post-1987 period, "young" refers to Medicare beneficiaries under 70, and "old" refers to beneficiaries aged 70+. Regressions include fixed effects for DRG pairs. The weighted mean of the dependent variable is reported at the top of each column. The unit of observation is the DRG pair-year. All observations are weighted by the number of admissions in the 20-percent MedPAR sample. The sum of the weights is 1.9 million (young) and 5.0 million (old). Robust standard errors are reported in parentheses. * Signifies $p < 0.05$; ** signifies $p < 0.01$; *** signifies $p < 0.001$.

The main threat to the validity of this analysis is the possibility that changes in spread are correlated with omitted factors which may be driving the observed changes in fraction, i.e., that the changes in spread are not exogenous. To help rule out this possibility, I regress the change in $fraction_{pt}$ for young patients between 1985–1986, 1986–1987, and 1987–1988 on $\Delta spread_{d,88\text{-}87}$ and DRG fixed effects. I find coefficients (robust standard errors) of 0.002 (0.013), −0.017 (0.015), and 0.084 (0.034), respectively. These results indicate that the changes in spread between 1987 and 1988 are not correlated with preexisting trends in $fraction_{pt}$. As another robustness check, I reestimated equation (3) including individual DRG trends. The coefficient estimates on spread

[25] An alternative specification using $\Delta spread_{p,88\text{-}87} \cdot post$ as an instrument for $spread_{pt}$ yields similar results.

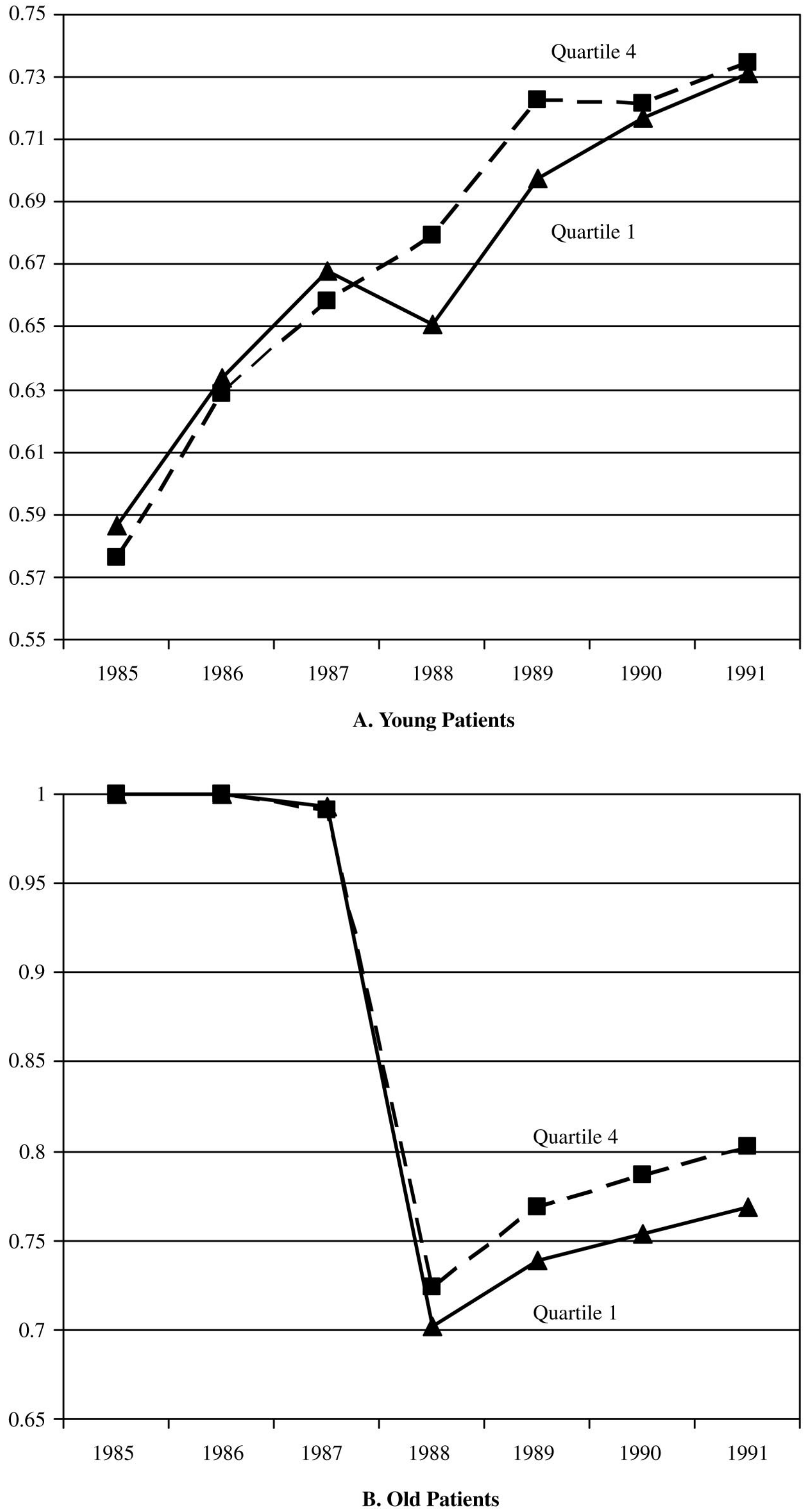

FIGURE 3. FRACTION OF ADMISSIONS IN TOP CODES BY AGE GROUP AND QUARTILE OF SPREAD CHANGE

Notes: DRG pairs experiencing spread changes under $598 are in quartile 1; DRG pairs experiencing spread changes over $1,375 are in quartile 4.

change little, and both remain statistically significant with $p < 0.05$. Thus, the identifying assumption is that changes in fraction and spread are not correlated with an omitted factor that *first* appeared in 1988.

The spread-related upcoding alone translates into a price increase of 0.7 percent and 0.9 percent for young and old patients admitted to DRG pairs, respectively.[26] The estimate for young patients rises to 1.5 percent if the jump between 1987 and 1989 is included, although this is an upper-bound estimate due to the potential role of confounding factors. Given average operating margins of 1 to 2 percent in the hospital industry, these figures are large. The estimates imply increased *annual* payments of \$330 to \$425 million.[27] This range is very conservative, as it excludes the effect of any average increase in upcoding of older patients across all DRG pairs, and older patients account for 70 percent of Medicare admissions.

HCFA's 1990 across-the-board reduction in DRG weights decreased annual payments by \$1.13 billion. However, while this reduction affected all hospitals equally, the rewards from upcoding accrued only to those hospitals engaging in it. The following section investigates the relationship between hospital characteristics and upcoding responses.

B. *Hospital Analysis*

To determine whether individual hospitals responded differently to the changes in upcoding incentives, I estimate equation (3) separately for subsets of hospitals. For example, I compare the results obtained using data solely from teaching hospitals with those obtained using the sample of nonteaching hospitals. I also consider stratifications by ownership type (for-profit, not-for-profit, government), financial status, region, size, and market-level Herfindahl index.[28] Hospitals with missing data for any of these characteristics are omitted from this analysis.

Table 4 presents estimates of $\boldsymbol{\delta}$ and ψ by hospital ownership type, financial status, and region. Figure 4 plots the $\hat{\boldsymbol{\delta}}$ from these specifications. For both young and old patients, there are no statistically significant differences in the response to Δ*spread* across the hospital groups. The discussion here, therefore, focuses on the results for young patients, for whom the year coefficients are relevant.

The main finding is that for-profit hospitals upcoded more than government or not-for-profit hospitals following the 1988 reform. Figure 4A illustrates that upcoding trends were the same for all three ownership types until 1987, but thereafter the trend for for-profits diverged substantially. By 1991, the fraction of young patients with complications had risen by 0.18 in for-profit hospitals, compared with ~0.13 for the other two groups. Given a universal mean of 0.65 in 1987, these figures are extremely large. For-profits' choice to *globally* code more patients with complications, rather than to manifest greater sensitivity to spread, is consistent with the reward system implemented by the nation's largest for-profit hospital chain, Columbia/HCA (now HCA). A former employee reported that hospital managers were specifically rewarded for upcoding patients into the more-remunerative "with complications" codes (Lucette Lagnado, 1997).

Hospitals with high debt:asset ratios (Figure 4B) and hospitals in the South (Figure 4C) also exhibited very large increases in *fraction*(*young*), although the graphs illustrate that these trends predated the policy change. Moreover, the strong presence of for-profits in the South and the tendency of for-profits to be highly leveraged suggests that for-profit ownership is driving the large fraction gains in these subsamples as well. All other hospital characteristics were not associated with changes in upcoding proclivity.

V. Real Responses

To investigate real responses to price increases, I create a dataset of mean intensity levels and price for each DRG pair and year

[26] These estimates are calculated using the average spread in 1988 (0.45), together with the average weights for DRG pairs in 1987 (1.05 for young patients, 1.13 for older patients).

[27] Dollar figures are calculated using PPS expenditures in 2000.

[28] The Herfindahl index is calculated as the sum of squared market shares for all hospitals within a health service area, as defined by the National Health Planning and Resources Development Act of 1974 (and reported in the AHA data).

TABLE 4—EFFECT OF POLICY CHANGE ON UPCODING OF YOUNG, BY HOSPITAL CHARACTERISTICS
($N = 650$)

	By ownership type			By financial state		By region			
	For-profit	Not-for-profit	Government	Distressed	Not distressed	Northeast	Midwest	South	West
$\Delta\text{spread}_{88\text{-}87} \cdot \text{post}$	0.071**	0.080***	0.058***	0.082***	0.074***	0.083***	0.062***	0.082***	0.079***
	(0.024)	(0.017)	(0.018)	(0.109)	(0.017)	(0.016)	(0.018)	(0.017)	(0.024)
Year fixed effects									
1986	0.038***	0.047***	0.040***	0.059***	0.041***	0.095***	0.027**	0.033***	0.035***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.012)
1987	0.081***	0.077***	0.078***	0.099***	0.073***	0.123***	0.054***	0.078***	0.052***
	(0.010)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.009)	(0.008)	(0.012)
1988	0.080***	0.055***	0.065***	0.083***	0.054***	0.104***	0.036***	0.063***	0.024***
	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.010)	(0.010)	(0.012)	(0.014)
1989	0.140***	0.094***	0.104***	0.127***	0.094***	0.131***	0.075***	0.111***	0.067***
	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.012)
1990	0.147***	0.114***	0.120***	0.144***	0.112***	0.148***	0.092***	0.132***	0.080***
	(0.011)	(0.009)	(0.010)	(0.009)	(0.010)	(0.008)	(0.009)	(0.010)	(0.013)
1991	0.179***	0.123***	0.136***	0.161***	0.124***	0.159***	0.103***	0.148***	0.091***
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.014)
$\hat{\delta}_{89} - \hat{\delta}_{87}$	0.059***	0.016	0.027***	0.027**	0.022*	0.007	0.021*	0.033***	0.015
	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)	(0.010)	(0.008)	(0.014)
Adj. R-squared	0.914	0.946	0.927	0.933	0.947	0.939	0.940	0.940	0.915

Notes: The table reports estimates of equation (3) in the text separately for subsets of hospitals. The specification is analogous to that used in column 1, Table 3. Regressions include fixed effects for DRG pairs. "Young" refers to Medicare beneficiaries under 70. "Distressed" denotes hospitals with 1987 debt:asset ratios at the seventy-fifth percentile or above. The unit of observation is the DRG pair-year. All observations are weighted by the number of admissions in the 20-percent MedPAR sample. Hospitals with missing values for any of the hospital characteristics are dropped. The sum of the weights is 1.45 million. Robust standard errors are reported in parentheses. * Signifies $p < 0.05$; ** signifies $p < 0.01$; *** signifies $p < 0.001$.

($N = 650$). Descriptive statistics for these variables are in Table 2. Although the elimination of the age criterion resulted in large price changes for individual DRGs, it would not be informative to investigate whether intensity levels rose (fell) for patients admitted to the top (bottom) code of DRG pairs, because the composition of patients admitted to each code changed as a result of the policy reform. Top codes, which were formerly assigned to all older patients as well as to young patients with CC, are now intended to be used exclusively for patients with CC, young or old. A finding that average intensity of care increased in top codes would not reveal whether hospitals increased intensity of care for patients with CC, the only patients for whom a price increase was enacted. Furthermore, policy-induced upcoding from bottom to top codes exacerbates the problem of compositional changes within each DRG code.[29] I therefore combine data from the top and bottom codes in order to keep the reference population constant before and after the policy reform.

To instrument for price, I exploit differences across DRG pairs in the magnitude of recalibration mistakes made by HCFA. I isolate this mechanical effect of the policy change because hospitals may respond differently to exogenous price changes than to the endogenous price changes produced via upcoding. (Recall that pure upcoding implies no change in patient care.)

A. *Aggregate Analysis*

To estimate the mechanical effect of the policy change, I create a Laspeyres price index for 1988 using the 1987 volumes of young patients in each code as the weights, e.g.,

$$\text{(4)}\qquad \text{Laspeyres price}_{\text{DRG 138/139,1988}} = \frac{\text{price}_{\text{DRG 138,1988}} \cdot N_{\text{DRG 138,1987}} + \text{price}_{\text{DRG 139,1988}} \cdot N_{\text{DRG 139,1987}}}{N_{\text{DRG 138,1987}} + N_{\text{DRG 139,1987}}}.$$

This fixed-weight index approximates the average price hospitals would have earned for an

[29] If the sample were restricted to patients under 70, the first of these compositional problems would not apply. However, the second would bias any intensity response estimated using individual DRGs as the unit of observation.

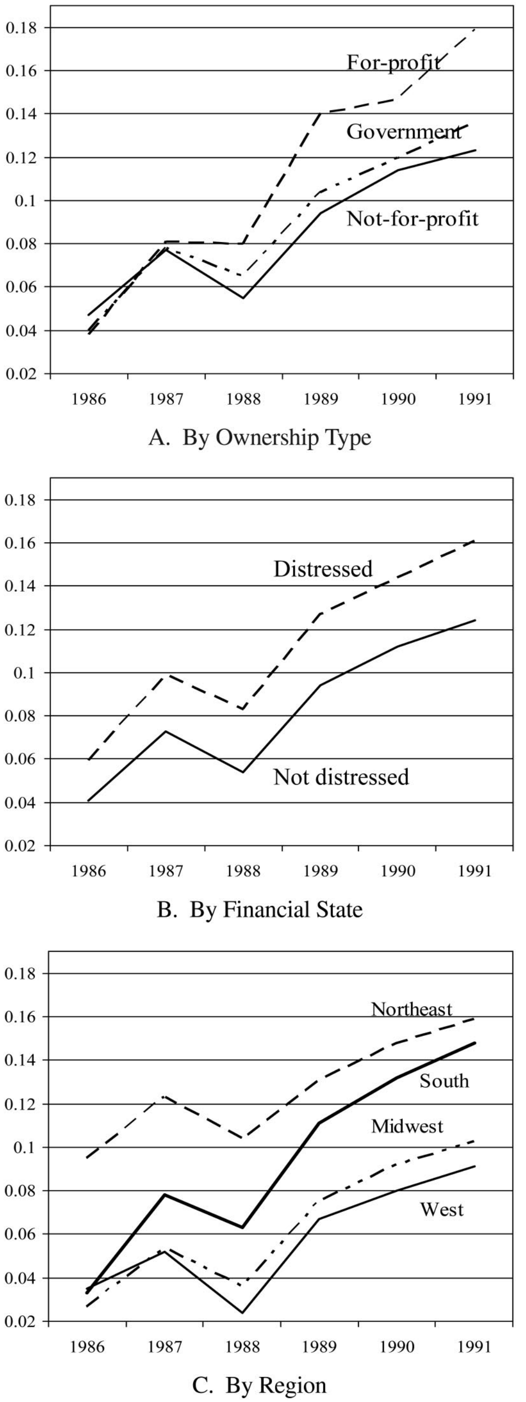

FIGURE 4. EFFECT OF POLICY CHANGE ON UPCODING OF YOUNG, BY HOSPITAL CHARACTERISTICS

Notes: The figures above plot the year coefficients from Table 4. "Young" refers to Medicare beneficiaries under 70.

admission to a given DRG pair in 1988 had the fraction of patients with CC remained constant at the 1987 fraction for young patients. (Because the fraction of old patients with CC cannot be ascertained in 1987, the fraction for young patients must proxy for this measure.) The mechanical effect of the policy change can then be estimated as the difference between the Laspeyres price in 1988 and the actual price in 1987. Figure 5 graphs the distribution of these changes, translated into 2001 dollars. The mean change is \$102 per admission, with a standard deviation of \$448 and an interquartile range of (\$131) to \$262. These figures are small relative to the average price per admission (\$4,376), but large relative to average operating margins for hospitals. Note also that this formula underestimates mechanical price changes because the share of old patients with CC is likely to be greater than that of young patients.

Assuming the measurement error in $\Delta ln(Laspeyres\ price)$ is small, its coefficient in the following first-stage regression should equal 1:

$$(5)\quad ln(price)_{pt} = \alpha + \boldsymbol{\varsigma} pair_p + \boldsymbol{\delta} year_t + \kappa_1 \Delta ln(Laspeyres\ price)_{p,88\text{-}87} \cdot post + \boldsymbol{\zeta} pair_p \cdot year + \varepsilon_{pt}.$$

$\hat{\kappa}_1$ may differ from 1 if upcoding or post-1988 price recalibrations are correlated with Δln(Laspeyres price). The inclusion of individual pair time trends should mitigate these potential biases, and indeed the estimate of $\kappa_1 = 0.925$ (0.093). As a robustness check, I subtract an estimate of the component of $\Delta ln(Laspeyres\ price)_{p,88\text{-}87}$ that is due to lagged cost growth.[30] The coefficient on this revised instrument is 1.120 (0.119).

In the second stage, I use five dependent variables to measure intensity: total costs (=total charges from MedPAR deflated by annual cost: charge ratios from the Cost Reports and converted to 2001 dollars using the hospital services CPI), length of stay, number of surgeries, number of ICU days, and in-hospital deaths.

[30] Because HCFA operates with a two-year data lag when recalibrating prices, I calculate this component using the estimated coefficients from $\Delta ln(Laspeyres\ price)_{p,88\text{-}87} = \alpha + \beta \Delta ln(price)_{p,86\text{-}85} + \varepsilon$.

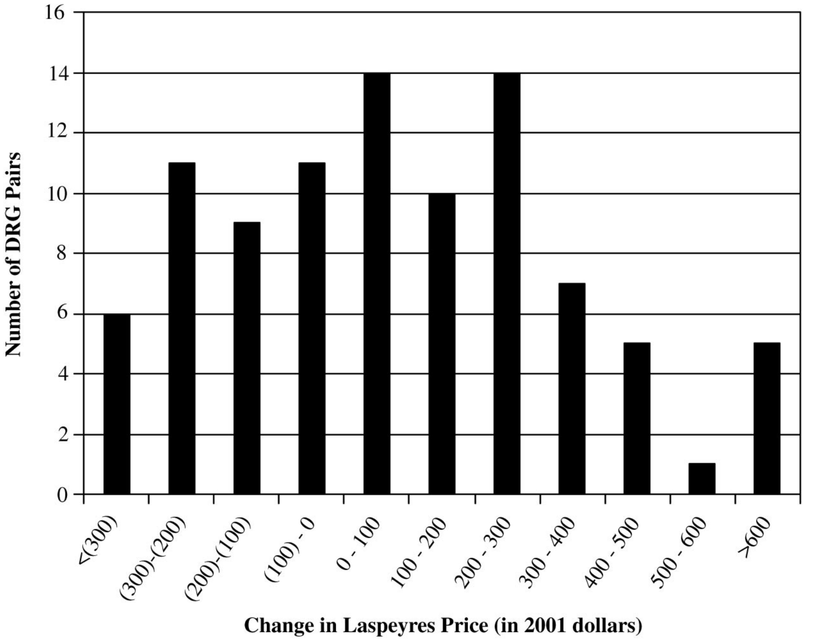

FIGURE 5. DISTRIBUTION OF CHANGES IN LASPEYRES PRICE, 1987–1988

Notes: The Laspeyres price is defined as the average DRG weight for a given pair and year, holding constant the fraction of patients coded with CC at the 1987 fraction for young patients in that DRG pair. "Young" refers to Medicare beneficiaries under age 70. Changes in Laspeyres price are converted to 2001 dollars using the base price for urban hospitals in 2001 (=$4,376).

All variables are normalized by the number of admissions in the relevant cell (i.e., average cost per patient in DRG pair 138/139 in 1987). The first four measures are strong indicators of hospital expenditures on behalf of patients.[31] Death rate is clearly an important, albeit limited, indicator of quality of care. Although these measures are commonly used in the health economics literature, they are imperfect. One of the most common measures, length of stay, could be correlated positively or negatively with quality of care. Better care may enable a patient to leave sooner; on the other hand, hospitals may discharge patients too early in order to cut costs. (The latter was of greater concern in the 1980s, as lengths of stay fell dramatically in response to PPS.) However, the consistency of the results across the variables suggests that the findings are robust.

Table 5 presents estimates of κ_2 from the reduced-form regressions

$$(6)\quad ln(intensity \text{ or } volume)_{pt} = \alpha + \boldsymbol{\varsigma} pair_p + \boldsymbol{\delta} year_t + \kappa_2 \Delta ln(Laspeyres\ price)_{p,88\text{-}87} \cdot post + \boldsymbol{\zeta} pair_p \cdot year + \varepsilon_{pt}.$$

This specification also controls for pair-specific trends in the dependent variable throughout the study period, ensuring that κ_2 does not reflect preexisting trends in intensity or volume. Table 5 offers little evidence that hospitals altered their treatment policies or increased their admissions differentially in DRG pairs subject to larger mechanical price changes. Two of the six point estimates indicate a *negative* response (costs and ICU days), and all are imprecisely

[31] Total charges deflated by hospital cost:charge ratios should be positively correlated with the services provided to patients; indeed, this is the measure HCFA uses to calculate DRG weights, so that diagnosis groups with higher average charges are reimbursed more than diagnosis groups with lower average charges.

TABLE 5—REAL RESPONSES TO CHANGES IN DRG PRICES
(N = 650)

	Dependent variable					
	ln(cost) mean = \$9,489	ln(LOS) mean = 9.37	ln(surgeries) mean = 1.15	ln(ICU days) mean = 0.51	ln(death rate) mean = 0.06	ln(volume) mean = 31,806
Reduced form						
Δ ln(Laspeyres price) · post	−0.207	0.073	0.009	−0.642	−0.381	0.245
	(0.133)	(0.146)	(0.158)	(0.380)	(0.352)	(0.272)
IV estimate						
ln(price)	−0.223	0.079	0.009	−0.694	−0.412	0.265
	(0.147)	(0.157)	(0.171)	(0.425)	(0.383)	(0.285)
Parametric tests of H_0: IV estimate $> =x$; H_1: IV estimate $< x$ (p-values are reported)[a]						
$x = 0.5$	0	0	0	0	0.41	0.21
$x = 1$	0	0	0	0	0.06	0.01
OLS estimate						
ln(price)	0.074	0.180***	−0.166*	0.106	0.151	N/A
	(0.053)	(0.049)	(0.068)	(0.136)	(0.134)	

Notes: The top panel of the table reports results from reduced-form regressions of ln(intensity) or ln(volume) on the change in ln(Laspeyres price) between 1987 and 1988 multiplied by an indicator for the post-1987 period ("post"). Intensity is measured by average costs, length of stay, number of surgeries, and the in-hospital death rate. The second and third panels report IV and OLS estimates, respectively, of the elasticity of DRG-pair intensity and DRG-pair volume with respect to DRG-pair price. All regressions include year fixed effects, DRG-pair fixed effects, and DRG-pair trends. The unlogged, weighted mean of the dependent variable is reported at the top of each column. The unit of observation is the DRG pair-year. All observations are weighted by the number of admissions in the 20-percent MedPAR sample. The sum of the weights is 6.9 million. Robust standard errors are reported in parentheses.

[a] For ln(death rate), the tests are H_0: IV estimate $< =x$; H_1: IV estimate $> x$ for $x = -0.5$ and $x = -1.0$.

* Signifies $p < 0.05$; ** signifies $p < 0.01$; *** signifies $p < 0.001$.

estimated. Given a precisely estimated first-stage coefficient of nearly 1, the corresponding IV estimates ($=\hat{\kappa}_2/\hat{\kappa}_1$) and standard errors are roughly the same. Due to the wide confidence intervals, intensity and volume responses cannot be ruled out, but the evidence for these responses is underwhelming. The results are not affected by the addition of a control variable for the severity of patients treated in each pair-year.[32]

To obtain upper bounds for the intensity elasticities, I estimated OLS regressions of ln(intensity) on ln(price), pair and year fixed effects, and pair-specific trends. These estimated elasticities, also reported in Table 5, are upward-biased due to the price recalibration method.[33] Notwithstanding this bias, the point estimates are extremely small. For example, the OLS estimate indicates that only 7 cents of every additional dollar of reimbursement within a DRG is spent on care for patients in that DRG. The elasticity of length of stay with respect to price (0.18) is similar to the estimate reported in Cutler (1990) (0.23), but the elasticity of surgeries is much smaller (−0.17 as compared to 0.23). Overall, Table 5 suggests that the flypaper effect is weak in this sector.

B. *Responses by Hospital and DRG Type*

The aggregate analysis captures the average intensity and volume responses across all admissions, but masks potentially different responses across hospitals. To determine whether individual hospitals responded differently, I estimate equation (6) separately for the various hospital subsamples. Due to the large volume of coefficients generated by these models, tables are not included here. Out of 126 regressions (6 dependent variables $*$ 21 subsamples), $\hat{\kappa}_2$ is statistically significant at $p < 0.05$ only once, and in this case it indicates a *negative* impact of price on mortality in hospitals with fewer than 100 beds.[34]

The point estimates suggest that for-profit hospitals decreased costs and length of stay

[32] To estimate patient severity, I computed the Charlson comorbidity index for each admission and calculated pair-year means. This index reflects the likelihood of one-year mortality, based on 19 categories of comorbidity that are captured in the diagnosis codes on each patient's record.

[33] One manifestation of this bias is the positive estimated elasticity of death rate with respect to price; the explanation for this paradoxical result is simply that DRG pairs that experience increases in death rates receive higher reimbursements because in-hospital care for the dying is costly.

[34] Tables are available upon request.

more than hospitals of other ownership types, although an F-test does not reject equality of $\hat{\kappa}_2$ across the groups. This pattern, together with a larger estimated elasticity of volume with respect to price [0.418 (0.318)], is consistent with the extensive upcoding by for-profits documented in the previous section. Financially distressed hospitals also exhibit more negative cost and length-of-stay elasticities, though again the standard errors are too large to reject a zero response overall or equality with financially stable hospitals. Finally, there is no evidence that elasticities are larger in more competitive markets, as measured by the 1987 Herfindahl indices in each hospital's Health Service Area. Rather, the elasticity of mortality with respect to price is negative and significant at $p < 0.10$ for hospitals in the least-competitive markets [−1.05 (0.63)].

There are several potential explanations for the scant evidence of real responses to the very real price increases described in Section V. First, hospitals may be unable to select different intensity levels for each diagnosis (i.e., intensity is "lumpy" across diagnoses). New technologies or practice patterns, once put in place, may be difficult to apply to only a select group of patients. Second, patients may respond to a hospital's *overall* choice of intensity rather than intensity at the diagnosis level, providing little incentive for hospitals to allocate funds precisely where they are earned.[35] Third, if intensity choices are not initially in equilibrium, a hospital may allocate new funds earned in certain diagnoses to overdue investments in other areas.

To address these possibilities, I first reestimated (5) and (6) using Medicare's Major Diagnostic Categories (MDCs) in place of DRG pairs.[36] There were 25 MDCs in 1987, 16 of which contained one or more DRG pairs. If hospitals alter intensity at this level of aggregation, and patients in turn respond, then intensity and volume responses should be positive and sizeable. The point estimates again suggest small or negative responses, and the standard errors are of course larger due to the decline in the number of observations (from 650 to 112).

Next, I consider the possibility that hospitals may have responded only in diagnoses in which patients are likely to be quality-elastic. All admissions in the MedPAR files are assigned to one of five categories: emergency (admitted through the ER, 44 percent of admissions in 1987); urgent (first available bed, 29 percent); elective (23 percent); newborn (0.1 percent); unknown (4 percent). I assigned each DRG to the group accounting for the plurality of its admissions in 1987, and then estimated both stages of the intensity analysis separately by group. Again, I find no conclusive evidence of real responses in any group. The intensity responses do appear to be strongest (and correctly signed) in the elective diagnoses, but they cannot be statistically distinguished from the responses in other categories.

Thus, in contrast to previous studies, I find little evidence of intensity responses to price changes at the diagnosis level. In addition, I do not find conclusive evidence that hospitals were "pushing" lucrative admissions during this time period.

C. *Where Did the Money Go?*

Given the lack of intensity responses at the DRG level, the question arises: where did the money go? One possibility is that hospitals spread the funds across all admissions. To investigate this hypothesis, I aggregate the individual data into hospital-year cells. The relationship of interest is the elasticity of hospital intensity with respect to hospital price, which can be estimated from

$$(7) \qquad ln(intensity)_{ht} = \alpha + \boldsymbol{\mu} hospital_h + \boldsymbol{\delta} year_t + \beta ln(price)_{ht} + \varepsilon_{ht}.$$

However, there are two sources of bias in the OLS estimate of β: (1) the DRG recalibration method; and (2) the omission of an annual hospital-level measure of patient severity.[37] As with the previous analyses, the policy change

[35] Note that patients themselves need not have detailed knowledge of intensity levels; their primary care physicians and specialists may provide referrals based on their assessments of intensity.

[36] Tables are available upon request.

[37] Because true severity is difficult to capture with discharge data, this bias remains even if measures such as the Charlson index are included as controls.

can be used to identify β, but variation at the hospital level is required. Because hospitals with a large fraction of admissions in the "with CC" DRGs benefited the most from the policy change, the interaction between this measure and a dummy for the post-reform years can serve as an instrument for average price in equation (7).[38]

In constructing this instrument, I use the 1987 share of Medicare patients who are young (under 70) and coded with CC (hereafter called *share CC*). I select the pre-shock year because contemporaneous *share CC* would be affected by post-shock upcoding responses, and I use young patients because the data do not indicate whether old patients had CC before the policy change. This measure should be correlated with the mechanical component of the *hospital-level* price increase: hospitals with a large *share CC* in 1987 enjoyed larger increases in their average DRG price independently of their upcoding response to the policy change.

Table 6 gives the results from the first-stage regression of ln(price) on *share CC · post*

$$(8) \quad ln(price)_{ht} = \alpha + \boldsymbol{\mu} hospital_h + \boldsymbol{\delta} year_t + \tau_1 shareCC_h \cdot post_t + \varepsilon_{ht}$$

where $hospital_h$ is a vector of hospital fixed effects. All hospitals that appear in the MedPAR sample in 1987 are included in this (unbalanced) panel. The mean (standard deviation) of *share CC* in 1987 is 0.086 (0.043), and $\hat{\tau}_1$ is 0.233. A two-standard-deviation increase in *share CC* is therefore associated with a 2-percent increase in the average price paid to a hospital following the policy change. To illustrate that *share CC* is uncorrelated with changes in average hospital prices in the pre-reform years, column 2 presents the results from a regression of *ln*(*price*) on *share CC* · year dummies.

[38] An alternative instrument for hospital price is $\Delta ln(Laspeyres\ price)_{h,88\text{-}87}$. However, because the actual DRGs sampled for each hospital vary substantially over time, and DRG controls cannot be included in this specification, *share CC* is a much more accurate measure of the mechanical component at this level of aggregation.

TABLE 6—EFFECTS OF POLICY CHANGE ON AVERAGE HOSPITAL PRICES
(N = 36,651)

	Dependent variable is ln(price) mean(price) = 1.27	
Share CC · post	0.233***	
	(0.021)	
Share CC · year dummies		
1986		−0.022
		(0.040)
1987		−0.015
		(0.038)
1988		0.229***
		(0.038)
1989		0.212***
		(0.039)
1990		0.174***
		(0.040)
1991		0.270***
		(0.047)
Year dummies		
1986	0.039***	0.041***
	(0.001)	(0.004)
1987	0.057***	0.058***
	(0.001)	(0.004)
1988	0.063***	0.064***
	(0.002)	(0.004)
1989	0.088***	0.090***
	(0.002)	(0.004)
1990	0.094***	0.099***
	(0.002)	(0.004)
1991	0.119***	0.116***
	(0.002)	(0.004)
Adj. R-squared	0.890	0.890

Notes: The table reports results from two specifications of the first-stage regression relating ln(price) to *share CC*, the 1987 share of a hospital's Medicare patients who are under 70 and assigned to the top code of a DRG pair. In column 1, *share CC* is multiplied by an indicator for the post-1987 period ("post"). In column 2, *share CC* is interacted with individual year dummies. Both regressions include hospital fixed effects. The unit of observation is the hospital-year. All observations are weighted by the number of admissions in the 20-percent MedPAR sample. The sum of the weights is 13.7 million. Robust standard errors are reported in parentheses. * Signifies $p < 0.05$; ** signifies $p < 0.01$; *** signifies $p < 0.001$.

Coefficient estimates from the reduced-form equation

$$(9) \quad ln(intensity)_{ht} = \alpha + \boldsymbol{\mu} hospital_h + \boldsymbol{\delta} year_t + \tau_2 shareCC_h \cdot post_t + \varepsilon_{ht}$$

are presented in Table 7, followed by IV and OLS estimates of equation (7). The IV estimates for the elasticity of hospital intensity with respect to average hospital price are positive for

TABLE 7—REAL RESPONSES TO CHANGES IN AVERAGE HOSPITAL PRICE

	Dependent variable					
	ln(cost) mean = \$9,014	ln(LOS) mean = 8.81	ln(surgeries) mean = 1.21	ln(ICU days) mean = 0.81	ln(death rate) mean = 0.06	ln(volume) mean = 778
Reduced form						
Share CC · post	0.234**	0.069*	0.067	0.684***	0.122	0.403***
	(0.075)	(0.034)	(0.104)	(0.186)	(0.098)	(0.052)
IV estimate						
ln(price)	0.998***	0.296*	0.291	3.457***	0.536	1.728***
	(0.312)	(0.141)	(0.445)	(0.950)	(0.423)	(0.276)
Parametric tests of H_0: IV estimate $> =x$; H_1: IV estimate $< x$ (p-values are reported)[a]						
$x = 0.5$	0.96	0.06	0.31	1.0	0	1.0
$x = 1$	0.50	0	0.04	1.0	0	1.0
OLS estimate						
ln(price)	0.769***	0.350***	0.867***	1.483***	0.601***	−0.022
	(0.027)	(0.011)	(0.036)	(0.065)	(0.031)	(0.018)
N	36,169	36,651	35,897	28,226	34,992	36,651

Notes: The top panel of the table reports results from reduced-form regressions of ln(intensity) or ln(volume) on *share CC* multiplied by an indicator for the post-1987 period ("post"). *Share CC* is the 1987 share of a hospital's Medicare patients who are under 70 and assigned to the top code of a DRG pair. Intensity is measured by average costs, length of stay, number of surgeries, and the in-hospital death rate. The second and third panels report IV and OLS estimates, respectively, of the elasticity of hospital intensity and hospital volume with respect to hospital price. All regressions include year and hospital fixed effects. The unlogged, weighted mean of the dependent variable is reported at the top of each column. The unit of observation is the hospital-year. All observations are weighted by the number of admissions in the 20-percent MedPAR sample. The sum of the weights is 13.7 million. Robust standard errors are reported in parentheses.

[a] For ln(death rate), the tests are H_0: IV estimate $< =x$; H_1: IV estimate $> x$ for $x = -0.5$ and $x = -1.0$.

* Signifies $p < 0.05$; ** signifies $p < 0.01$; *** signifies $p < 0.001$.

four of the five intensity measures, and statistically significant for three. The exception is the in-hospital death rate, for which estimated elasticity is negative, but insignificant (a positive coefficient on death rate implies a negative intensity response). The elasticity results reveal that *an additional dollar of reimbursement goes wholly toward patient care.* Extra reimbursement is associated with longer stays, more surgeries, more ICU days, and possibly worse outcomes.[39] The intensity increases are consistent with previous studies that have examined hospital-wide responses to changes in the update factor (e.g., Jack Hadley et al., 1989).[40]

Hospitals benefiting disproportionately from the policy change also enjoyed increases in market share. Given the time resolution of the data, however, it is difficult to determine whether intensity increases are the signal that elicited these gains. The intensity results are therefore consistent with (at least) two distinct models of hospital behavior: competition in overall inten-

[39] Due to computing constraints, it is not possible to estimate equations (8) and (9) with individual hospital-year trends. If *share CC* is correlated with pre-existing trends in the dependent variables, the estimates in Table 7 will be upward-biased.

[40] The results can also be reconciled with Duggan (2000), who finds that private hospitals invested funds from the expansion of California's Disproportionate Share Program (DSH) in financial assets rather than patient care. There are at least two plausible reasons for this difference. First, the DSH payments were substantially larger, representing up to 30 percent of hospital revenues. Hospitals may react differently to such large budget shocks, particularly because the marginal benefit of dollars directed to patient care likely declines with the amount spent. Second, the DSH payments include a behavioral response: Duggan shows that private hospitals obtained their DSH funds by cream-skimming newly profitable patients away from public hospitals. Funds obtained in this manner may be spent differently than payments that are "mechanically" increased. Duggan's results suggest that the upcoding-induced payments I examine in Section IV may not have been allocated to patient care.

sity, and maximization of overall intensity subject to a budget constraint. The preponderance of the evidence does not, however, support the commonly assumed model of intensity competition at the diagnosis level. The lack of diagnosis-specific intensity responses contrasts with earlier research and helps to explain why diagnosis specialization is very limited in inpatient care.

VI. Conclusion

As public and private healthcare insurers continue to strengthen financial incentives for efficiency in the production of healthcare, it is critical to understand what the implications of such incentives are for health care quality and expenditures. The fixed-price system used by many insurers makes hospitals the residual claimants of profits earned on inpatient stays. These profits differ by diagnosis, creating incentives for hospitals to increase the volume of admissions in profitable diagnoses relative to unprofitable diagnoses. Soliciting the most lucrative type of business may have few deleterious consequences in other fixed-price settings (e.g., utilities), but it is potentially dangerous in the healthcare industry. For example, doctors at Redding Medical Center, a for-profit hospital operated by Tenet Healthcare Corporation in Redding, California, are currently under criminal investigation for performing lucrative open-heart surgeries in place of medically managing symptoms of heart disease (Kurt Eichenwald, 2003).

Resolving the question of how hospitals respond to changes in DRG prices, which are simply shocks to the profitability of certain diagnoses or treatments, is therefore critical from a policy standpoint. In addition, these responses provide a window into industry conduct. In theory, quality erosion is kept in check by competition among hospitals.[41] Responses to individual price changes can reveal whether this competition occurs at the diagnosis level.

This study illustrates how a simple change in the DRG classification system in 1988 generated large and exogenous price changes for 40 percent of DRG codes, accounting for 43 percent of Medicare admissions. Hospitals responded to these price changes by upcoding patients to DRG codes associated with large reimbursement increases, garnering \$330–\$425 million in extra reimbursement annually. They proved quite sophisticated in their upcoding strategies, upcoding more in those DRGs where the reward for doing so increased more. Additionally, while all subsamples of hospitals upcoded in response to the policy change, for-profit facilities availed themselves of this opportunity to the greatest extent.

Whereas coding behavior proved very responsive to diagnosis-specific financial incentives, admissions and treatment policies did not. I do not find convincing evidence that hospitals increased admissions differentially for those diagnoses with the largest price increases, although this practice may have become more prevalent in recent years. The results also suggest that healthcare insurers cannot effect an increase in the quality of care provided to patients with a particular diagnosis simply by increasing reimbursement rates for that diagnosis. However, there is evidence that hospitals spend extra funds they receive on patient care in *all* DRGs. This suggests that hospitals do not (or cannot) optimize quality choices by product line, which may explain the relative lack of specialization in the hospital industry. One anticipated benefit of PPS was that hospitals would specialize in admissions in which they are relatively cost-efficient. If, however, hospitals do not balance costs and benefits within individual product lines, such specialization is unlikely to occur.

This research exposes the difficulties inherent in implementing a fixed-price system in which the output is difficult to verify. As Medicare and other insurers continue to extend prospective payment systems to additional areas, they would do well to consider the evidence that providers of all ownership types may behave strategically to garner additional resources.

[41] Of course, physicians also play an important role in ensuring appropriate care for their patients, as highlighted by Kenneth J. Arrow (1963).

Appendix

Table A1—Descriptive Statistics for Hospital Characteristics

Variable	Mean	*SD*	Min	Max
Ownership				
For-profit	0.14	0.35	0	1
Non-profit	0.58	0.49	0	1
Government	0.28	0.45	0	1
Financial distress measures				
Debt:asset ratio	0.52	0.32	0	2.17
Medicare bite	0.37	0.11	0	1.00
Medicaid bite	0.11	0.08	0	0.84
Region				
Northeast	0.14	0.35	0	1
Midwest	0.29	0.46	0	1
South	0.38	0.49	0	1
West	0.18	0.38	0	1
Size				
1–99 beds	0.46	0.50	0	1
100–299 beds	0.37	0.48	0	1
300+ beds	0.17	0.37	0	1
Service offerings				
Teaching program	0.06	0.23	0	1
Open-heart surgery	0.13	0.34	0	1
Trauma facility	0.19	0.39	0	1
ICU beds (except neonatal)	10.27	12.29	0	194
Market concentration				
HSA Herfindahl (AHA)	0.07	0.05	0	1
HSA Herfindahl (*Dartmouth Atlas of Health Care*)	0.67	0.35	0	1

Notes: The table reports descriptive statistics for hospitals with complete data. Hospitals with debt:asset ratios in the 1% tails of the distribution are also excluded. $N = 5352$. The analyses in Tables 4, 6, and 7 utilize data from this subset of hospitals; the analyses in Tables 3 and 5 do not require hospital characteristics and therefore include data from all hospitals financed under PPS.

REFERENCES

Arrow, Kenneth J. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review*, 1963, *53*(5), pp. 941–73.

Carter, Grace M.; Newhouse, Joseph P. and Relles, Daniel A. "How Much Change in the Case Mix Index Is DRG Creep?" *Journal of Health Economics*, 1990, *9*(4), pp. 411–28.

Coulam, Robert F. and Gaumer, Gary L. "Medicare's Prospective Payment System: A Critical Appraisal." *Health Care Financing Review Annual Supplement*, 1991, *12*, pp. 45–77.

Cutler, David M. "Empirical Evidence on Hospital Delivery under Prospective Payment." Unpublished Paper, 1990.

Cutler, David M. "The Incidence of Adverse Medical Outcomes under Prospective Payment." *Econometrica*, 1995, *63*(1), pp. 29–50.

Cutler, David M. "Cost Shifting or Cost Cutting? The Incidence of Reductions in Medicare Payments," in James M. Poterba, ed., *Tax policy and the economy*, Vol. 12. Cambridge, MA: MIT Press, 1998, pp. 1–27.

Dafny, Leemore S. and Gruber, Jonathan. "Public Insurance and Child Hospitalizations: Access and Efficiency Effects." *Journal of Public Economics*, 2005, *89*(1), pp. 109–29.

Dartmouth Medical School, Center for the Evaluative Clinical Sciences. *The Dartmouth atlas of health care*. Chicago: American Hospital Publishing, 1996.

Dranove, David. "Rate-Setting by Diagnosis Related Groups and Hospital Specialization." *RAND Journal of Economics*, 1987, *18*(3), pp. 417–27.

Dranove, David. "Pricing by Non-Profit Institutions: The Case of Hospital Cost-Shifting." *Journal of Health Economics*, 1988, *7*(1), pp. 47–57.

Duggan, Mark G. "Hospital Ownership and Public Medical Spending." *Quarterly Journal of Economics*, 2000, *115*(4), pp. 1343–73.

Duggan, Mark G. "Hospital Market Structure and the Behavior of Not-for-Profit Hospitals." *RAND Journal of Economics*, 2002, *33*(3), pp. 433–46.

Eichenwald, Kurt. "Operating Profits: Mining Medicare; How One Hospital Benefited on Questionable Operations," *The New York Times*, August 12, 2003, A1.

Ellis, Randall P. and McGuire, Thomas G. "Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects." *Journal of Health Economics*, 1996, *15*(3), pp. 257–77.

Gilman, Boyd H. "Hospital Response to DRG Refinements: The Impact of Multiple Reimbursement Incentives on Inpatient Length of Stay." *Health Economics*, 2000, *9*(4), pp. 277–94.

Hadley, Jack; Zukerman, Stephen and Feder, Judith. "Profits and Fiscal Pressure in the Prospective Payment System: Their Impacts on Hospitals." *Inquiry*, 1989, *26*(3), pp. 354–65.

Hodgkin, Dominic and McGuire, Thomas G. "Payment Levels and Hospital Response to Prospective Payment." *Journal of Health Economics*, 1994, *13*(1), pp. 1–29.

Lagnado, Lucette. "Columbia/HCA Grades Its Hospitals on Severity of Their Medicare Claims," *The Wall Street Journal*, May 30, 1997, A1, A6.

Lichtenberg, Frank R. "The Effects of Medicare on Health Care Utilization and Outcomes," in Alan M. Garber, ed., *Frontiers in health policy research*. Vol. 5. Cambridge, MA: MIT Press, 2002, pp. 27–52.

Murray, Lauren A. and Eppig, Franklin J. "Insurance Trends for the Medicare Population, 1991–1999." *Health Care Financing Review*, 2002, *23*(3), pp. 9–16.

Pstay, Bruce M.; Boineau, Robin; Kuller, Lewis H. and Luepker, Russell V. "The Potential Costs of Upcoding for Heart Failure in the United States." *American Journal of Cardiology*, 1999, *84*(1), pp. 108–09.

Shleifer, Andrei. "A Theory of Yardstick Consumption." *RAND Journal of Economics*, 1985, *16*(3), pp. 319–27.

Silverman, Elaine M. and Skinner, Jonathan S. "Medicare Upcoding and Hospital Ownership." *Journal of Health Economics*, 2004, *23*(2), pp. 369–89.

Silverman, Elaine M.; Skinner, Jonathan S. and Fisher, Elliott S. "The Association between For-Profit Hospital Ownership and Increased Medicare Spending." *New England Journal of Medicine*, 1999, *341*(6), pp. 420–26.

Staiger, Douglas and Gaumer, Gary L. "The Impact of Financial Pressure on Quality of Care in Hospitals: Post-Admission Mortality under Medicare's Prospective Payment System." Unpublished Paper, 1992.

Steinwald, Bruce and Dummit, Laura A. "Hospital Case-Mix Change: Sicker Patients or DRG Creep?" *Health Affairs*, 1989, *8*(2), pp. 35–47.

U.S. Census Bureau, Statistical Compendia Branch. *Statistical abstract of the United States*. Washington, DC: U.S. Government Printing Office, 1987, 1991, 1998, 2001.

U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services. *2002 data compendium*. Washington, DC: U.S. Government Printing Office, 2002.

Office of the Federal Register, National Archives and Records Service. *Federal register*. Washington, DC: U.S. Government Printing Office, 1984–2000.

Weisbrod, Burton A. "Why Private Firms, Governmental Agencies, and Nonprofit Organizations Behave Both Alike and Differently: Application to the Hospital Industry." Northwestern University, Institute for Policy Research Working Papers: No. WP-04-08, 2004.