WP-03-17

HOW DO HOSPITALS RESPOND TO PRICE CHANGES?

Leemore Dafny
Northwestern University
IPR
NBER

September 2003

Abstract

This paper investigates whether hospitals respond in profit-maximizing ways to changes in diagnosis-specific prices, as determined by Medicare's Prospective Payment System and other public and private insurers. Previous studies have been unable to isolate this response because changes in reimbursement amounts (prices) are typically endogenous: they are adjusted to reflect changes in hospital costs. I exploit an exogenous 1988 policy change that generated large price changes for 43 percent of all Medicare admissions. I find that hospitals responded to these price changes by "upcoding" patients to diagnosis codes associated with large reimbursement increases, garnering \$330-\$425 million in extra reimbursement annually. This response was particularly strong among for-profit hospitals. With the important exception of elective diagnoses, I find little evidence that hospitals increased the intensity of care in diagnoses subject to price increases, where intensity is measured by total costs, length of stay, number of surgical procedures, and number of intensive-care-unit days. Neither did hospitals increase the volume of patients admitted to more remunerative diagnoses, notwithstanding the strong a priori expectation that such a response should prevail in fixed-price settings. Taken together, these findings suggest that, for the most part, hospitals do not alter their treatment or admissions policies based on diagnosis-specific prices; however, they employ sophisticated coding strategies in order to maximize total reimbursement. The results also suggest that models of quality competition among hospitals may be inappropriate at the level of specific diagnoses ("products").

---

e-mail: l-dafny@kellogg.northwestern.edu. I am grateful to Jonathan Gruber, James Poterba, David Cutler, and Glenn Ellison for excellent guidance. I thank Josh Angrist, William Collins, Joseph Doyle, Mark Duggan, Julian Jamison, David Levine, Joseph Newhouse, Scott Stern, Nancy Rose, and seminar participants at several universities and the NBER Universities Research Conference for helpful comments. 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 Introduction

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 non-negative 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 establishing fixed prices for Medicare hospitalizations. Other public and private insurers soon followed suit, wresting price-setting control from providers and imposing yardstick competition.

A large literature documents hospitals' responses to the introduction of fixed prices, but few studies have explored reactions to *changes* in these prices. Yet once the transition to a fixed-price regime is completed, 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? 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.

This study focuses on inpatient care for Medicare beneficiaries, who account for 37 percent of hospital discharges and 31 percent of total revenues.[1] Since 1984, hospital reimbursement for Medicare patients has been governed by the Prospective Payment System (PPS), which provides a

[1] 2002 Data Compendium, *Centers for Medicare and Medicaid Services*, and author's tabulations from the 2000 Survey of Hospitals (administered by the American Hospital Association). Figures are for 2000.

fixed payment for each Medicare patient in a given hospital and diagnosis-related group (DRG). Standard models of hospital behavior, reviewed in section 2, predict that hospitals will respond to a diagnosis-specific price increase by raising the intensity of care provided to patients in that diagnosis. According to these models, hospitals behave much like multiproduct firms, where the products are DRGs and the choice variables are not prices but rather intensity of care within each DRG. Both patient volume and hospital costs are assumed to increase in the intensity of care provided. A price increase for a given DRG raises the profitability of that DRG, creating an incentive to attract more patients by increasing the intensity of care that is provided. Indeed, the few studies that investigate the effect of DRG-level price changes on intensity levels all find a positive relationship, where intensity of care is measured by length of stay, number of surgical procedures, and/or death rates (Cutler 1990, 1995; Gilman 2000). 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.

All of the aforementioned studies utilize data from a transition to a prospective payment system, either PPS or one of the many systems implemented by state Medicaid programs. These studies therefore 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. By investigating responses to 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. Estimating responses to price changes in post-implementation eras is difficult, however, because price changes are typically endogenous: they are adjusted to reflect changes in hospital costs. Thus, positive associations between changes in price and changes in spending or intensity 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 an exogenous 1988 policy change that generated large price changes for 43 percent of Medicare admissions. The policy change was simply the elimination of “age over 69” and “age under 70” in the descriptions for the diagnosis-related groups (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 change, which is described in greater detail in Section 3, actually led to large increases in reimbursement for patients assigned to DRG codes with these qualifiers (“affected DRGs”), as compared to patients in other codes (“unaffected DRGs”).

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 simply entice hospitals to “upcode,” or switch patients from lower-paying DRGs into 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 tradeoffs in their upcoding practices: although the policy shock created a blanket incentive to increase upcoding in dozens of diagnoses, hospitals upcoded more in those diagnoses where the incentive to do so was larger. The upcoding response was also strongest among for-profit hospitals, a finding that is consistent with prior research.

Using the unaffected DRGs as a control group, I find that hospitals did not increase the intensity or quality of care provided to patients in affected DRGs, where intensity is measured by total costs, length of stay, number of surgical procedures, and number of intensive-care-unit (ICU) days, and quality by the in-hospital death rate. DRGs in which the plurality of admissions are elective were the sole exception: hospitals did increase their spending in affected DRGs relative to unaffected DRGs in this category, although this increased spending did not translate into significant increases in the other dimensions of intensity that I measure. Across the board, hospitals did not increase the volume of patients admitted to more remunerative diagnoses, a finding that is theoretically consistent with the general intensity non-response, but perhaps surprising given theoretical predictions of firm behavior in fixed-price settings. I do find evidence that hospitals spent the extra funds they earned on patient care, but these funds were spread across *all* admissions. Correspondingly, overall hospital volume growth was also stronger for hospitals with larger price gains (and therefore intensity increases) arising from the policy change.

Taken together, these findings indicate that hospitals generally do not alter their treatment or admissions policies based on diagnosis-specific prices; however, they employ sophisticated coding strategies in order to maximize total reimbursement. 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 their reimbursement rates for that diagnosis. Another important implication is that models of quality competition among hospitals may be inappropriate at the level of specific diagnoses. Finally, this research illustrates the difficulties inherent in regulating prices in an industry where the products are hard to define.

The remainder of the paper is organized into 5 sections. Section 2 describes PPS and prior related research, and introduces a hospital objective function that provides a theoretical framework for the empirical sections that follow. Section 3 gives a detailed explanation of the 1988 policy change. The data are presented in Section 4, followed by an evaluation of the aggregate impact of the policy change on price levels in Section 5. Section 6 quantifies the share of the price change attributable to mistakes by price-setting authorities (the exogenous or *mechanical* component), as compared to true changes in patient mix (the *severity* component), and hospital upcoding (the *upcoding* component). Section 7 explores the intensity and volume responses to the exogenous component of the price change, and Section 8 concludes.

2 Background

2.1 A PPS Primer

The Prospective Payment System (PPS) for hospitalizations of Medicare beneficiaries was implemented in October 1984 by the Health Care Financing Administration (HCFA), now known as the Centers for Medicare and Medicaid Services (CMS). The defining element of the system 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 Diagnosis-Related Groups (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

$$P_{hd} = P_h \cdot (1 + IME_h) \cdot (1 + DSH_h) \cdot \text{DRG weight}_d$$

where P_h is a hospital-specific amount (inflated annually by a Congressionally-approved "update factor"), IME represents an adjustment for indirect medical education (teaching), and DSH adjusts payment levels to compensate hospitals with a disproportionate share of indigent patients.[2] Most of the variation in P_{hd} is due to the DRG weights, which range between .09 (DRG 448 for allergic reactions) to 22.8 (DRG 480 for liver transplants).[3] CMS uses hospital charge data (deflated by hospital cost:charge ratios) to recalibrate the weights annually, raising weights for DRGs that experience relative increases in average charges, and reducing weights for DRGs with relative decreases in average charges. The average DRG weight per hospital admission has risen substantially over time, from 1.13 in 1984 to 1.36 in 1996.[4] This phenomenon has been termed "DRG creep," as patients are

[2] This simplified formula appears in Cutler (1995).
[3] The range for DRG weights is given for 1985-1996.
[4] Steinwald and 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.

increasingly coded into DRGs with higher weights. A one-percent increase in the average case weight is associated with an additional $930 million in *annual* Medicare payments to hospitals.[5]

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. Section 2.2 illustrates this incentive more formally.

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 3.

2.2 Hospital Objective Functions

To illustrate how changes in DRG prices might affect hospital behavior, it is helpful to introduce a simple model for the hospital objective function. I begin with the traditional assumptions that hospitals attach non-negative weights to both patient care (often called "intensity" or quality) and profits, and that the objective function is separable in these arguments:

$$\max\ G_h = \alpha_h f(I_h) + (1 - \alpha_h)\pi_h$$

where $0 < \alpha < 1$, h is a hospital index, I denotes intensity, and π denotes profits.

The PPS system effectively defines D "product lines" for every hospital, where D is the number of DRGs. Each hospital selects an intensity level I_{hd} for each DRG d, attracting $N_{hd}(I_{hd}, I_{\sim hd})$ patients, where ~h denotes hospital h's competitors. Patient demand is increasing in a hospital's own intensity

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

level (at a decreasing rate), and decreasing in that of its competitors. Because higher intensity levels attract sicker patients, the severity of patients served, $S_{hd}(I_{hd})$, is also increasing in a hospital's intensity level. For each admission, the hospital earns $P_{hd} - C_{hd}(I_{hd}, S_{hd}(I_{hd}))$, where P_{hd} is as defined above, C_{hd} is the average cost per patient

assigned to DRG d, and $\frac{\partial C_{hd}}{\partial I_{hd}}$ and $\frac{\partial C_{hd}}{\partial S_{hd}}$ are greater than zero.[6] Thus, the hospital's problem becomes

$$\max\ G_h = \alpha_h f(I_{h1}, I_{h2} \ldots I_{hD}) + (1-\alpha_h) \sum_{d=1}^{D} \left([P_{hd} - C_{hd}(I_{hd}, S_{hd})] N_{hd}(I_{hd}, I_{\sim hd}) \right),$$

and the first-order condition for I_{hd}, taking competitors' behavior as given, is

$$\frac{\partial G_h}{\partial I_{hd}} = \alpha_h \frac{\partial f}{\partial I_{hd}} + (1-\alpha_h) \left[(P_{hd} - C_{hd}) \frac{\partial N_{hd}}{\partial I_{hd}} - N_{hd} \left(\frac{\partial C_{hd}}{\partial I_{hd}} + \frac{\partial C_{hd}}{\partial S_{hd}} \bullet \frac{\partial S_{hd}}{\partial I_{hd}} \right) \right] = 0$$

For every DRG, the hospital equates the marginal benefit of intensity with its marginal cost. This expression implicitly defines the optimal intensity choice, I_{hd}^*. To illustrate that an increase in P_{hd}

raises optimal intensity, I set $\frac{\partial G_h}{\partial I_{hd}}(I_{hd}^*, P_{hd}) = 0$, differentiate with respect to P_{hd}, and solve for $\frac{dI_{hd}^*}{dP_{hd}}$.

Under the assumptions that G_h is twice differentiable and concave in I_{hd} $\forall$ d, and that I_{hd} and $I_{\sim hd}$ are strategic complements $\forall$ d,

$$\frac{dI_{hd}^*}{dP_{hd}} = \frac{-(1-\alpha_h)(\partial N_{hd}/\partial I_{hd}) - \left(\partial^2 G_h / \partial I_{hd} \partial I_{\sim hd} \bullet dI_{\sim hd}/dP_{hd}\right)}{\partial^2 G_h / \partial I_{hd}^{\ 2}} > 0.$$ [7]

This result suggests that price increases should be associated with a "flypaper effect" of the sort widely-documented in the public sector: additional funds are not treated as general income but are spent where they are raised. My primary empirical objective is to test this prediction explicitly by

---

[6] This model is based on Dranove (1987), Hodgkin and McGuire (1994), Ellis and McGuire (1996), and Gilman (2000).
[7] I adopt the definition of Bulow, Geanakoplos, and Klemperer (1985) by using $\partial^2 G_h / \partial I_{hd} \partial I_{\sim hd} > 0$ to denote strategic complements.

investigating whether hospital costs and other measures of intensity increased more for DRGs that were more highly reimbursed after the policy change. This analysis is presented in Section 7.

Section 7 tests another prediction that follows from the flypaper effect: the *volume* of admissions in DRGs subject to price increases should grow.[8] If intensity levels rise as a result of price increases, by assumption volume should increase as well. This is the classical response expected in fixed-price industries: when price increases, so long as it exceeds marginal cost, firms will want to produce more.

There are several reasons these results may not obtain. First, $\partial N_{hd}/\partial I_{hd}$ may be very small, reducing the effect of a price increase on intensity levels. Patients may respond to a hospital's overall choice of intensity ("I_h"), but not to I_{hd}, which is more difficult to ascertain.[9] Second, hospitals may be unable to select different intensity levels for each DRG (i.e., intensity is "lumpy" across DRGs). New technologies or practice patterns, once put in place, may be difficult to apply to only a select group of patients. Third, if intensity choices are not initially in equilibrium, a hospital may allocate new funds earned in affected DRGs to overdue investments in unaffected DRGs. Finally, hospitals may maximize objectives that are not captured in the functional form above, such as the total volume of patients.

The objective function G_h is quite general, allowing for heterogeneity in hospitals' responses to the same payment incentives. Any characteristic that affects the parameter α_h will affect the intensity response to a price increase. For example, for-profit hospitals should place a higher weight on profits (lower α_h), as should hospitals under financial duress. The "mission" of a hospital, reflected by such characteristics as teaching status, may also affect the tradeoff between intensity and profits. Alternatively, different hospitals with the same α_h may be differentially-equipped to respond to reimbursement incentives. Small hospitals in particular lack the resources needed to reoptimize

[8] Strictly speaking, this is true so long as the price-induced changes in a hospital's own intensity have a greater impact on its volume than the price-induced changes in the intensity of its competitor(s).
[9] Note that patients themselves need not have detailed knowledge of intensity levels; their primary care physicians and specialists may refer them to hospitals based on their assessments of intensity.

quickly in the face of price changes. Finally, there are important regional differences in hospital behavior, although there are few theoretical explanations for this phenomenon apart from "cultural norms."

Differences across hospitals are one possible source of variation in intensity responses; differences across DRGs are another. For example, 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 is therefore larger for such admissions, raising hospitals' incentives to increase quality in the face of price increases. Thus, the same price increase may elicit different intensity responses across DRGs. I explore differences in intensity and volume responses across hospitals and admission types in sections 7.2.1 and 7.2.2, respectively.

2.2.1 Incorporating Upcoding

The general model outlined above can be easily expanded to include upcoding effects. Using U_{hd} to denote an "upcoding index," the number of patients N_{hd} can be redefined as an increasing function of U_{hd} and a decreasing function of $U_{h\sim d}$, the degree of upcoding in other DRGs. Holding the number of patients constant, if more patients are upcoded into DRG d, fewer patients are assigned to other DRGs. Upcoding a patient to DRG d also reduces average severity in DRG d (else it would not be upcoding); the effect on average severity in the original DRG is ambiguous. To summarize,

$$N_{hd} = N_{hd}(I_{hd}, I_{\sim hd}, U_{hd}, U_{h\sim d}),\ \frac{\partial N_{hd}}{\partial I_{hd}} > 0, \frac{\partial N_{hd}}{\partial I_{\sim hd}} < 0, \frac{\partial N_{hd}}{\partial U_{hd}} > 0, \frac{\partial N_{hd}}{\partial U_{h\sim d}} < 0$$

$$S_{hd} = S_{hd}(I_{hd}, U_{hd}, U_{h\sim d}),\ \frac{\partial S_{hd}}{\partial I_{hd}} > 0, \frac{\partial S_{hd}}{\partial U_{hd}} < 0, \frac{\partial S_{hd}}{\partial U_{h\sim d}} <> 0\ .$$

Adding a probability of detection μ_h that is increasing in the level of upcoding, a penalty T_h if the hospital is caught upcoding, and a total cost of upcoding R, the objective function becomes

$$G_h = \alpha_h f(I_{h1}, I_{h2}..I_{hD}) + (1-\alpha_h)\begin{bmatrix} \sum_{d=1}^{D}\left([P_{hd} - C_{hd}(I_{hd}, S_{hd}(I_{hd}, U_{h1}, U_{h2}..U_{hD}))]N_{hd}(I_{hd}, I_{\sim hd}, U_{h1}, U_{h2}..U_{hD})\right) \\ -\mu(U_{h1}, U_{h2}...U_{hD})T_h - R(U_{h1}, U_{h2}...U_{hD}) \end{bmatrix}$$

with the following first-order condition for U_{hd}:

$$\frac{\partial G_h}{\partial U_{hd}} = (1-\alpha_h)\left[\sum_{j=1}^{D}\left((P_{hj} - C_{hj})\frac{\partial N_{hj}}{\partial U_{hd}} - N_{hj}\left(\frac{\partial C_{hj}}{\partial S_{hj}} \bullet \frac{\partial S_{hj}}{\partial U_{hd}}\right)\right) - \frac{\partial \mu}{\partial U_{hd}}T_h - \frac{\partial R}{\partial U_{hd}}\right] = 0.$$

Hospitals trade off the added revenue (less any change in treatment costs) from shifting patients into higher-weighted DRGs 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. Holding the penalties and costs associated with upcoding constant, a price increase for a given DRG increases the incentive to upcode patients into that DRG.[10]

The coding of patient conditions is performed by administrative staff, who use hospital charts and the ICD-9 diagnosis codes provided by physicians to map patient conditions into DRGs (Silverman and Skinner 2000). Upcoding costs therefore depend upon the availability of multiple DRG codes for similar diagnoses. It is theoretically possible to assign a patient with bronchitis to the heart transplant DRG, but such overt upcoding requires altering or misinterpreting medical records substantially and increases the risk of detection later on.[11] The policy change I study involves DRGs that are particularly susceptible to upcoding because these are DRGs in which the coding of patient complications results in a substantially higher price. One former manager from the largest for-profit hospital chain,

[10] The conditions for this prediction to hold are analogous to those in section 2.2: G_h must be twice differentiable and concave in U_{hd}, and the cross-partial $\partial^2 G_h / \partial U_{hd} \partial I_{hd} \geq 0$. This cross-partial can reasonably be expected to equal zero, as the marginal benefit of intensity should not vary with upcoding.

[11] Regulatory agencies known as "Peer Review Organizations" regularly audit DRG assignments. CMS works with the Office of the Inspector General (OIG), the FBI, and the US Attorney's Office to levy fines, recover funds, and prosecute providers who defraud the Medicare program. There are *qui tam* provisions to reward and protect whistle-blowers.

Columbia/HCA (now HCA), reported that hospital managers were rewarded for upcoding patients with these diagnoses into the more-remunerative "with complications" codes (Lagnado 1997). Section 6 presents results on upcoding following the 1988 policy change.

As with intensity levels, there are many reasons that upcoding behavior may differ across hospitals and DRGs. Hospitals with a lower α_h should upcode more, while hospitals with a greater penalty T_h (real or perceived, monetary or otherwise) or a higher probability of detection μ_h should upcode less. There are a number of theories of the effect of hospital ownership on upcoding, but few consensus predictions (see Silverman and Skinner 2000 for a comprehensive discussion). Hospitals experiencing financial distress should be more willing to risk detection, all things equal, while larger hospitals may be "savvier" in training their coding personnel. Practices of competitors may also affect upcoding indirectly through pressure on hospital profits, or directly via the dissemination of upcoding practices.[12]

Finally, upcoding may also vary across DRGs. Diagnoses based on subjective interpretations of patient conditions are more prone to upcoding, as are diagnoses for which minor variations (e.g., presence of a complication) are associated with large reimbursement differences. The upcoding analysis in section 6 focuses on diagnoses in this latter group. Within this subset of conditions, I also investigate the relationship between the extent of upcoding in a particular diagnosis and the financial incentive to upcode.

2.3 Previous Research

2.3.1 Average Reimbursement Effects

Virtually all of the papers that evaluate the impacts of PPS do not distinguish between the effects due to changes in marginal reimbursement (during the phase-in of the system) and those due to changes in

[12] Several recent studies document this indirect channel, e.g., Duggan (2002), which finds that not-for-profit hospitals respond more strongly to financial incentives to treat indigent patients in markets with greater for-profit penetration.

average reimbursement levels (P_{dh}).[13] The first papers to distinguish these effects at the diagnosis level are Cutler (1990) and Cutler (1995).[14] Cutler (1990) studies 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 .2, Cutler does not find a corresponding volume response.[15] 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 affects all DRGs equally. However, intensity reductions may be easier to make in certain DRGs and/or hospitals, and to the extent that price reductions were more prevalent in such DRGs and/or hospitals (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.[16]

The two additional studies addressing DRG-specific intensity responses to price changes employ different identification strategies but reach the same conclusion. Gilman (2000) investigates the impact of a 1994 reform to Medicaid DRGs for HIV diagnoses in New York. He finds that length of stay increased in procedure-based DRGs, which were subject to price increases, and decreased in

[13] Hodgkin and McGuire (1994) provide an excellent overview of empirical research on this subject.
[14] Studies of *hospital*-level responses to changes in average reimbursement amounts include Hadley, Zuckerman, and Feder (1989) and Staiger and Gaumer (1992). These works find positive intensity responses as measured by length of stay and patient survival, respectively. Cutler (1998) studies responses to average payment reductions implemented through the annual update factor. He finds cost-shifting to private payors in the early PPS era (1985-1990), and cost-cutting through capacity and nursing staff reductions in the later PPS era (1990-1995).
[15] Such a result could be consistent with a model in which volume is not a function of intensity, and hospitals simply maximize intensity within each DRG subject to a DRG-specific breakeven constraint.
[16] Cutler's methodology for calculating the change in average payment incentives following the implementation of PPS is likely to lead to 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 55-64 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. Such hospitals may have had "more fat to trim" in terms of intensity provision.

non-procedure-based DRGs, which were subject to price decreases. Assuming the controls for patient severity adequately capture the severity changes in the patient population for both admission types, these results also suggest that hospitals adjust DRG-specific intensity in response to price changes. Newhouse (1989) finds some evidence that private hospitals successfully shifted patients in unprofitable DRGs to public hospitals following the implementation of PPS; the mechanism for this shift is not specified, but the finding is consistent with real responses to incentives at the DRG level.[17] As with the Cutler studies, these works investigate simultaneous changes in marginal and average reimbursement incentives. The policy change I assess affects only average reimbursement levels, eliminating the need to disentangle the responses to changes in marginal incentives. In addition, because the policy change affected a large proportion of DRG codes (40 percent), the analysis produces representative estimates of DRG-specific intensity responses.

2.3.2 Upcoding

Because the single largest source of increased hospital spending by Medicare is the rapid rise in the average case weight, the subject of upcoding has generated a substantial literature. Coulam and Gaumer (1991) review this 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.

Carter, Newhouse, and Relles (1990) use the 'gold standard' in chart review to estimate the role of upcoding in the case-mix increase between 1986 and 1987: they send a nationally representative sample of discharge records from 1986 and 1987 to an expert coding group (called the "SuperPRO") that regularly reviews samples of discharges to enforce coding accuracy. They find that one-third of

[17] Newhouse specifically considers the possibility that private hospitals transferred unprofitable patients to public hospitals after admission, but does not find any evidence to support this mechanism for case redistribution.

the case-mix increase was due to upcoding, although the standard error of this estimate is large. More recently, Psaty et al (1999) use 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 non-medical 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 period, the results are again difficult to interpret; patient severity changed dramatically due to changes in patient composition following the implementation of PPS.

A recent study by Silverman and Skinner (2000) 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 that are assigned to the most expensive DRG possible, Silverman and Skinner document large increases in upcoding, despite a downward trend in mortality rates. Interestingly, the authors find that for-profit hospitals upcode the most, and that not-for-profit hospitals are more likely to engage in upcoding when area market share of for-profit hospitals is higher, independently of financial distress and other control variables. This finding is consistent with a contagion model like that described in Cutler and Horwitz (1999), or with the "cultural norms" hypothesis. In addition, Silverman and Skinner find that hospitals under financial distress upcode *less* than financially sound institutions.

My upcoding analysis takes a similar approach, but the policy change I analyze offers two important advantages. First, I study an abrupt change in upcoding incentives that should be met with a similarly abrupt change in upcoding if hospitals are responsive to these incentives. Second, because the policy change created upcoding incentives that vary by diagnosis, I am able to investigate not only whether hospitals respond to upcoding incentives in general, but also whether they respond to upcoding incentives *on the margin*, upcoding more when the payoff is greater.

3 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).[18] 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 the affected diagnoses, HCFA eliminated the age over 69/under 70 criterion beginning in 1988. 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 gives the three most commonly-coded pairs and their DRG weights before and after the policy change.[19] 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

[18] ProPAC, now incorporated into MedPAC (Medicare Payment Advisory Commission), was an independent federal agency that reported to Congress on all PPS matters.
[19] 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.

bottom code was 6.2 percent. In the final notice of the policy change, HCFA clearly states that the goal of the recalibration was to ensure no overall 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] It is worth emphasizing, however, that while annual recalibrations are intended to be revenue-neutral *overall*, there is no requirement that they be revenue-neutral for any subset of DRGs.

Indeed, as the analysis in Section 5 reveals, this policy change resulted in a large relative price increase (7 percent) for discharges coded in DRG pairs, and a moderate absolute price increase (1-2 percent). There are three sources of this price increase: a *mechanical* component, an *upcoding* component, and a *severity* component. The mechanical component is the effect of the recalibration on prices, holding the incidence of reported complications constant – essentially, it captures mistakes made by HCFA in its recalibration. The upcoding component captures the opportunistic coding of complications, while the severity component is associated with an increase in the true incidence of complications. In 1989, HCFA published its own (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 .93 percentage points could be attributed to faulty recalibration of DRG weights for 1988, and an additional .29 percentage points to similar errors in 1987. These 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 large relative effects on the DRG pairs were unabated.

The 1988 policy change provides an excellent opportunity to study hospital responses to changes in DRG-specific prices. After describing the data, I analyze the effects of this price shock in three parts. First, I estimate the magnitude of the shock to prices for affected DRGs. Second, I disaggregate this price increase into its mechanical, upcoding and severity components. Third, I

[20] There were only a few minor changes to the GROUPER program between 1987 and 1988 that were not associated with the elimination of the age criterion.

investigate the elasticity of DRG-specific intensity and volume with respect to price, using the mechanical component as an instrument for price.

4 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. 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). Appendix Table 1 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.

[21] The Cost Reports also contain an indicator for whether a hospital is paid under the PPS system (certain hospitals are exempted). I omit exempt hospitals from my sample.

5 Assessing the Magnitude of the Price Shock

The elimination of the age criterion resulted in large price changes for individual DRGs, as described in section 3. However, 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 yield information on 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.[22] In order to keep the reference population constant before and after the policy reform, I combine data from the top and bottom codes, effectively creating a single DRG for each pair. It is therefore critical to illustrate that the average price paid for patients in these newly-created paired DRGs did indeed increase following the 1988 elimination of the age criterion.

To assess the magnitude of this price increase, I employ a differences-in-differences technique, comparing the time-series changes in price for the paired DRGs (henceforth the "affected DRGs") with the changes in price for the single DRGs (the "unaffected" DRGs). While prices for unaffected DRGs are given annually by HCFA, prices for affected DRGs must be calculated by taking a weighted average of the prices for the top and bottom codes in each pair. For example,

$$\text{price}_{\text{DRG138/139,1988}} = \frac{\text{price}_{\text{DRG 138,1988}} * \text{N}_{\text{DRG 138,1988}} + \text{price}_{\text{DRG 139,1988}} * \text{N}_{\text{DRG 139,1988}}}{\text{N}_{\text{DRG138,1988}} + \text{N}_{\text{DRG 139,1988}}}$$

$$= \frac{.8535 * 35{,}233 + .5912 * 16{,}829}{35{,}233 + 16{,}289} = .7687$$

[22] 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.

where N denotes the number of admissions in the MedPAR sample. I use this formula to calculate prices for the affected DRGs in every year. To evaluate the aggregate impact of the policy change, I assemble a dataset of annual prices for the affected and unaffected DRGs between 1985 and 1991, and estimate the following specification:

$$(1)\ \ln(\text{price})_{dt} = \alpha + \varsigma \text{DRG}_d + \delta \text{year}_t + \gamma \text{affected DRG}_d \bullet \text{post}_t + \varepsilon_{dt}$$

where d indexes DRGs and t indexes years, affected DRG is a dummy variable that equals one for the treatment group (DRGs affected by the policy change), post is an indicator for the years following the policy change (1988-91), and the dimensions of the coefficient vectors are ς (1 x 387), δ (1 x 6), and γ (1 x 1).[23] Note that the affected DRG main effect is absorbed by the inclusion of the DRG fixed effects. The coefficient of interest, γ, captures the average price change for paired DRGs *relative to* single DRGs during the post period. Each observation is weighted by the number of discharges for that DRG-year cell.

The results from specification (1) are displayed in column 1 of Table 3. Column 2 adds a time trend for affected DRGs, and column 3 includes individual DRG trends. The $\hat{\gamma}$ reveal a robust and statistically significant price increase of 7 percent for affected DRGs in the post-shock period. To illustrate the time path of this change, I replace affected•post in specification (1) with individual affected•year dummies. The coefficients on these dummies, graphed in Figure 1, demonstrate that prices for affected DRGs did not display a different trend from prices for unaffected DRGs in the years prior to the shock. This finding supports the contention that the price change was in fact exogenous, and cannot be attributed to different pre-existing trends in costs for the two groups. The relative price increase of 7 percent and the absolute price increase of 1-2 percent (obtained by summing the year and affected•post coefficients) are considerable because they represent pure profits in an industry in which total profit margins are on the order of 1-2 percent.

[23]Of the 95 DRG pairs and 300 single DRGs in place by 1991, 2 pairs are dropped because the age criterion was eliminated one year early for these pairs, and 5 single DRGs are dropped because there were no admissions coded in these DRGs in the MedPAR sample.

6 Decomposing the Price Shock

6.1 The Mechanical Component

To estimate the mechanical component of the price increase, as described in section 3, I replace price for DRG pairs in 1988-1991 with a Laspeyres price index, calculated using the 1987 volumes of young patients in each code as the weights, i.e.

$$\text{Laspeyres price}_{\text{DRG138/139},t} = \frac{\text{price}_{\text{DRG 138},t} \bullet \text{N(young)}_{\text{DRG 138},1987} + \text{price}_{\text{DRG 139},t} \bullet \text{N(young)}_{\text{DRG 139},1987}}{\text{N(young)}_{\text{DRG138},1987} + \text{N(young)}_{\text{DRG 139},1987}}.$$

This fixed-weight index approximates the average price hospitals would have earned in each post-reform year 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.[24]

Estimating specification (1) with this dependent variable produces a coefficient of .046 (.011), implying that .046/.071=65 percent of the aggregate relative price increase is associated with price recalibrations. In Section 7, I use this mechanical component, $\Delta\ln(\text{Laspeyres price})_{dt} = \ln(\text{Laspeyres price})_{dt} - \ln(\text{price})_{d,1987}$, as an instrument for price. Because this instrument includes all of the effects of annual recalibrations, the appendix details the methods used to eliminate the component related to lagged cost growth, leaving only the price increase associated with the policy change. Estimates of

$$\begin{aligned}(2)\ \ln(\text{price})_{dt} = {} & \alpha + \varsigma \text{DRG}_d + \delta \text{year}_t + \gamma \text{affected DRG}_d \bullet \text{post}_t + \\ & \kappa_1 \text{affected DRG}_d \bullet \text{post}_t \bullet \Delta\ln(\text{Laspeyres price})_{dt} + \varepsilon_{dt}\end{aligned}$$

are given in Table 3, column 4. Columns 5 and 6 present results with an affected DRG trend and individual DRG trends, respectively. Rather than pool the affected DRGs into one treatment group, as in specification (1), specification (2) exploits the fact that the policy change imposed different mechanical price increases for each affected DRG. The positive and significant estimates of κ_1 indicate

[24] During the post-policy period, the correlation between $\text{fraction(old)}_{dt}$ and $\text{fraction(young)}_{dt}$ is .94.

that this refined policy variable captures the differences in the treatment across the affected DRGs. This variable will permit more precise estimates of the elasticity of intensity with respect to price.

6.2 The Upcoding Component

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 affected DRGs, 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. Because all older patients were assigned to the top codes during the pre-shock years, upcoding older patients is the easier of the two options; a hospital assigning a large proportion of older patients to the top codes following the policy change could argue that its older patients had always been relatively complicated. After all, it was not necessary to code complications for older patients during the pre-shock period, so a comparison of pre/post behavior would not be conclusive. Upcoding among the young requires *shifting* patients into the top codes, and is therefore easier to detect. For this reason, my identification strategy provides upper and lower bounds for upcoding among the young, but only lower bounds for upcoding among the old.

6.2.1 Aggregate Upcoding Analysis

The dependent variable for this analysis is $fraction_{dt}$, the share of admissions to pair d in year t that is assigned to the top code in that pair. Because this variable can only be defined for DRG pairs, single DRGs cannot serve as a control group. For young patients, time-series identification is a possibility; a discrete jump in the fraction of patients coded with complications after 1988 suggests an upcoding response to the classification change. However, confounding factors such as an increasing trend in the true severity of patients' conditions may also generate increases in $fraction_{dt}$. For old patients, it is impossible to use the time-series decline in fraction_{dt} to estimate the upcoding response because the magnitude of the decline that would have occurred in the absence of upcoding cannot be determined. I therefore introduce a new independent variable,

$$\text{spread}_{dt} = \text{DRG weight in top code}_{dt} - \text{DRG weight in bottom code}_{dt}\,,$$

e.g. $\text{spread}_{\text{DRG 138/139, 1988}} = \text{weight}_{\text{DRG 138, 1988}} - \text{weight}_{\text{DRG 139, 1988}}$

$= .8535 - .5912 = .2623.$

spread_{dt} is simply a measure of the upcoding incentive in pair d at time t. Between 1987 and 1988, mean spread increased by .20, approximately \$875.[25] The standard deviation of this increase was .16, however, indicating substantial variation in spread changes across DRGs. In the regression

$$(3)\ \ \text{fraction}_{dt} = \alpha + \varsigma \text{DRG}_d + \delta \text{year}_t + \psi \Delta\text{spread}_{d,88\text{-}87} \bullet \text{post}_t + \varepsilon_{dt}$$

δ captures the *average* impact of the policy reform on all DRGs, while ψ captures the *marginal* effect of differential upcoding incentives. $\hat{\psi} > 0$ signifies that hospitals upcoded more in DRGs where the incentive to do so increased more. The estimation results for equation (3) are reported separately by age group in Table 4. For older patients, I include $\text{fraction(young)}_{d,87}$•post as an estimate of the underlying complication rate in each DRG pair.[26]

Table 4 reveals that upcoding is sensitive to changes in spread, even after controlling for the large average increase in spread between 1987 and 1988. As hypothesized, the upcoding response appears to be larger for older patients: the coefficient estimates imply a spread-induced increase of .022 in the fraction of old patients coded with CC, as compared to .015 for younger patients. The year coefficients indicate that the fraction of older patients assigned to the top code declined in 1988 as expected, but this decline was least where the incentive to retain patients in the top code was greatest. Young patients experienced an increase of .02 in reported complications between 1987 and 1989;

[25] This dollar amount is based on P_h for urban hospitals in 2001, which was \$4,376.
[26] An alternative specification using $\Delta\text{spread}_{d,88\text{-}87}$•post as an instrument for spread_{dt} yields similar results.

however, due to the strong upward trend in fraction_{dt} throughout the study period, this increase cannot be unequivocally attributed to the policy change.

The spread-related upcoding alone translates into a price increase of .7 percent and .9 percent for young and old patients admitted to DRG pairs, respectively.[27] 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. These figures imply increased *annual* payments of $330 to $425 million, a substantial reward for altering coding practices.[28] Table 5 summarizes the contributions of the various components of the total relative price increase for DRG pairs. The severity component is the residual remaining after the mechanical and upcoding components are taken into account. Table 5 shows that the vast majority of the relative price increase for DRG pairs can be attributed to HCFA's recalibration errors, despite the large and costly upcoding response. Thus, the policy-induced price change remains an excellent instrument for DRG price even after the upcoding and severity components are removed.

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

6.2.2 Hospital Upcoding 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 non-teaching

---

[27] These estimates are calculated using the average spread in 1988 (.45), together with the average weights for DRG pairs in 1987 (1.05 for young patients, 1.13 for older patients).

[28] These estimates are conservative because upcoding among the old is underestimated. This is likely to be important both because the old account for 70 percent of Medicare admissions, and because upcoding is more prevalent in this group. Dollar figures are calculated using PPS expenditures in 2000.

hospitals.[29] I also consider stratifications by ownership type (for-profit, not-for-profit, government), financial status, region, size, and market-level Herfindahl index.[30]

Table 6 presents estimates of δ and ψ by hospital ownership type, financial status, and region. Figure 2 plots the $\hat{\delta}$ from these specifications. For both young and old patients, there are no statistically significant differences in the response to Δspread across the hospital groups. The discussion here is therefore limited to 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 facilities following the 1988 reform. Consistent with the incentive provided to some for-profit managers to *globally* code more patients with complications, the heightened for-profit response is manifested in the time-series increase in fraction_{dt}, *not* in the spread coefficient. Figure 2 illustrates that upcoding trends were the same for all three ownership types until 1987, but thereafter the trend for for-profits diverges substantially. By 1991, the fraction of young patients with complications had risen by .18 in for-profit hospitals, compared with ~.13 for the other two groups. Given a universal mean of .65 in 1987, these figures are extremely large.

Hospitals with high debt:asset ratios and hospitals in the South also exhibited very large increases in fraction, although Figure 2 illustrates that these trends pre-date 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.

To summarize, HCFA's decision to increase the difference between the prices for complicated and uncomplicated patients with the same diagnosis unleashed a substantial upcoding response. I

[29] One alternative to this approach is to disaggregate the data into hospital-DRG-year cells, and to estimate the following equation separately for each subset of hospitals: $\text{fraction}_{hdt} = \alpha + \mu\text{hospital}_h + \varsigma\text{DRG}_d + \delta\text{year}_t + \psi\Delta\text{spread}_{d,88\text{-}87}\cdot\text{post}_t + \varepsilon_{hdt}$, where hospital_h is a vector of hospital fixed effects. However, since the independent variable of interest varies at the drg-year level, using drg-year cells is the more conservative approach. Moreover, the size of the dataset precludes estimation of an analogous equation for the intensity response (section 7.2), hence for consistency I employ specification (3).

[30] The Herfindahl index is calculated as the sum of squared market shares for all hospitals within a health service area. I constructed two such measures, one using the health service areas reported in the AHA data and another using the health service areas defined by the Dartmouth Atlas on Health Care (1996).

estimate that upcoding generated by the 1988 recalibration alone increased the average price for patients in DRG pairs by approximately one percent. These estimates come from an especially robust and comprehensive empirical investigation; I study not only the time-series response to an unanticipated policy reform, but also differential responses across 93 DRG pairs.

7 Intensity and Volume Responses

Given that HCFA's recalibration mistakes following the 1988 policy change resulted in substantial mechanical price increases for affected DRGs, intensity and volume responses to price changes can be identified using a differences-in-differences specification. If the flypaper effect operates in this setting, intensity levels should rise in affected DRGs relative to unaffected DRGs after 1988. Furthermore, this response should be greater in those DRGs subjected to larger mechanical price increases.

I use five different measures of intensity and quality of patient care to investigate this response: total costs (=total charges from MedPAR deflated by annual cost:charge ratios from the Cost Reports and converted to $1990 using the hospital services CPI), length of stay, number of surgeries, number of ICU days, and in-hospital deaths. All variables are normalized by the number of admissions in the relevant cell (i.e., average cost per patient in "DRG" 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

[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.

was of greater concern in the 1980s, as lengths of stay fell dramatically in response to PPS.) However, the consistency of the aggregate results across all of the variables suggests that the findings are robust.

Given the model outlined in section 2, another way to identify an intensity response is to look at the volume of patients admitted. If hospitals do increase intensity of care within affected DRGs, they should also admit more patients in these DRGs. Stated another way, hospitals seeking to increase volume in affected DRGs following the price shock must increase their investment in intensity.[32]

7.1 Aggregate Intensity and Volume Responses

To examine the effect of the policy change on intensity and volume, I estimate the same specification used to study the effect of the policy change on price, replacing ln(price) with ln(intensity) or ln(admissions):

$$\begin{aligned}(4)\ \ln(\text{intensity or admissions})_{dt} &= \alpha + \varsigma \text{DRG}_d + \delta \text{year}_t + \gamma \text{affected DRG}_d \bullet \text{post}_t \\ &+ \kappa_2 \text{affected DRG}_d \bullet \text{post}_t \bullet \Delta \ln(\text{Laspeyres price})_{dt} + \varepsilon_{dt}.\end{aligned}$$

γ captures the average response to the policy change, while κ_2 allows this response to vary with the magnitude of the mechanical price increase. To ensure that $\hat{\gamma}$ and $\hat{\kappa}_2$ are not capturing pre-existing trends in intensity or volume, I again estimate this specification with separate affected DRG•year dummies in place of affected DRG•post, an affected DRG trend, and finally individual DRG time trends. For each dependent variable, the results from the latter specification are reported in Table 7, in the row labeled "Reduced Form."[33]

I find no evidence that hospitals altered their treatment policies or increased their admissions differentially for patients in affected DRGs as a result of the 1988 classification change. The

---

[32] Note that advertising can certainly be one component of intensity, although I do not have data on such expenditures.

[33] Observations with a value of zero for the unlogged dependent variable are dropped. Regressions of 1(intensity>0) reveal no relationship with the year and affected•year dummies; hence, Tobit estimates using the unlogged dependent variables did not differ from OLS results for the same specifications. The interpretation is that there are some DRGs for which an intensity measure is typically zero, such as death rate in the DRG for tonsillectomy, and excluding such DRGs from the intensity analyses does not affect the estimation.

nonresponse appears to be uniform across affected DRGs, regardless of the size of the mechanical price increase. The point estimates of κ_2 are fairly small, statistically insignificant, and of the wrong sign for 3 of the 6 equations. The corresponding IV estimates of the elasticity of intensity and volume with respect to price ($= \hat{\kappa}_2 / \hat{\kappa}_1$ from Table 3, column 6) are therefore small and imprecisely estimated. To obtain upper bounds for these elasticities, I run OLS regressions of intensity on price. These estimated elasticities, also reported in Table 7, are upward-biased due to the price recalibration method.[34]

Notwithstanding this bias, the point estimates are extremely small. For example, the OLS estimate indicates that only 13 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 (.18) is similar to the estimate reported in Cutler (1990) (.23), but the elasticity of surgeries is much smaller (-.03 as compared to .23), and there is no evidence that in-hospital mortality rates decline in price, as reported in Cutler (1995). Overall, Table 7 suggests that the flypaper effect is very weak in this sector.

7.2 Intensity and Volume Responses Across Hospitals and DRGs

7.2.1 Responses Across Hospitals

The aggregate analysis captures the average intensity and volume responses across all admissions, but masks potentially different responses across hospitals. According to the model defined in section 2.2, hospitals with stronger profit objectives and/or more quality-elastic demand should increase intensity (and therefore volume) more in response to price increases. To determine whether individual hospitals responded differently, I estimate equation (4) separately for the various hospital subsamples. Individual DRG trends are included in all analyses to control for differences in underlying trends across DRGs.

The results provide little evidence of real responses to price increases during the study period. Due to the large volume of coefficients generated by these models, tables are not included here. Out of

[34] One manifestation of this bias is the positive estimated elasticity of death rate with respect to price; the explanation for this puzzling result is simply that those DRGs that experience increases in death rates receive higher reimbursements because in-hospital care for the dying is very expensive.

126 regressions (6 dependent variables * 21 subsamples), $\hat{\kappa}_2$ is statistically significant in only 9, and in most of these cases, the responses are not consistent across the various intensity measures.[35] The sole exceptions are large hospitals (300+ beds) and teaching hospitals (90 percent of which have more than 300 beds). Hospitals in these subsamples increased relative ICU days substantially in response to the price increases. The IV estimates of ICU elasticity are 1.01 (.451) and 1.47 (.63) for large hospitals and teaching hospitals, respectively. The elasticity of total costs with respect to price is correspondingly positive and statistically significant for these samples as well: .49 (.20) and .33 (.17), respectively. There was no significant volume response for these or any other subset of hospitals; indeed, the sign of the volume response was negative in 125 of the 126 specifications.

Section 2 offers several possible explanations for the scant evidence of real responses to the very real price increases documented in Table 3. The arguments focus on the potential inability of hospitals to alter intensity at the DRG level, and of patients in turn to respond. The quality elasticity of demand is paramount in generating an intensity response, and there is reason to believe that this elasticity is extremely low for certain diagnoses. For example, even if hospitals invest in improving care for amputees, these investments are unlikely to yield additional volunteers for the surgery. This reasoning suggests that it may be more fruitful to examine intensity responses separately by DRG type.

7.2.2 Responses Across DRGs

All admissions in the MedPAR files are assigned to one of 5 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). To see how intensity and volume responses differ across these admission types, I assign each DRG to the group accounting for the plurality of its admissions in 1987. I then perform both stages of the intensity analysis (equations 2 and 4) separately by group. The elasticity estimates ($= \hat{\kappa}_2 / \hat{\kappa}_1$) are reported in Table 8.

[35] Tables are available upon request.

Again, the intensity and volume responses are fairly weak. Notwithstanding the strong financial incentive to attract more patients in affected DRGs, hospitals did not increase volume differentially for affected DRGs in *any* admission category following the 1988 relative price increase. (The lack of a volume response also implies that hospitals neglected to upcode *across* DRGs by shifting patients from unaffected to affected DRGs.) However, the estimated intensity elasticities are largest – and in one case, statistically significant – for elective diagnoses. The point estimates provide suggestive evidence that hospitals channeled extra funds to these quality-elastic admissions, but were not rewarded with additional patients or improved outcomes.

Overall, there is little robust evidence of real responses to the changes in DRG prices documented in sections 5 and 6. There is no evidence of volume responses in any subsample of the data, indicating that concerns about hospitals "pushing" certain procedures are unfounded during this time period. To the extent that an intensity response occurred, it was concentrated in elective diagnoses, where patients are likeliest to respond, and in large and/or teaching hospitals, whose operations are more conducive to fine-tuning at the diagnosis level.

7.3 Why Didn't Hospitals Respond?

Given the simultaneous price increase for top codes and decrease for bottom codes within DRG pairs, one possibility is that hospitals may not have realized they were receiving a relative price increase for the pairs as a whole.[36] Even if hospitals were cognizant of the price increase in affected DRGs, their response may have been muted because of the simultaneous price decrease in unaffected DRGs. The net result was that average prices for all admissions did not increase by much. A positive intensity response would therefore involve a decrease in intensity for unaffected DRGs, and to the extent that decreases are more difficult to implement than increases, the coefficients I obtain may underestimate the true intensity-price relationship. This explanation, though certainly a possibility, is by no means a

[36] I thank David Cutler for this insight.

certainty: immediately following the implementation of PPS, hospitals showed themselves quite capable of reducing overall intensity in all of the dimensions I explore.

Another possibility is that hospitals optimize overall intensity, rather than intensity by DRG. 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

$$(5)\ \ln(\text{intensity})_{ht} = \alpha + \mu \text{hospital}_h + \delta \text{year}_t + \beta \ln(\text{price})_{ht} + \varepsilon_{ht}.$$

However, there are two sources of bias in the OLS estimate of $\hat{\beta}$: (1) the DRG recalibration method; (2) the omission of an annual hospital-level measure of patient severity. As with the previous analyses, the policy change can be used to identify β, but hospital-level variation in the impact of the policy change is required – a differences-in-differences strategy comparing affected and unaffected DRGs cannot be implemented with hospital-year data. Because hospitals with a large fraction of admissions in the "with CC" DRGs benefited the most from the policy reform, the interaction between this measure and a dummy for the post-reform years can serve as an instrument for average price in equation (5).[37]

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 only because the data do not indicate whether old patients had CC before the policy change. This instrument captures the mechanical, or exogenous, 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 and any change in the true severity of

[37] An alternative instrument for hospital price is $\Delta\ln(\text{Laspeyres price})_{ht} = \ln(\text{Laspeyres price})_{ht} - \ln(\text{price})_{h,1987.}$ However, because the actual DRGs sampled for each hospital varies 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.

patients. Eliminating the upcoding and severity components from the instrument ensures that the IV estimate will be unbiased even if these components are associated with intensity decisions.

Table 9 gives the results from the first-stage regression of ln(price) on *share CC*•post,

$$(6)\ \ln(\text{price})_{ht} = \alpha + \mu \text{hospital}_h + \delta \text{year}_t + \tau_1 \text{share}\,\text{CC}_h \bullet \text{post}_t + \varepsilon_{ht},$$

where hospital_h is a vector of hospital fixed effects. The mean (standard deviation) of *share CC* in 1987 is .086 (.043). A two-standard-deviation increase in *share CC* is associated with a two percent increase in the average price paid to a hospital following the policy change. To illustrate that *share CC* is uncorrelated with average hospital prices in the pre-reform years (after hospital fixed effects are included), column 2 presents the results from a regression of ln(price) on *share CC* •year dummies.

Coefficient estimates from the reduced-form equation,

$$(7)\ \ln(\text{intensity})_{ht} = \alpha + \mu \text{hospital}_h + \delta \text{year}_t + \tau_2 \text{share}\,\text{CC}_h \bullet \text{post}_t + \varepsilon_{ht},$$

are presented under "Reduced Form" in Table 10, followed by IV and OLS estimates of equation (5). The IV estimates for the elasticity of hospital intensity with respect to average hospital price are positive for 4 of the 5 intensity measures, and statistically significant for 3. 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.

Hospitals subjected to price increases not only increased their intensity of care, but also succeeded in drawing in additional patients: for every one-percent increase in price, total admissions increased by 1.7 percent. This volume response also explains the large, positive coefficient on in-hospital mortality: if greater intensity of care attracts sicker patients, as posited in section 2, outcomes may actually deteriorate.

The intensity results in Table 9 are consistent with two distinct models of hospital behavior: competition in overall intensity, and maximization of overall intensity subject to a budget constraint.

The fact that volume responds to increases in intensity provides a motive for the former, but does not rule out the latter. 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.

8 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. If hospitals respond to these incentives, we may see them encouraging certain types of admissions and discouraging others, a practice that could be innocuous in other fixed-price industries (e.g., utilities), but is potentially dangerous in this setting. 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 (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.[38] Responses to individual price changes can reveal whether this competition occurs at the level of the DRG, or product line.

This study illustrates how a simple change in the DRG classification system in 1988 generated large and exogenous relative price increases 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. Finally, while all subsamples of hospitals upcoded more following the policy change, for-profit facilities availed themselves of this opportunity to the greatest extent.

Whereas coding behavior proved very responsive to financial incentives, admissions and treatment policies did not. Using a differences-in-differences identification strategy, I find no evidence of a relative increase in admissions to DRGs subjected to price increases, and very limited evidence of increases in intensity of care. However, I find strong evidence that hospitals spent the extra funds they received on patient care in *all* DRGs. This finding suggests that hospitals do not (or cannot) optimize intensity choices by product line, and may compete instead in overall quality levels.[39]

These results may help to 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 were relatively cost-efficient. If, however, hospitals do not balance costs and benefits within individual product lines, such specialization is unlikely to occur. Another implication of these results is that insurers may be unable to use prices to encourage quality improvements in specific diagnoses. More generally, this research suggests that better models of hospital behavior are necessary for anticipating the impacts of public and private-sector actions in this important industry.

[38] Of course, physicians also play an important role in ensuring appropriate care for their patients, as highlighted by Arrow (1963).
[39] Previous studies have also found a positive relationship between *overall* hospital intensity and financial pressure; see footnote 14.

References

Arrow, Kenneth J., 1963, "Uncertainty and the Welfare Economics of Medical Care," American Economic Review 53, No. 5: 941-973.

Bulow, J., Geanakoplos, J., and P. Klemperer, 1995, "Multimarket Oligopoly: Strategic Substitutes and Complements," Journal of Political Economy, 93: 488-511.

Carter, Grace M., Newhouse, Joseph P., and Daniel A. Relles, 1990, "How Much Change in the Case Mix Index is DRG Creep?" Journal of Health Economics, 9(4): 411-428.

Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 1996. The Dartmouth Atlas of Health Care (Chicago: American Hospital Publishing).

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

Cutler, David M., 1990, "Empirical Evidence on Hospital Delivery Under Prospective Payment," MIT mimeo.

Cutler, David M., 1995, "The Incidence of Adverse Medical Outcomes under Prospective Payment," Econometrica 63, No. 1: 29-50.

Cutler, David M., 1998, "Cost Shifting or Cost Cutting?: The Incidence of Reductions in Medicare Payments," in J. Poterba, ed., Tax Policy and the Economy, Volume 12, Cambridge, MA: MIT Press.

Cutler, David M. and Jill R. Horwitz, 1999. "Converting Hospitals from Not-for-Profit to For-Profit Status: Why and What Effects?" in D. Cutler, ed., The Changing Hospital Industry: Comparing Not-for-Profit and For-Profit Institutions, Chicago: University of Chicago Press and NBER.

Data Compendium, 2002. Baltimore, MD.: U.S. Dept. of Health and Human Services, Health Care Financing Administration, Bureau of Data Management and Strategy.

Dranove, David, 1987, "Rate-Setting by Diagnosis Related Groups and Hospital Specialization," RAND Journal of Economics 18: 417-427.

Duggan, Mark, 2000, "Hospital Ownership and Public Medical Spending," Quarterly Journal of Economics, 115(4): 1343-1373.

Duggan, Mark, 2002, "Hospital Market Structure and the Behavior of Not-for-Profit Hospitals," Rand Journal of Economics, Autumn.

Eichenwald, Kurt, 2003, "How One Hospital Benefited on Questionable Operations," The New York Times (August 12).

Ellis, Randall P. and Thomas McGuire, 1996, "Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects," Journal of Health Economics, 15: 257-277.

Federal Register, 1984-2000. U.S. Office of the Federal Register, National Archives and Records Service, General Services Administration, Washington, D.C.:U.S. Government Printing Office.

Gilman, Boyd H., 2000, "Hospital Response to DRG Refinements: The Impact of Multiple Reimbursement Incentives on Inpatient Length of Stay," Health Economics, 9: 277-294.

Hadley, Jack, Zuckerman, Stephen, and Judith Feder, 1989, "Profits and Fiscal Pressure in the Prospective Payment System: Their Impacts on Hospitals," Inquiry, 26(3): 354-365.

Hodgkin, Dominic and Thomas G. McGuire, 1994, "Payment Levels and Hospital Response to Prospective Payment," Journal of Health Economics, 9: 1-29.

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

Newhouse, Joseph P., 1989, "Do Unprofitable Patients Face Access Problems?" Health Care Financing Review, 11(2): 33-42.

Psaty, Bruce M., Robin Boineau, Lewis H. Kuller, and Russell V. Luepker, 1999, "The Potential Costs of Upcoding for Heart Failure in the United States," The American Journal of Cardiology, (July 1) 84:108-109.

Silverman, Elaine and Jonathan Skinner, 1999. "The Association between For-Profit Hospital Ownership and Increased Medicare Spending," New England Journal of Medicine, 341(6): 420-426.

Silverman, Elaine and Jonathan Skinner, 2000. "Are For-Profit Hospitals Really Different? Medicare Upcoding and Market Structure," NBER Working Paper W8133.

Staiger, Douglas and Gary L. Gaumer, "The Impact of Financial Pressure on Quality of Care in Hospitals: Post-Admission Mortality Under Medicare's Prospective Payment System," Working Paper prepared for the Health Care Financing Administration. Cambridge, MA. Abt Associates, Inc.

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

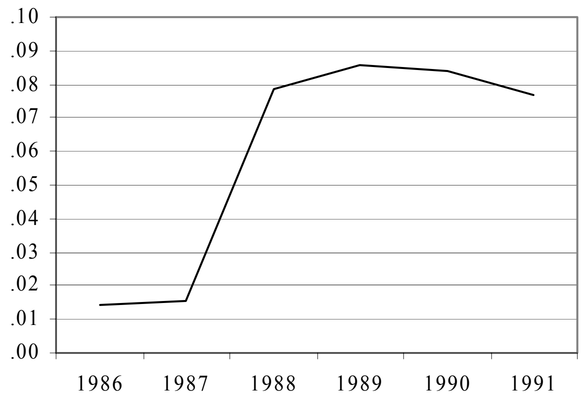
Statistical Abstract of the United States, 2001, 1998, 1991, 1987. U.S. Census Bureau, Administrative and Customer Services Division, Statistical Compendia Branch. Washington, D.C.: U.S. Government Printing Office.

Appendix

The total change in the price paid to hospitals for admissions to affected DRGs following the elimination of the age criterion can be subdivided into three components: *mechanical, upcoding*, and *severity*. As noted in the text, the mechanical component is the effect of the recalibration on prices, holding the incidence of complications constant – essentially, it captures mistakes made by HCFA in its recalibration. To estimate this component, I construct a Laspeyres price index for each paired DRG in each year after 1987 using 1987 admissions of young patients as the fixed weights. I then subtract the 1987 price to obtain a measure of the price increase in each DRG-year, holding constant the fraction of patients with CC. This measure (denoted $\Delta\ln(\text{Laspeyres price})_{dt}$ in the text) incorporates *all* of the price changes associated with the annual recalibrations after 1987 – i.e., that due to the policy change, and that due to differences across DRGs in annual charge growth. To eliminate the latter from this instrument, I regress the instrument on the lagged charges actually used in the updating process, and use the residuals as the final instrument. Because HCFA operates with a 2-year lag for updating, this process amounts to the following:

for each t>87, regress $\Delta\ln(\text{Laspeyres price})_{dt}=(\ln(\text{Laspeyres price})_{dt} - \ln(\text{price})_{d,1987})$ on $(\ln(\text{charges})_{d,t-2} - \ln(\text{charges})_{d,85})$ and a constant. The residuals from each regression constitute the final instrument.

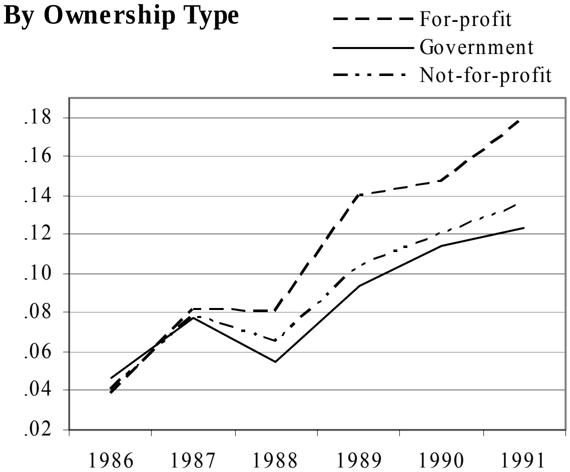
Figure 1. Effects of Policy Change on Relative Prices for Affected DRGs, By Year



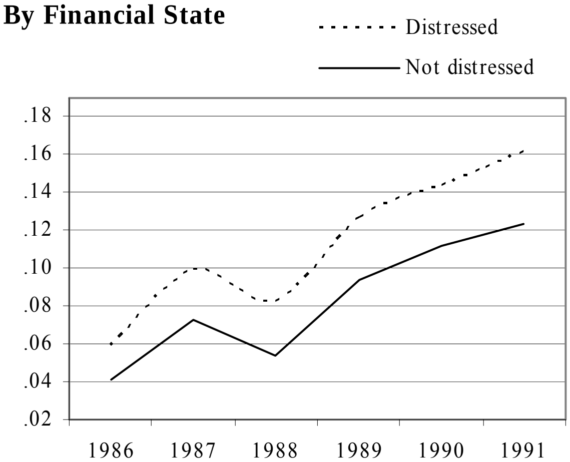
Notes: Estimates of γ from $\ln(\text{price})_{dt} = \alpha + \varsigma DRG_d + \delta year_t + \gamma \text{affected } DRG_d \bullet year_t + \varepsilon_{dt}$

Figure 2. Effect of Policy Change on Upcoding of Young, by Hospital Characteristics

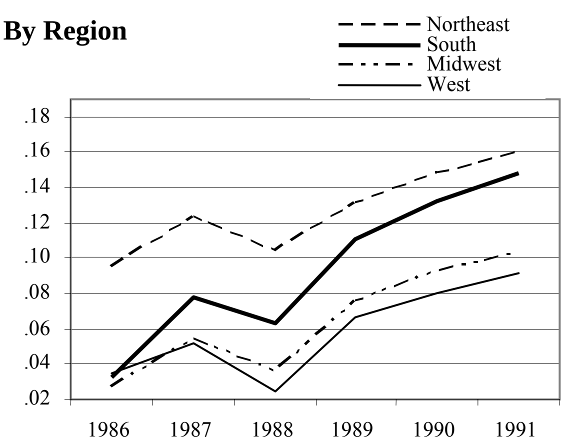
By Ownership Type



By Financial State



By Region



Source: Year coefficients from Table 6.

Table 1. Examples of Policy Change

DRG code	Description in 1987 (Description in 1988)	1987 weight	1988 weight	% change in weight	1987 volume (20% sample)	1988 volume (20% sample)	% change in volume
96	bronchitis and asthma age>69 and/or CC (bronchitis and asthma age>17 with CC)	.8446	.9804	16%	44,989	42,314	-6%
97	bronchitis and asthma age 18-69 without CC (bronchitis and asthma age>17 without CC)	.7091	.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)	.8136	.8535	5%	45,080	35,233	-22%
139	cardiac arrhythmia and conduction disorders age<70 without CC (cardiac arrhythmia and conduction disorders without CC)	.6514	.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)	.8271	.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)	.6984	.5791	-17%	2,033	12,363	508%

Notes: Of the 95 affected pairs, these three occur most frequently in the 1987 20% MedPAR sample.

Table 2. Descriptive Statistics

Unit of Observation	**DRG-year**			**Hospital-year**		
	N	Mean	SD	N	Mean	SD
price (DRG weight)	2482	1.26	(.91)	36651	1.27	(.19)
Laspeyres price	2482	1.25	(.90)			
observations per cell	2482	6128	(12817)	36651	373	(389)
Nominal responses						
fraction(young) in top code	650	.66	(.14)			
fraction(old) in top code	650	.85	(.15)			
Real Responses						
mean cost ($)	2474	6000	(4889)	36169	6450	(3005)
mean LOS (days)	2482	10.64	(5.64)	36651	8.81	(2.21)
mean surgeries	2450	1.15	(.73)	35897	1.21	(.55)
mean ICU days	2290	.72	(1.18)	28226	.81	(.59)
death rate	2123	.07	(.10)	34992	.06	(.02)
mean admissions	2482	32921	(30981)	36651	778	(538)
Instruments						
1988 spread-1987 spread	650	.20	(.16)			
(1988 spread-1987 spread) •post	650	.12	(.16)			
affected	2482	.45	(.50)			
affected•post	2482	.26	(.44)			
Δln(Laspeyres price)	368	.01	(.05)			
affected•post•Δln(Laspeyres price)	2478	.00	(.03)			
share CC				36651	.09	(.03)
share CC•post				36651	.05	(.05)

Notes: Nominal responses are calculated for DRG pairs only. Means are weighted by the number of observations in the 20 percent MedPAR sample, with the exception of observations per cell.

Table 3. Total Effect of Policy Change on DRG Prices

	Dependent Variable is ln(price) mean(price) = 1.26					
Affected•post	.071 ***	.066 ***	.065 ***	.062 ***	.064 ***	.064 ***
	(.012)	(.016)	(.013)	(.011)	(.014)	(.013)
Affected•post•Δln(Laspeyres price)				1.233 ***	1.234 ***	.629 ***
				(.092)	(.092)	(.124)
Year dummies						
1986	-.017	-.018	-.009	-.017	-.017	-.008
	(.014)	(.015)	(.012)	(.013)	(.014)	(.012)
1987	-.017	-.018	.000	-.017	-.016	.001
	(.014)	(.015)	(.019)	(.013)	.015	(.019)
1988	-.049 ***	-.049 ***	-.019	-.048 ***	-.048 ***	-.019
	(.015)	(.015)	(.027)	(.015)	(.014)	(.027)
1989	-.045 **	-.045 **	-.005	-.043 **	-.043 **	-.004
	(.016)	(.016)	(.035)	(.015)	(.016)	(.035)
1990	-.055 ***	-.057 ***	-.006	-.056 ***	-.056 ***	-.006
	(.016)	(.017)	(.044)	(.015)	(.017)	(.044)
1991	-.061 ***	-.062 ***	-.001	-.062 ***	-.062 ***	-.000
	(.016)	(.018)	(.053)	(.015)	(.018)	(.052)
DRG fixed effects	Y	Y	Y	Y	Y	Y
Affected DRG trend	N	Y	N/A	N	Y	N/A
DRG trends	N	N	Y	N	N	Y
Adj. R-squared	.977	.977	.990	.979	.979	.990
N	2482	2482	2482	2478	2478	2478

Notes: The unit of observation is DRG-year (where "DRG" refers to single DRGs as well as to DRG pairs). All observations are weighted by the number of admissions in the 20% MedPAR sample. The sum of the weights is 15.2 million. Standard errors are robust.

* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 4. Effect of Policy Change on Upcoding

	fraction(young) mean = .66	**fraction(old)** mean = .85
$\Delta spread_{88\text{-}87}$•post	.077 ***	.108 ***
	(.016)	(.015)
fraction(young)$_{87}$•post		.731
		(.020) ***
Year dummies		
1986	.044 ***	.000
	(.008)	(.005)
1987	.077 ***	-.011 *
	(.008)	(.005)
1988	.058 ***	-.813 ***
	(.011)	(.014)
1989	.097 ***	-.780 ***
	(.009)	(.014)
1990	.115 ***	-.764 ***
	(.009)	(.014)
1991	.128 ***	-.748 ***
	(.010)	(.014)
Adj. R-squared	.948	.960
N	650	650

Notes: Regressions include DRG fixed effects. "Young" refers to Medicare beneficiaries under 70; "Old" refers to beneficiaries aged 70+. The unit of observation is DRG-year. Single DRGs are not included. All observations are weighted by the number of admissions in the 20% MedPAR sample. The sum of the weights is 1.9 million (young) and 5.0 million (old). Standard errors are robust.
* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 5. Decomposition of Price Change for Affected DRGs

	Total Price Change	Mechanical Component	Upcoding Component	Severity Component
Conservative estimate	7.1%	4.6%	0.8%	1.7%
Percent of total	*100.0%*	*64.8%*	*11.7%*	*23.5%*
Liberal estimate	7.1%	4.6%	1.1%	1.4%
Percent of total	*100.0%*	*64.8%*	*14.9%*	*20.3%*

Table 6. Effect of Policy Change on Upcoding of Young, by Hospital Characteristics

	By Ownership Type			By Financial State		By Region			
	For-profit	Not-for-profit	Government	Distressed	Not distressed	Northeast	Midwest	South	West
$\Delta spread_{88\text{-}87} \bullet post$	.071 **	.080 ***	.058 ***	.082 ***	.074 ***	.083 ***	.062 ***	.082 ***	.079 ***
	(.024)	(.017)	(.018)	(.109)	(.017)	(.016)	(.018)	(.017)	(.024)
Year fixed effects									
1986	.038 ***	.047 ***	.040 ***	.059 ***	.041 ***	.095 ***	.027 **	.033 ***	.035 ***
	(.009)	(.009)	(.008)	(.008)	(.009)	(.008)	(.009)	(.009)	(.012)
1987	.081 ***	.077 ***	.078 ***	.099 ***	.073 ***	.123 ***	.054 ***	.078 ***	.052 ***
	(.010)	(.008)	(.008)	(.008)	(.008)	(.007)	(.009)	(.008)	(.012)
1988	.080 ***	.055 ***	.065 ***	.083 ***	.054 ***	.104 ***	.036 ***	.063 ***	.024 ***
	(.012)	(.011)	(.012)	(.011)	(.011)	(.010)	(.010)	(.012)	(.014)
1989	.140 ***	.094 ***	.104 ***	.127 ***	.094 ***	.131 ***	.075 ***	.111 ***	.067 ***
	(.011)	(.009)	(.009)	(.009)	(.009)	(.009)	(.010)	(.009)	(.012)
1990	.147 ***	.114 ***	.120 ***	.144 ***	.112 ***	.148 ***	.092 ***	.132 ***	.080 ***
	(.011)	(.009)	(.010)	(.009)	(.010)	(.008)	(.009)	(.010)	(.013)
1991	.179 ***	.123 ***	.136 ***	.161 ***	.124 ***	.159 ***	.103 ***	.148 ***	.091 ***
	(.011)	(.011)	(.010)	(.010)	(.010)	(.010)	(.011)	(.010)	(.014)
$\hat{\delta}_{89} - \hat{\delta}_{87}$	.059 ***	.016	.027 ***	.027 **	.022 *	.007	.021 *	.033 ***	.015
	(.010)	(.009)	(.008)	(.009)	(.009)	(.008)	(.010)	(.008)	(.014)
Adj. R-squared	.914	.946	.927	.933	.947	.939	.940	.940	.915
N	650	650	650	650	650	650	650	650	650

Notes: Regressions include DRG fixed effects. "Young" refers to Medicare beneficiaries under 70. "Distressed" denotes hospitals with 1987 debt:asset ratios at the 75th percentile or above. The unit of observation is DRG-year. Single DRGs are not included. All observations are weighted by the number of admissions in the 20% MedPAR sample. Hospitals with missing values for any of the hospital characteristics are dropped. The sum of the weights is 1.45 million. Standard errors are robust.
* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 7. Real Responses to Changes in DRG Prices

	Dependent Variable					
	ln(cost)	ln(LOS)	ln(surg)	ln(ICU)	ln(death rate)	ln(volume)
	mean=$5,995	mean=10.63	mean=1.15	mean=.72	mean=.07	mean=32,944
Reduced Form						
Affected•post	.007	.012	-.005	-.023	-.019	.040
	(.011)	(.017)	(.016)	(.034)	(.038)	(.025)
Affected•post•	.057	.119	-.094	.395	.619	-.107
Δln(Laspeyres price)	(.118)	(.126)	(.127)	(.271)	(.437)	(.373)
IV Estimate						
ln(price)	.090	.190	-.149	.627	.984	-.171
	(.189)	(.197)	(.194)	(.450)	(.758)	(.604)
Parametric Tests of H_0: IV estimate>=x; H_1: IV estimate<x (p-values are reported)						
x = .5	.02	.06	.00	.62	.03	.13
x = 1	.00	.00	.00	.21	.00	.03
OLS Estimate						
ln(price)	.126 ***	.182 ***	-.029	.253 **	.258 *	-.048
	(.037)	(.043)	(.047)	(.089)	(.115)	(.092)
N	2470	2478	2446	2286	2119	2478

Notes: Regressions include year fixed effects, DRG fixed effects, and DRG trends. Unlogged means are reported. The unit of observation is DRG-year (where "DRG" refers to single DRGs as well as to DRG pairs). All observations are weighted by the number of admissions in the 20% MedPAR sample. The sum of the weights is 15.2 million. For ln(death rate), the tests presented are H_0: IV estimate<=x; H_1: IV estimate>x for x=-.5 and x=-1. Standard errors are robust.
* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 8. Real Responses to Changes in DRG Prices, by DRG Type

	Dependent Variable					
	ln(cost)	ln(LOS)	ln(surg)	ln(ICU)	ln(death rate)	ln(volume)
Emergency DRGs						
IV Estimate						
ln(price)	-.181	.164	-.211	.404	-.031	.003
	(.236)	(.298)	(.260)	(.582)	(.539)	(.528)
N	1334	1334	1324	1255	1186	1334
Urgent DRGs						
IV Estimate						
ln(price)	.193	-.530	.182	-.668	.909	.506
	(.461)	(1.562)	(.533)	(1.605)	(1.517)	(.614)
N	240	245	229	188	161	245
Elective DRGs						
IV Estimate						
ln(price)	.977 *	-.010	.340	1.937	3.550	-.939
	(.451)	(.349)	(.210)	(1.089)	(2.397)	(1.600)
N	896	899	893	843	772	899

Notes: Elasticities are estimated from regressions of the following form:

$\ln(\text{intensity or admissions})_{dt} = \alpha + \varsigma DRG_d + \delta year_t + \omega DRG\ trends_{dt} + \gamma affected\ DRG_d \bullet post_t + \beta \ln(price)_{dt} + \varepsilon_{dt}$

where the instrument for ln(price) is affected $DRG_d \bullet post_t \bullet \Delta\ln(\text{Laspeyres price})_{dt}$.

The unit of observation is DRG-year (where "DRG" refers to single DRGs as well as to DRG pairs). All observations are weighted by the number of admissions in the 20% MedPAR sample. The sum of the weights is 10.9 million (emergency DRGs), .83 million (urgent DRGs), or 3.4 million (elective DRGs). Standard errors are robust.

* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 9. Effects of Policy Change on Average Hospital Prices

Dependent Variable is ln(price) mean(price) = 1.27		
Share CC•post	.233 ***	
	(.021)	
Share CC•year dummies		
1986		-.022
		(.040)
1987		-.015
		(.038)
1988		.229 ***
		(.038)
1989		.212 ***
		(.039)
1990		.174 ***
		(.040)
1991		.270 ***
		(.047)
Year dummies		
1986	.039 ***	.041 ***
	(.001)	(.004)
1987	.057 ***	.058 ***
	(.001)	(.004)
1988	.063 ***	.064 ***
	(.002)	(.004)
1989	.088 ***	.090 ***
	(.002)	(.004)
1990	.094 ***	.099 ***
	(.002)	(.004)
1991	.119 ***	.116 ***
	(.002)	(.004)
Adj. R-squared	.890	.890
N	36,651	36,651

Notes: Regressions include hospital fixed effects. The unit of observation is 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. Share CC•post = (1987 share of a hospital's Medicare patients who are under 70 and assigned to the top code of a DRG pair)•(indicator variable for year>1987). Standard errors are robust.
* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Table 10. Real Responses to Changes in Average Hospital Price

	Dependent Variable					
	ln(cost) mean=$9,014	ln(LOS) mean=8.81	ln(surg) mean=1.21	ln(ICU) mean=.81	ln(death rate) mean=.06	ln(volume) mean=778
Reduced Form						
Share CC•post	.234 ***	.069 *	.067	.684 ***	.122	.403 ***
	(.075)	(.034)	(.104)	(.186)	(.097)	(.052)
IV Estimate						
ln(price)	.998 ***	.296*	.291	3.457 ***	.536	1.728 ***
	(.312)	(.141)	(.445)	(.950)	(.423)	(.276)
Parametric Tests of H_0: IV estimate>=x; H_1: IV estimate<x (p-values are reported)						
x = .5	.96	.06	.31	1.0	.00	1.0
x = 1	.50	.00	.04	1.0	.00	1.0
OLS Estimate						
ln(price)	.769 ***	.350 ***	.867***	1.483 ***	.601 ***	-.022
	(.027)	(.011)	(.036)	(.065)	(.031)	(.018)
N	36,169	36,651	35,897	28,226	34,992	36,651

Notes: Regressions include year fixed effects and hospital fixed effects. Unlogged means are reported. The unit of observation is hospital-year. All observations are weighted by the number of admissions in the 20% MedPAR sample. Share CC•post = (1987 share of a hospital's Medicare patients who are under 70 and assigned to the top code of a DRG pair)•(indicator variable for year>1987). The sum of the weights is 13.7 million. For ln(death rate), the tests presented are H_0: IV estimate<=x; H_1: IV estimate>x for x=-.5 and x=-1. Standard errors are robust.
* signifies p<.05, ** signifies p<.01, *** signifies p<.001

Appendix Table 1. Descriptive Statistics for Hospital Characteristics

Variable	Mean	SD	Min	Max
Ownership				
For-profit	.14	.35	0	1
Non-profit	.58	.49	0	1
Government	.28	.45	0	1
Financial Distress Measures				
Debt:asset ratio	.52	.32	0	2.17
Medicare bite	.37	.11	0	1.00
Medicaid bite	.11	.08	0	.84
Region				
Northeast	.14	.35	0	1
Midwest	.29	.46	0	1
South	.38	.49	0	1
West	.18	.38	0	1
Size				
1-99 beds	.46	.50	0	1
100-299 beds	.37	.48	0	1
300+ beds	.17	.37	0	1
Service Offerings				
Teaching program	.06	.23	0	1
Open heart surgery	.13	.34	0	1
Trauma facility	.19	.39	0	1
ICU beds (except neonatal)	10.26	12.29	0	194
Market Concentration				
HSA Herfindahl (AHA)	.07	.05	0	1
HSA Herfindahl (Dartmouth Atlas)	.65	.36	0	1

Notes: N=5,336. Excludes hospitals with missing values for any of the variables, or with debt:asset ratios in the 1% tails of the distribution.

Sources: HCFA Cost Reports (1987), American Hospital Association Annual Survey of Hospitals (1987)