Volume Title: Advances in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90302-8

Volume URL: http://www.nber.org/books/wise96-1

Conference Date: May 6-9, 1993

Publication Date: January 1996

Chapter Title: Medicare Reimbursement and Hospital Cost Growth

Chapter Author: Mark B. McClellan

Chapter URL: http://www.nber.org/chapters/c7322

Chapter pages in book: (p. 149 - 190)

5 Medicare Reimbursement and Hospital Cost Growth

Mark B. McClellan

In 1984, the Health Care Financing Administration implemented the prospective payment system (PPS) to reimburse hospitals for treating Medicare patients. PPS features payment on the basis of the diagnosis-related group (DRG), an essentially fixed payment for the entire bundle of medical services produced during a hospital admission. DRG outlier payments provide some insurance to hospitals for treating patients with extraordinarily high costs or long stays within the DRG. Other DRG adjustments, such as higher DRG payments for patients with versus without "complicating conditions" and higher payments for teaching hospitals, provide limited opportunities for hospitals to receive additional compensation for higher costs of care. But DRGs have been thought to provide high-powered incentives for efficient production of hospital care: reimbursement for an admission in a given DRG should not change much as input use varies.

In many respects, Medicare's actual experience with hospital costs under PPS has not worked out as hoped. Most troubling has been a continued steady increase in intensity per admission, and consequently in average per capita expenditures for hospital services in the Medicare program. Between 1984 and 1990, the Prospective Payment Assessment Commission (1991) estimated that the case-mix index per discharge increased by almost 20%, after adjusting for price updates and other policy changes. A 1% increase in the case-mix index translates into an additional $750 million per year in Medicare payments; this increase in intensity accounts for the bulk of Medicare hospital expenditure increases. As the Prospective Payment Assessment Commission (1993) has

Mark B. McClellan is professor of economics and medicine at Stanford University and a faculty research fellow of the National Bureau of Economic Research.

The author thanks Jonathan Feinstein, Joseph Newhouse, Andrew Samwick, and Jonathan Skinner for useful discussions, and the Agency for Health Care Policy Research (grant R03-HS07638-01) and the National Institute on Aging (grant 5-R29-AG11706-02) for financial support.

noted, growth in the case mix index is the principal determinant of hospital payment growth; its effect has been "greater than the combined effect of . . . all other policy changes implemented under PPS" (20–21).

From the standpoint of controlling Medicare hospital expenditure growth, the PPS experience has clearly been disappointing. In a related paper (McClellan 1993), I have argued that a few basic implications of the nature of DRG payments, coupled with some understanding of the internal organization of hospital production, can explain the "puzzles" of PPS in the context of a standard analysis of production incentives with price regulation. Here I review and expand the previous analysis, describing some distinctive institutional features of hospital production in more detail and summarizing the results of the static model of hospital production designed to capture these features. I then develop a dynamic framework to describe how price and cost growth interact in hospital production, which suggests that a reimbursement system such as PPS may encourage technological change endogenous to the payment rules. Finally, I turn to some preliminary evidence on the nature of expenditure-increasing technological change in the Medicare program, which appears to be consistent with the dynamic framework.

5.1 Institutional Characteristics of Hospital Production

This section addresses two related topics that have received limited attention in the literature on hospital behavior: the appropriate definition of the "good" produced for patients demanding hospital care, and the nature of the internal organization of hospitals. These issues have important implications for the production and regulation of medical care.

Patients do not demand hospital treatments per se; they demand treatment for their particular health problems. "Production" of health care involves both diagnostic and therapeutic treatments, the most intensive of which are administered to patients who have been admitted to a hospital. For most health problems, many technological choices exist, and hence many alternative treatment intensities are possible. As a result, a hospital admission per se does not constitute a well-defined product. The following examples, drawn from some of the most common and costly diseases in the elderly, illustrate that (1) hospitalizations may or may not be required for treatment; (2) when a patient is hospitalized, multiple treatment courses involving quite different medical technologies are possible; and (3) these treatment courses can result in different payment levels, even in "prospective" reimbursement systems such as Medicare PPS. The brief clinical summaries are not intended to imply that all patients with each of the health problems are equally appropriate candidates for all treatment alternatives. However, the net benefits of more intensive treatment may be modest or uncertain for a considerable fraction of patients with a health problem, providing considerable opportunities for physician discretion in treatment choices (Park et al. 1986; McClellan and Brook 1992).

Coronary heart disease. Coronary heart disease is the leading cause of death in the United States. In most cases, death is directly or indirectly related to a heart attack, or acute myocardial infarction (AMI). Patients with symptoms of a heart attack such as crushing chest pain will be treated at a hospital to "rule out" AMI or to manage its complications. The diagnostic and therapeutic management of all heart disease patients may include several intensive procedures: cardiac catheterization, percutaneous transluminal coronary angioplasty (PTCA), and coronary artery bypass graft (CABG) surgery. These technologies are substantially more costly than the less intensive alternatives, which involve noninvasive diagnostic studies, drug treatments, and lifestyle counseling.

Cancer. Treatment modalities for most types of cancer involve some combination of surgery, radiation therapy, and chemotherapy. For example, for localized breast cancer, treatment may involve total or partial mastectomy, followed by radiation or chemotherapy. Various cancer treatment regimens have different requirements for the intensity of these treatments and for the frequency of inpatient hospital stays.

Hip fracture. Patients with hip fracture will be admitted to the hospital. Depending on the nature of the fracture, the treatment for setting the fracture may involve open reduction (a surgical procedure) or closed reduction (a nonsurgical procedure).

Gall bladder disease. Patients with chronic cholecystitis (gall bladder disease) can be managed "medically," with drugs and dietary modification, or "surgically," with removal of the gall bladder. For patients managed surgically, several types of cholecystectomy procedures are possible. Cholecystectomy may include intraoperative radiographic imaging of the biliary ducts with contrast dye enhancement to identify any remaining gallstones. An alternative cholecystectomy technique for uncomplicated cases, laparoscopic cholecystectomy, involves a more limited surgical incision and potentially shorter postoperative recovery.

Benign prostatic hypertrophy. Enlargement of the prostate gland, leading to urinary retention, frequency, and hesitancy, is extremely common in elderly males. One treatment is prostatectomy (primarily transurethral prostatectomy) to remove gland tissue. Alternatively, especially in milder cases, patients can simply endure symptoms (e.g., having to get up to urinate in the middle of the night). Patients who choose not to have surgery require hospitalization only for severe complications.

Back pain. Back pain is an extremely common chronic health problem. Surgical treatments to reduce symptoms include spinal decompression (for pain, tingling, or numbness arising from nerve compression), intervertebral diskectomy and spinal fusion, and laminectomy. Patients who are managed without these surgical procedures (through drugs, musculoskeletal therapies, and other regimens) may require hospital treatments such as traction only for severe complications.

Arthritis. Rheumatoid arthritis and osteoarthritis are leading causes of functional limitations in the elderly, especially elderly females. Many drug regimens of increasing strength (and side effects), as well as other nonsurgical treatments such as heat baths, are available to reduce pain and limit deterioration in affected joints. These treatments generally do not require hospitalization. A more intensive hospital-based alternative is surgical replacement of the knee or hip.

The preceding illustrations, representing a spectrum of common diseases, demonstrate that any analysis of hospital production should consider incentives for treatment of a health problem rather than for production of a particular kind of treatment. However, virtually all of the alternative treatments just described lead to different DRG classifications. The structure of DRGs consequently supports a form of indirect cost sharing, in which payments depend not on reported costs but on actual use of particular technologies that have important implications for the cost of an admission. In this sense, the DRG system continues a historical trend in hospital payment mechanisms toward dependence on real measures of resource use. Indeed, hospital payments under the Tax Equity and Fiscal Responsibility Act (TEFRA) system that PPS replaced were primarily based on allowed per diems, that is, hospital payments that depended on the number of days provided. While PPS moved further toward an admission-based payment system, it retained some features of intensity-related payment.

Why is PPS structured so that use of specific intensive procedures for a health problem leads to more generous reimbursement? If health care were produced competitively, then additional payments would not be required to encourage hospitals to produce enough intensive treatments that are valued by patients. A fixed payment per admission or even per individual would be adequate; hospitals would simply maximize (costly) intensity subject to the budget constraint implied by the payment level. Cost sharing implies a concern about imperfect competition: for whatever reason, hospitals would tend to underproduce intensity unless incentives were softened. Myriad reasons for market failures have been proposed in health care, and it is beyond the scope of this paper to review all of them here. I concentrate on one important tension that has been noted by a few other researchers: on one hand, the tension between physician incentives, financial and otherwise, to maximize patient welfare and, on the other, hospital incentives to maximize financial and other rents.

Theoretical papers on the optimal regulation of hospitals have mostly assumed that the incentives facing hospitals match the incentives of the physicians practicing in them, at least up to the (costly) implementation of appropriate internal incentives.[1] One exception is Harris (1977), who describes hospital production as a noncooperative game between physicians engaged in specific relationships with patients on the one hand and hospital managers making investment decisions involving capacity, equipment, and staff hiring on the other. Physicians generally are not employees of the hospital,[2] but rather are participants in highly incomplete contracts with both the hospital and their patients. When a patient seeks care for a health problem, the patient contracts for "appropriate" medical treatment, whatever the health problem turns out to be. Because these treatment contingencies are too numerous to describe ex ante or to verify ex post, patients and physicians cannot sign a complete contract that specifies how the physician will treat every possible medical contingency that may arise in the course of the treatment episode. Just as contingent production decisions are not specified in the contract between a medical expert and a patient, they are not specified in the physician's contractual arrangement with the hospital. Hospital contracts for admitting privileges specify few explicit restrictions on the use of hospital capital that is within the purview of the physician's specialty.

Harris emphasizes the transactions costs of all these treatment intensity decisions and the importance of physician preferences (as opposed to those of hospital managers) in resource allocation decisions for individual patients. But the problem is even more fundamental than transactions costs. Specific investments are substantial: a physician treating a particular patient develops detailed knowledge of the idiosyncrasies of the patient's case. In theory, physicians or hospital managers could hold up patients for the value of these specific investments at critical junctures in the course of treatment, when the patient's life may literally be at stake. These characteristics of hospital production suggest that markets or explicit contracts would represent costly approaches to fostering efficient production.

In this view, the traditional organizational structure of the hospital may rep-

1. An exception is Ellis and McGuire (1986), who model hospital production as a reflection of treatment decisions by physicians. However, the marginal rate of substitution between patient and hospital objectives (that is, the relative weight placed on maximizing hospital profits versus maximizing patient welfare) and hence the extent of the agency problem is treated as exogenous in the model, there is no role for hospital investment decisions in influencing physician decisions, and hospital payment is modeled as fixed regardless of treatment choice. These features may have important implications for hospital production decisions.

2. The physicians that tend to be contractual employees—radiologists, anesthesiologists, and pathologists—prove the rule, since these physicians specialize in technical aspects of hospital production rather than making decisions as agents for particular patients.

resent not so much a reflection of physicians' desire for autonomy in treating patients (Starr 1984) as an institutional mechanism to limit the power of incentives in the physician's other role, as an agent for the hospital. Physicians do not control investment decisions for the hospital or own hospital capital, but they largely retain the rights of control for utilization of hospital capital equipment and other resources.[3] Altering these residual rights of control over treatment decisions, for example through direct employment relationships between physicians and hospitals, would probably alter these residual rights of control and hence treatment decisions for hospitalized patients.[4]

Thus, if a hospital finds that treating patients in the DRG for congestive heart failure (CHF) is unprofitable, it cannot order its admitting physicians not to treat patients with CHF. Physicians control specific admission decisions. Even if such a prohibition were technically possible, it might impose considerable externalities on hospital production. Many patients with CHF initially arrive at the emergency department with shortness of breath, water weight gain, feelings of weakness and lightheadedness, or even vaguer symptoms. Only after some evaluation of the specific case can the diagnosis of CHF be made and treatment initiated. By this time, physicians practicing at the hospital have invested in a relationship with the patient that involves some specific knowledge of the idiosyncrasies of the case. Similarly, CHF is a problem that frequently accompanies other common diseases—such as a recent heart attack, diabetes, or hypertension—which also may require hospitalization (possibly at the same time) and which may represent relatively profitable DRGs for the hospital. Prohibiting or restricting admissions for the diagnosis would limit gains from relationship-specific investments between the patient and physicians that permit the delivery of more effective care for many other diagnoses.

Hospital managers can, however, affect the environment in which physicians make these specific treatment decisions. For diseases such as CHF that involve "nonspecific" technologies, hospital choices about capacity investments in hospital beds, laboratory equipment, and support staffing levels may indirectly affect physician decisions on treatment intensity, but only at the cost of affecting treatment decisions for many health problems. In contrast, for diagnoses and procedures requiring investments in specific capacity, the choices of hospital managers can significantly affect use of the treatments and admissions at the DRG level. For example, if a hospital finds cardiac catheterization unprofitable, it need not prohibit its physicians from catheterizing specific patients. It need only close its catheterization laboratory. By deciding its invest-

3. Some other capital-intensive professions in which the appropriateness of specific production decisions is difficult for managers to observe have similar arrangements. In many research laboratories, for example, the capital equipment may belong to the academic institution, but the researchers who use it are paid largely through grants obtained independently rather than through employee relationships.

4. The consequences of the nature of the employment relationship for production choices and efficiency are examined in detail in Grossman and Hart (1986), which provides some of the foundation for this discussion.

ment level in the technologies needed to produce catheterization, the hospital can effectively specialize or not in catheterization DRGs. Similarly, by determining whether to provide the specialized equipment and personnel required for cancer treatments or particular kinds of surgical procedures, hospital managers can significantly influence admissions in these DRGs.

Regardless of the nature of medical technologies, the physician faces inherent agency conflicts when patient and hospital interests do not coincide. They will not coincide if intensive treatments are costly and if hospitals are not fully insured against costs of care. Vesting the residual rights of production decisions in the physician, and separating physician reimbursement incentives from hospital reimbursement incentives, clearly reduces the strength of the physician-hospital agency relationship. To the extent that physicians are reimbursed at cost for the services they provide, so that the trade-offs they face between patient welfare and their own net profits (including effort costs) are minimized, then physician agency for patients may be much stronger than physician agency for the hospital at the margin.

Whether this allocation of noncontractible rights of control is optimal depends on hospital and patient incentives and on the nature of the alternative technologies that may be involved in patient care. Today, it is easy to consider other allocations of residual control. However, in both traditional insurance arrangements and most managed-care plans, patients typically face limited out-of-pocket prices for medical services; Medicare inpatient services are approximately fully insured.[5] If patients were fully informed and controlled marginal treatment decisions, they would consequently demand all services with positive net benefits. The implication is that residual control by physicians—supported by loose contractual relationships with hospitals and physician payment systems in which the physician tends not to be the residual claimant for substantial costs—strengthens patient agency, leading physicians to prefer more intensive treatment patterns than hospital managers prefer under fixed-price reimbursement.[6]

Together, these institutional characteristics imply that a model of the internal organization of hospital production decisions must incorporate the following features (McClellan 1993):

1. Physicians—who have the most knowledge about the benefits of alterna-

5. The Part A deductible is limited to the cost of one hospital day. While Part B insurance covers only 80% of physician allowed charges, over 80% of beneficiaries have supplemental insurance that eliminates this residual cost sharing.

6. Another advantage of explicitly modeling the internal organization of hospital production is that the consequences of other employment arrangements can be considered explicitly. Strong residual control by physicians has other costs besides limiting the power of hospital managers to influence use of some hospital technologies. For example, the costs of coordinating treatment for an illness beyond the level of an admission may be higher than it would be if the hospital (or HMO) employed all physicians involved in treating the illness. The costs and benefits of the various possibilities for vertical and horizontal integration in the production of treatment for an illness have not been explored much by economists. In this paper, I consider only "traditional" arrangements, which currently describe the production of care for most Medicare beneficiaries.

tive medical technologies for specific patients—are dual agents, for patients in the production of medical treatments and for hospitals in the use of hospital capital. The incomplete contractual arrangements that form the basis for specific production relationships between physicians and patients, and between physicians and hospitals, leave the residual rights of control for decisions to fulfill these production relationships primarily to the physician.

2. These open-ended contractual arrangements, coupled with generous demand-side insurance and a physician payment system that is largely independent of hospital profits, suggest that physicians involved in specific agency relationships with patients weigh patient welfare more than hospital profits, and in any case more than hospital managers weigh patient versus hospital objectives.

3. The nature of the medical technologies used to treat particular illnesses influences how hospital payment incentives influence equilibrium production decisions. For treatments dependent on specific hospital investments, hospital managers have more control over production levels; for treatments involving nonspecific technologies that are difficult to ration, physicians have more control.

5.2 A Model of the Production of Medical Treatment Intensity

McClellan (1993) formalizes the noncooperative game of hospital production described in the previous section. Here I sketch the main results for investment in a "simple" intensive technology, one that may be used in the treatment of a single health problem. There are four parties: fully insured *patients* with a health problem who decide which hospital to visit for treatment; *physicians* who learn the patient's expected benefits from alternative treatments and make treatment intensity decisions; *hospitals* that provide medical technologies for treatment; and a *regulator* who sets the price schedule for hospitals. The structure of the game is as follows:

1. The regulator sets prices for less intensive treatment (without the intensive technology) and more intensive treatment for the health problem.
2. The hospital invests in medical technologies that constitute the environment for medical treatment.
3. Physicians observe hospital investment choices, contract with hospitals to provide medical services, and define their equilibrium treatment intensity choices.
4. Patients observe some measure of hospital quality related to its investment decisions and, possibly with guidance from physicians, choose a hospital for treatment.
5. The physician observes the patient's expected benefit from alternative treatments and chooses treatment intensity. There are two treatment intensity choices, $r = 0$ (less intensive) and $r = 1$ (more intensive). The performance of

$r = 0$ or $r = 1$ is observable and verifiable, and hence can be used as the basis for a reimbursement contract.

6. Patients receive treatment and hospitals are reimbursed for services.

Suppose that the nature of the technology is such that the hospital can choose a capacity level; for example, staffing and equipment choices in a cardiac catheterization laboratory can effectively determine the number of patients with heart disease who receive catheterization. Because investment in the capacity to produce the intensive treatment is costly, hospital managers will invest in intensive capacity until the marginal benefits of investment (in terms of additional demand attracted and patients treated profitably) equal the marginal costs (in terms of the costs of additional capacity and the potential losses on the fraction of patients treated intensively). In turn, as more capacity becomes available, physicians are able to treat more patients intensively (the "marginal" patient treated intensively has a lower perceived benefit), and more intensive treatment patterns attract more patients. Denote the regulated payment level as $\bar{p}^0$ and $\bar{p}^1$ as the regulated prices for less intensive and more intensive treatment, c^0 and c^1 as the corresponding marginal costs, I^1 as the hospital investment in intensive capacity, $\hat{\theta}(I^1)$ as the expected net benefit (perceived by the physician) of more intensive treatment for the marginal patient receiving it, $F(\hat{\theta})$ denotes the share of patients treated with $r = 0$, and $q(\hat{\theta})$ denotes aggregate demand (decreasing in $\hat{\theta}$, i.e., increasing in intensity). The hospital will choose capacity I^1 such that

$$\frac{1}{\eta_{0\theta}}\cdot[\bar{p}^0 - c^0] + \frac{1}{\eta_{1\theta}}\cdot[\bar{p}^1 - c^1] = \frac{1}{\eta_{q\theta}}\cdot[(\bar{p}^0 - c^0) - (\bar{p}^1 - c^1)],$$

where $\eta_{0\theta} = (\partial F/\partial\theta)\cdot(\theta/F)$ is the elasticity of treatment with $r = 0$ with respect to $\hat{\theta}$; $\eta_{1\theta} = -(\partial(1 - F)/\partial\theta)\cdot(\theta/(1 - F))$ is the elasticity of treatment with $r = 1$ and with respect to $\hat{\theta}$; and $\eta_{q\theta} = -q'\cdot\theta/q$ is the elasticity of demand with respect to intensity $q(I)$. All of these "quality elasticities" are defined to be positive.[7]

The equation shows that equilibrium intensity increases and equilibrium profits[8] decrease as the elasticity of demand with respect to intensity increases. In the extreme case of perfectly elastic demand, the zero-profit constraint holds and intensity is maximized subject to the price constraints. For a given $\bar{p}^0$, equilibrium intensity also increases as $\bar{p}^1$ increases, and the difference in margins $[\bar{p}^0 - c^0] - [\bar{p}^1 - c^1]$ decreases (if $\bar{p}^0$ is set high enough for production to occur). Equilibrium intensity will only be first-best for a particular combination of prices, and actual intensity would vary across markets as a function of demand elasticity.

This model of equilibrium intensity choices for hospital investment in a specific technological capacity demonstrates that the quality and cost of hospital

7. McClellan (1993) describes the technical properties of this solution in more detail.

8. For nonprofit hospitals, economic rents are possible even if profits are not.

production need not depend on *physician* preferences in any direct way. However, it is easy to imagine other kinds of medical technologies where physician preferences would be much more important in determining equilibrium choice. Specific investments in capacity are an extreme case; for other technologies, physician objectives are more relevant to equilibrium intensity choices.

Many technologies involve a more or less binding capacity choice but are not specific to the treatment of a particular health problem. For example, hospital managers typically make decisions about opening a hospital bed, purchasing a computerized tomography (CT) scanner, or providing a chemical analyzer for laboratory tests, but they have much less control over how these technologies are used in practice. Because the hospital cannot direct their use to specific kinds of patients, it can only choose an optimal aggregate intensity level for all the health problems. It does so considering how physicians will allocate marginal units of capacity among patients with the various health problems, and the implications of these decisions for profits. In addition, some specific and nonspecific investments by hospitals may represent "line of business" decisions rather than capacity choices. For example, a hospital can decide whether to stock specific drugs such as thrombolytic agents (for heart attacks) or less specific drugs such as antihypertensive and pain-reducing drugs (for many types of patients), but it cannot easily choose a capacity for these treatments. Hospitals can make binary choices about whether to adopt these technologies, knowing that physicians will choose the treatment pattern that they prefer.

McClellan (1993) formalizes a model of these hospital investment decisions as well. For example, in deciding whether to adopt a "line of business" technology, the hospital's adoption decision is based on profits if it does or does not adopt; if the hospital adopts, then physician preferences and not hospital preferences determine equilibrium intensity levels for use of the technology. Even if hospital production is not very competitive, physician decisions can thus lead to more intensive treatment (i.e., lower rents and more spending on desired technology) than hospital managers would prefer.

Most medical technologies involve physician effort costs between the extremes of investments in fixed capacity and in a line of business. For example, congestion may make physician effort costs an increasing function of their intensity choice, as scanners or operating rooms become more crowded. Hospitals must trade off higher unit production costs associated with congestion against the effects of congestion on limiting resource use; the model is qualitatively similar to the capacity-constraint case. Hospitals may also make costly investments in technologies that influence physician use of other technologies. For example, hospitals may invest in monitoring systems that require physicians to justify particular admissions or long stays, in computer technologies that make it easier to order certain tests or drugs but harder to order others, and other "utilization management" investments that are not only costly to the hospital but also to the physician. Such investments are not first-best, but may be the best available option to hospital managers when their preferences conflict with those of the physicians.

This discussion suggests that optimal regulation of medical prices should consider not only price levels but also the extent of price aggregation. *Any* administered price system will lead to efficient production of the particular bundles of services for which prices are set exogenously: the fixed price makes the producer the residual claimant for all savings resulting from cost minimization in production of the services. However, disaggregating payments for treatment of a particular health problem so that the more intensive treatment receives a higher payment encourages more use of the intensive treatment than would be supported by a single, aggregated price. Administered price systems for medical care, as for other industries, are not new. The so-called cost-based hospital reimbursement system that Medicare PPS replaced was such a disaggregated administered price system. Hospitals could not charge any desired price for services billed to Medicare; instead, hospital regulators reviewed reported hospital costs for their "reasonableness" in relation to the services provided. Many large private insurers continue to rely on such disaggregated administered price systems today for hospital services; they are also widely used for reimbursement of physician services. At the opposite extreme, some health maintenance organizations (HMOs) use fixed-price per capita contracts with hospitals or physicians for providing all medical services for a given population.

Medicare PPS is between these extremes. The system provides a single payment per hospitalization, which is a more aggregated level of administered prices than the cost-based reimbursement system that it replaced. However, some particular treatments or types of treatments—primarily intensive surgical procedures—still define administered price groups. In contrast to the current DRG system, one can imagine a more aggregated DRG system at the level of diagnoses only.[9]

If hospitals have market power and can effectively control treatment intensity, increasing the aggregation of the pricing system reduces the costs of producing care for a health problem through intensity reductions. Such incentives may be too high-powered. On the other hand, setting price equal to cost for each alternative treatment provides no financial incentive for the hospital to limit treatment intensity; this is the same kind of problem that "prospective" payment was intended to address. There is no conflict between physician agency for patients and for hospitals with respect to the intensity choices that are each reimbursed at cost. Thus, pricing admissions for an intensive treat-

9. Diagnoses themselves may also be more or less aggregated. DRGs for most major diagnostic categories include a number of specific diagnoses (e.g., AMI, deep venous thrombophlebitis), some less specific diagnoses (e.g., chest pain), and catchall residual categories (other circulatory system diagnoses). Here I focus on the effects of allowing DRG classification to depend on treatment intensity decisions.

ment at its cost will tend to encourage its use to the point that marginal benefits are close to zero, though the treatment itself will be produced efficiently.

An intuitive solution to this dilemma is for the regulator to "split the difference" in setting prices for the alternative treatments. If regulators have some knowledge of the demand function and the distribution of net benefit levels from more intensive treatment, they can use the equation derived in McClellan (1993) to solve for the set of prices that implements the desired intensity level. Returning to the notation of the model, since treatment intensity is increasing in $(\bar{p}^1 - \bar{p}^0)$, this procedure involves setting a price for $r = 0$ (the less intensive treatment) greater than c^0 and a price increment for $r = 1$ less than the cost difference $(c^1 - c^0)$. While such a "first-best" rule seems to have strong informational requirements, changes in the Medicare price structure that are clearly welfare-improving can presumably be accomplished with much more limited information. In particular, moving away from paying the full cost differential (on average) for a more intensive treatment should reduce incentives to provide the more intensive treatment in cases of minimal expected benefit.

5.3 Medicare Price Regulation and the Dynamics of Medical Treatment Intensity

In this section, I illustrate the dynamic implications of the Medicare price rule based on the model reviewed in section 5.2. I assume no exogenous changes in technology over time to focus on endogenous technological change resulting from the PPS price rules. These DRG pricing rules include unbundled DRGs for some intensive treatments, and price updates that essentially make the current DRG price a function of average charges for patients in the DRG in a previous year (lagged "yardstick" competition).

The dynamic implications can be illustrated qualitatively with phase diagrams. Consider a true DRG; that is, reimbursement is a fixed price for the diagnosis regardless of the treatment intensity level chosen. Figure 5.1 illustrates the current Medicare pricing rule for a DRG of setting $\bar{p}_{t+1} = c_t$, where c_t is the average cost in the DRG in period t.[10] If there are a large number of identical hospitals, the impact of intensity choices by a particular hospital will have no effect on price in subsequent periods, so hospitals will not behave strategically, and all will choose the same equilibrium intensity and cost for a given price. At X, $p_t(X)$ is lower than $c_t(X)$, so $p_{t+1}(X)$ increases to offset the difference. At Y, $p_t(Y)$ is higher than $c_t(Y)$, so $p_{t+1}(Y)$ falls to offset the difference. At Z, $p_t(Z)$ equals $c_t(Z)$, so $p_{t+1}(Z)$ is unchanged. Assuming for now that regulators can measure hospital costs accurately, it is evident that a line with

10. The actual Medicare pricing rule is to set price equal to a conversion factor times the relative weight of a particular DRG, where the weight is determined by the ratio of charges in the DRG to average charges for all Medicare patients in year $t - 2$. The conversion factor is set to achieve a particular total expenditure target. Thus the price update rule is exact only if the conversion factor updates hospital output prices by the rate of increase in hospital input prices.

unitary slope through the origin traces out the set of (p_t, c_t) combinations for which p_{t+1} will be unchanged; this is the $dp/dt = \dot{p} = 0$ equilibrium line. To the northwest, $\dot{p} > 0$; to the southeast, $\dot{p} < 0$.

Figure 5.2 illustrates a complementary relationship for the dynamics of equilibrium cost. If cost happened to be very low in a period when a very high price was offered, cost will increase (the higher price increases the returns to attracting more demand). Thus, in the southeast corner of figure 5.2, for example at point Y, $c_{t+1}(Y) > c_t(Y)$. In contrast, if equilibrium cost happened to be very high and a low price were introduced, cost would fall through intensity reductions. Thus, in the northeast corner, for example at point X, $c_{t+1}(X) < c_t(X)$. At point Z, cost and price are such that no equilibrium adjustment is necessary, so $\dot{c} = 0$. A sample $\dot{c} = 0$ line is traced out in the figure; the line reflects an equilibrium cost function that becomes less sensitive to price at higher price levels. Different demand functions, intensity costs, distributions of expected net benefits to patients, and degrees of physician control over intensity choices will generate different forms for $\dot{c} = 0$. Intensity is increasing in costs, so the $\dot{c} = 0$ mapping also implies an equilibrium intensity level $\hat{\theta}$ as a function of price.

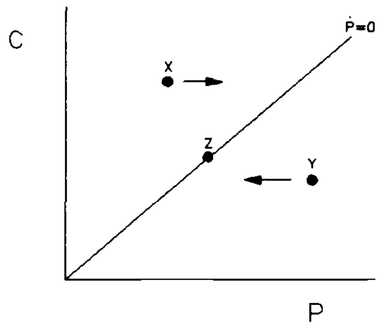

Fig. 5.1 Dynamics of Medicare DRG prices

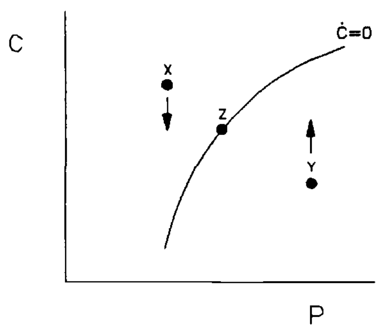

Fig. 5.2 Dynamics of hospital costs: an example

Combining the $\dot{c} = 0$ and $\dot{p} = 0$ relationships, as in figure 5.3, illustrates how the PPS might influence the dynamics of equilibrium cost and intensity levels in the DRG. In this example, two points comprise equilibrium price-cost relationships, *A* and *B*. *A* is an unstable equilibrium, since any perturbation away from *A* will tend to move farther from *A*. *B* is a stable equilibrium, since local perturbations tend to return to *B*. A key feature is that at *B* the slope of the $\dot{p} = 0$ function is greater than that of the $\dot{c} = 0$ function. The phase diagram does not depict time to convergence, but the time course of a sample intensity path from point *W* to point *B* is illustrated in figure 5.4. Cost and intensity first fall then rise to converge on *B* (that is, intensity first rises, then falls).

Note that, if costs are measured accurately, the $\dot{p} = 0$ line also comprises the zero-profit line for hospital production. One feature of figure 5.5, that $\dot{c} = 0$ equilibria could occur in a range where $p > c$, may seem implausible. However, it is possible if cross-subsidies exist (e.g., hospital investments are

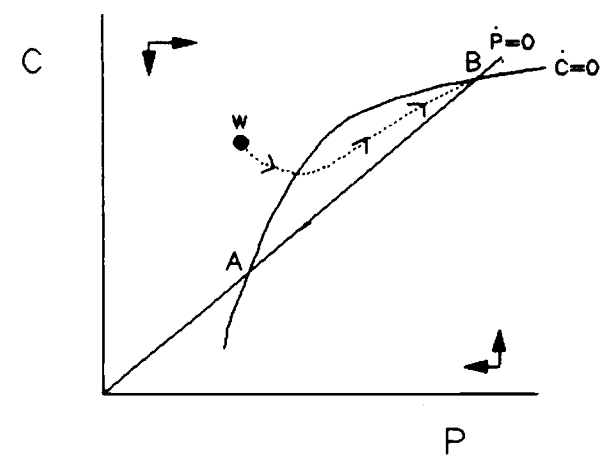

Fig. 5.3 Stable dynamic equilibrium

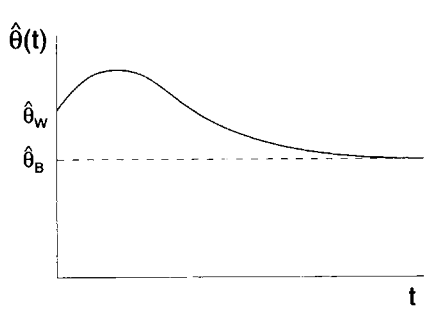

Fig. 5.4 Intensity choice over time

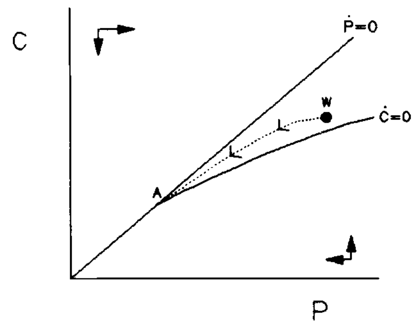

Fig. 5.5 Low-intensity dynamic equilibrium

nonspecific), if hospital costs are subsidized from other sources (e.g., separate reimbursement of capital costs, as described below, or subsidies from non-Medicare patients), if hospital-reported cost estimates are biased upward, or if other arguments in the hospital objective function besides patient welfare can be disguised as necessary costs of care. I return to some of these possibilities momentarily.

If none of these possibilities exist, figure 5.5 represents a more realistic path in the case where hospitals have market power. Point *A* represents a stable equilibrium, occurring at the *minimal* intensity level. At any given price, a hospital with market power can earn positive price-cost margins on average by decreasing quality slightly below the level that maximizes consumer welfare (and earns zero profits). Consequently, average cost is less than price. Thus the $\dot{c} = 0$ curve will generally have a flatter slope and be located to the right of the $\dot{p} = 0$ curve over the range of prices in which hospitals can choose intensity levels above the minimum. The updated price (equal to observed cost in the previous period) will be lower, leading to a lower optimal intensity level, and this downward spiral continues until the minimum intensity level is reached. In practice, of course, there is no distinct minimal intensity level of medical care; at the extreme, hospital care would simply move to an outpatient setting.

However, higher equilibrium intensity levels (or at least a slower rate of decline) are possible if the technologies used at equilibrium are such that physicians acting as patient agents control intensity or if hospitals face near-perfect competition. Whenever hospital demand is perfectly elastic or hospitals maximize patient welfare for other reasons, *any* point along the $\dot{p} = 0$ line is a stable equilibrium. Intensity is maximized subject to the zero-profit constraint, so price equals cost for *any* price at least as high as the minimum intensity level of hospital production requires. Thus, as shown in figure 5.6, an infinite number of stable equilibria are possible; the particular equilibrium that exists depends only on the initial price level.

Perfect competition or physician choices leading to intensity maximization for a given price result in stable equilibria. In contrast, figure 5.7 presents a

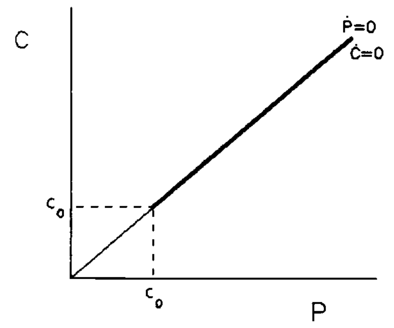

Fig. 5.6 Multiple dynamic equilibria with perfect competition

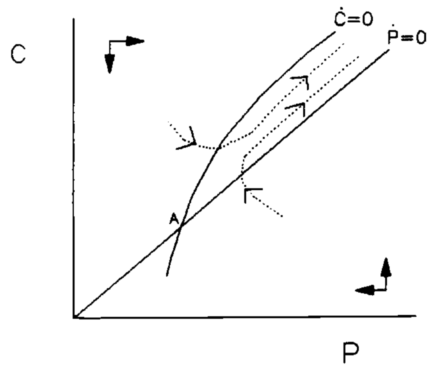

Fig. 5.7 Unstable dynamic intensity growth

more worrisome situation, in which the price rule encourages dynamic increases in intensity. The equilibrium $\dot{c} = 0$ curve is above the equilibrium $\dot{p} = 0$ curve in the relevant range of prices and costs. Equilibrium A is unstable: from any initial condition with cost or price beyond A, cost and price will continue to increase over time as long as the $\dot{c} = 0$ line is above the $\dot{p} = 0$ line. This instability is a general characteristic of pricing rules that increase the steepness of the $\dot{c} = 0$ line. Three scenarios that tend to lead to such relationships are considered below: separate cost-based reimbursement of specific dimensions of intensity, nonprofit hospital objectives, and overestimation of hospital marginal production costs in the reimbursement formulas.

In some respects the Medicare pricing rule is less high-powered than suggested by the fixed-price rules described in the preceding figures. Consider the extreme case of disaggregated pricing at average cost for each treatment, preserving the incentive to produce the treatment efficiently but making hospital profits zero regardless of intensity choice. Then every point on the $\dot{c} = 0$ line can be a dynamic equilibrium. (In fact, any intensity level in which price equals cost is sustainable.) Medicare actually reimburses only certain dimensions of treatment intensity separately, but this can still encourage broader intensity and cost increases in the treatment of a disease. In general, as the previous discussion indicated, increasing the price differential between alternative intensity levels appears to increase the steepness of the market equilibrium ($\dot{c} = 0$) curve, increasing the probability of unstable cost growth.[11]

Another possible cause of unstable intensity growth, even if measured costs are correct, is the existence of nonprofit objectives in hospital objective functions. Only a small fraction of hospitals are for-profit, so that most hospitals should not be expected to show much total accounting profit. To the extent that such nonprofit objectives translate into higher observed costs of treating Medicare patients, they will tend to push hospital $\dot{c} = 0$ curves toward—if not above—the $\dot{p} = 0$ curve, increasing the probability of dynamic cost and intensity increases in the absence of technological change.

An additional potential source of instability is mismeasurement of marginal production costs. A large literature documents the difficulty of correctly estimating "true" production costs in multiproduct firms in other industries. Fundamental cost estimation problems such as allocating costs in joint production or calculating true economic depreciation rates for assets that may be expensed or depreciated according to standard accounting formulas are no different for hospitals. Such detailed cost estimation is practically infeasible anyway. Medicare's cost estimates for setting DRG prices are obtained from "ghost" bills for hospital charges that would have been generated if the beneficiary were a private patient.[12] Further, because DRG prices are updated based on average charges in the DRG relative to national average charges,[13] the reimbursement system tends to favor treatments for which true cost-to-charge ratios are low.

In particular, the reimbursement system favors DRGs that rely more heavily on capital-intensive treatments such as operative procedures or laboratory tests. These DRGs have the lowest measured cost-to-charge ratios (Rogowski and Byrne 1991). Other features of PPS also favor more capital-intensive treatments. Teaching hospitals receive an indirect medical education adjustment to their DRG payments proportional to their ratio of interns and residents to hospital beds. Since the adjustment is proportional, it provides larger increments to reimbursement for admissions in more intensive DRGs. Finally, Medicare payments for hospital capital expenditures were cost-based through 1991:

11. Of course, reimbursing some dimensions of treatment intensity at higher levels also affects the form of the $\dot{p} = 0$ curves for treatment of the health problem in ways that cannot easily be captured in a two-dimensional phase diagram. In general, as the static model indicated, smaller intensity price differentials lead to higher equilibrium intensity levels; if the resulting cost increases lead to further average price increases, the cycle may continue over time.

12. The relationship of hospital charge data to "true" costs may be declining. Most hospitals provide special discounts to many large payers (e.g., Blue Cross, HMOs, preferred provider organizations [PPOs]). Charges to residual third-party payers may be higher as a result, and in any event may bear less of a direct relationship to actual hospital costs of care.

13. Charge-based weights were first used in the fiscal year 1988 PPS DRG updates. Prior to that, cost-based weights were used, where costs were calculated from reported hospital charges using hospital-specific cost-to-charge ratios.

Medicare reimbursed hospitals for 85% of its "share" of capital purchases (i.e., 0.85 times the proportion of Medicare admissions at the hospital times the capital purchase). The pass-through payment system for capital costs is being gradually phased out over a ten-year period ("prospective" capital payments will be included in DRG payments), but pass-throughs will still be larger for the more intensive, treatment-based DRGs. Thus, particularly for intensive DRGs requiring substantial investments in equipment, reported costs (and prices) may be greater than true marginal costs to the hospital.

The dynamic instability of treatment intensity is further complicated by the mechanism for updating Medicare prices. The "conversion factor" between DRG relative weights and dollar payment levels is set so that if the current period's DRG admission patterns are identical to the prior period's DRG admission patterns, aggregate hospital payments will increase at a target growth rate. This updating mechanism has consequences for the dynamic reallocation of resources to health problems between single-DRG health problems and health problems for which DRGs are available for multiple intensity levels. Suppose that all treatments for health problem A are aggregated into a single DRG at price $\bar{p}_A(t)$, and that treatments for health problem B are divided into a less intensive DRG at price $\bar{p}_B^0(t)$ and a more intensive DRG at price $\bar{p}_B^1(t) > \bar{p}_B^0(t)$. Assume that the number of patients with the health problem in each period n_A and n_B is fixed and that production of care for A is competitive, so that $\hat{c}_A(t) = \bar{p}_A(t)$ (the dynamic problem is worse if it is not). Assume that the current (price, cost) state is such that the use of the more intensive treatment for problem B is increasing (that is, $\hat{\theta}_B$ is decreasing). The current price for a DRG is a function of its relative weight $\omega(t-1)$ in the previous period (average cost per admission in the DRG divided by the average cost of all admissions) times the conversion factor $\phi(t)$, which solves

$$\phi(t)\cdot\{n_A \cdot \omega_A(t-1) + n_B \cdot[F(\hat{\theta}_B(t-1))\cdot\omega_B^0(t-1)] \\ + [1 - F(\hat{\theta}_B(t-1))\omega_B^1(t-1)]\} = \hat{c}(t-1)\cdot g(t),$$

where $\hat{c}(t-1)$ is total cost in the previous period, and $g(t)$ is the exogenous target growth rate for the current period. If the rate of change of $F(\hat{\theta}_B)\cdot(c_B^1 - c_B^0)$ exceeds g, then $\phi(t)$ will decrease over time, leading to a decline in price (and cost) for treatment of health problem A and a reallocation of resources to health problem B. Even if the production of A is competitive, the intensity of treatment for problem A will decline over time due to an externality effect of increasing intensity for problem B induced by the update rule.[14]

This theoretical evaluation of intensity dynamics indicates that PPS rules are unlikely to lead to dynamic cost-intensity equilibria with any particularly

14. In recent years, PPS updates have been based on target growth rates lower than measured hospital cost inflation. To the extent that this reflects true price ratcheting, it can be captured in the phase diagrams above by a counterclockwise rotation of the $\dot{p} = 0$ line. Thus it tends to lead to treatment equilibria at lower intensity levels and is less likely to result in unstable growth; the qualitative conclusions are unchanged.

desirable properties, if they lead to cost equilibria at all. The next section presents some preliminary empirical evidence consistent with the model of reimbursement dynamics suggested here.

5.4 Preliminary Empirical Evidence

The model developed in sections 5.2 and 5.3 has a number of empirical predictions. In this section, I describe preliminary evidence on two of them. First, facing some competitive pressures, hospitals have minimal incentives to limit dimensions of treatment intensity that involve technologies reimbursed through treatment-based DRGs; if the treatments themselves are produced efficiently, these costs will be shared almost completely. Second, if demand for hospital services is not perfectly elastic in intensity, dynamic changes in treatment intensity are likely to occur *independently* of any exogenous technological change. In particular, dynamic intensity growth is likely for health problems where reimbursement differences reflect cost differences, leading to steeper equilibrium price-cost relationships, or where measured marginal production costs overstate true costs. Conversely, for health problems grouped into true diagnosis-based DRGs, declines in intensity over time are likely because the PPS price update mechanism tends to increase the share of reimbursement devoted to health problems in which cost increases can occur through shifts to more intensive DRGs.

To illustrate the practical consequences of the structure of the DRG reimbursement system, I first present evidence on one particular health problem, AMI. This very common health problem in the elderly accounts, directly or indirectly, for much of the mortality and hospital use associated with coronary heart disease. The data are derived from all hospital discharge abstracts filed over a two-year period by all elderly Americans hospitalized with a new AMI in 1987. The sample thus includes data from 522,506 hospitalizations for 205,021 elderly AMI patients (see McClellan and Newhouse 1993 for details of the data-set creation process). Table 5.1 summarizes demographic characteristics and some important treatment intensity decisions for these patients. In 1987, approximately 23% of elderly AMI patients underwent cardiac catheterization, an invasive procedure, and approximately 13% of these patients underwent a further intensive revascularization procedure, PTCA, or CABG surgery. Approximately 87% of patients spent at least one day in a specialized coronary- or intensive-care-unit bed (CCU/ICU), including 84% of patients who did not undergo any invasive procedure (second column).

These intensive treatments have different implications for hospital reimbursement. Admissions for AMI patients who do not undergo intensive procedures are categorized into DRG 121 (AMI with complicating conditions), 122 (AMI without complicating conditions), or 123 (AMI, expired). The three DRGs differ somewhat in reimbursement levels, reflecting patient disease severity (ideally independent of hospital treatment choices) that is correlated

Table 5.1 **Characteristics and Technology Use for Elderly Patients with Acute Myocardial Infarction (%)**

		Use of Invasive Procedures			
	All Patients (N = 205,021)	No Procedures (N = 155,880)	Catheterization Only (N = 22,902)	Catheterization and PTCA[a] (N = 10,837)	Catheterization and CABG[b] (N = 15,402)
Patient characteristics					
Female	50.36	53.75	41.07	41.02	36.37
Black	5.65	6.08	5.69	3.19	2.96
Age in years	76.11	77.49	71.71	71.51	71.47
(standard deviation)	(7.26)	(7.28)	(5.11)	(5.15)	(4.79)
Rural	29.45	30.56	26.40	24.45	26.32
Technology use					
Catheterization within 90 days	22.81	0	100.0	100.0	100.0
Percutaneous transluminal coronary angioplasty within 90 days	5.30	0	0	100.0	0
Coronary artery bypass graft within 90 days	7.74	0	0	0	100.0
Acute treatment in coronary care unit	86.56	83.99	92.07	96.70	97.21

[a]Percutaneous transluminal coronary angioplasty.

[b]Coronary artery bypass graft.

with cost of care. Both revascularization procedures, PTCA (DRG 112) and CABG (DRGs 106 and 107), involve distinct DRGs with much higher reimbursement levels that reflect the incremental costs of the intensive treatments. In contrast, although one might imagine a DRG for "AMI with CCU admission," there is no incremental reimbursement for treatment involving specialized CCU beds. The implications for DRG classification of the initial hospitalizations of AMI patients are presented in table 5.2, which shows that only 92% of all AMI patients and only 88% of male patients aged 65–74 years were initially hospitalized in an AMI-related DRG. The other AMI patients were mostly categorized in more intensive DRGs based on the use of invasive or surgical treatments. Table 5.2 also shows that the share of treatment-based DRGs is even larger when all hospitalizations within thirty days of the AMI are considered. AMI patients are specifically excluded from the catheterization DRGs (124 and 125), but over 1% of patients are hospitalized in these DRGs within a month of their AMI.

Table 5.3 reports means and standard deviations for hospital utilization, reported hospital costs, and reimbursement for AMI patients grouped by which treatments they received. These statistics are reported for all hospital admissions during three time intervals after AMI: within thirty days, within ninety

Table 5.2 Distribution of Admissions for Acute Myocardial Infarction among Diagnosis-Related Groups

	All Patients (N = 205,021)		Male Patients Aged 65–74 (N = 55,570)	
Diagnosis-Related Groups	Initial AMI Admission (N = 205,021)	All Admissions within 30 Days (N = 251,575)	Initial AMI Admission (N = 55,570)	All Admissions within 30 Days (N = 72,006)
Acute myocardial infarction (AMI) (121, 122, 123)	92.0	81.1	88.1	75.4
Coronary artery bypass graft (106, 107)	2.2	4.9	4.1	8.5
Percutaneous transluminal coronary angioplasty (112)	2.8	3.9	4.7	6.1
Catheterization without acute myocardial infarction (124, 125)	0	1.2	0	1.9
Other cardiovascular procedures (mainly 109, 115)	2.5	2.8	2.7	3.0
Other	0.5	6.1	0.4	5.2

Table 5.3 **Hospital Utilization, Costs, and Reimbursement for Elderly Patients with Acute Myocardial Infarction**

Time Interval after AMI	All Patients (N = 205,021)	Use of Invasive Procedures: No Procedures (N = 155,880)	Catheterization Only (N = 22,902)	Catheterization and PTCA[a] (N = 10,837)	Catheterization and CABG[b] (N = 15,402)
Hospital admissions[c]					
Within 30 days	1.22	1.12	1.46	1.53	1.71
	(0.48)	(0.35)	(0.62)	(0.67)	(0.69)
Within 90 days	1.46	1.29	1.83	1.90	2.23
	(0.77)	(0.63)	(0.90)	(0.94)	(0.97)
Within 1 year	1.95	1.78	2.38	2.42	2.67
	(1.39)	(1.31)	(1.53)	(1.48)	(1.46)
Hospital days[c]					
Within 30 days	12.06	10.85	13.46	13.43	21.31
	(10.59)	(9.92)	(8.99)	(9.24)	(14.50)
Within 90 days	14.13	12.45	15.91	15.91	27.22
	(13.70)	(12.76)	(12.27)	(11.93)	(17.90)
Within 1 year	18.38	16.76	20.37	19.64	30.93
	(19.69)	(19.19)	(18.62)	(16.93)	(22.87)
ICU/CCU days[c]					
Within 30 days	5.28	4.46	6.48	6.98	10.59
	(6.09)	(5.11)	(6.46)	(6.03)	(10.10)
Within 90 days	5.83	4.80	7.24	7.88	12.62
	(6.92)	(5.66)	(7.57)	(7.34)	(11.26)
Within 1 year	6.77	5.73	8.47	8.92	13.35
	(8.27)	(7.15)	(9.24)	(8.94)	(12.15)

Total hospital costs ($)[d]					
Within 30 days	7,702	6,073	8,992	11,413	19,663
	(7,529)	(5,505)	(6,880)	(7,136)	(12,634)
Within 90 days	9,096	6,953	10,697	13,614	25,221
	(9,188)	(6,694)	(8,699)	(8,985)	(13,285)
Within 1 year	11,364	9,198	13,416	16,024	26,964
	(7,529)	(9,878)	(12,187)	(11,958)	(15,456)
Total hospital reimbursement ($)[e]					
Within 30 days	6,810	5,418	7,141	9,880	18,235
	(5,704)	(3,501)	(4,985)	(5,471)	(9,690)
Within 90 days	8,098	6,177	8,749	12,146	23,708
	(7,066)	(4,443)	(6,384)	(6,888)	(8,794)
Within 1 year	10,078	8,118	11,155	14,479	25,208
	(9,196)	(7,338)	(9,161)	(9,425)	(10,497)

Note: Standard errors in parentheses.

[a]Percutaneous transluminal coronary angioplasty.

[b]Coronary artery bypass graft.

[c]Number of admissions, number of total hospital days, and number of ICU or CCU days were calculated from Medicare claims.

[d]Reported departmental costs for each hospitalization were calculated by multiplying the reported departmental charges for a hospitalization by the PPS cost-to-charge ratio for that hospital department and summing the resulting cost estimates across all departments. Reported routine costs were calculated by determining the average accounting cost per day for Medicare patients by bed type (standard, ICU, CCU) and multiplying by the number of days spent in each bed type during the admission. Total reported operating costs are the sum of these two components; DRG payments are intended to reimburse hospitals for these operating costs. As noted above, these reported costs are not perfect measures of "true" average incremental costs to the hospital and so should only be interpreted qualitatively.

[e]Reimbursement rates were calculated for each admission by summing DRG-based payments and DRG outlier payments (if any).

days, and within one year. Compared to patients not undergoing invasive procedures, patients undergoing one or more of the procedures used hospitals more intensively (in terms of number of admissions, total days, and intensive-care days) and incurred substantially higher costs. But reimbursement totaled a relatively constant proportion of reported costs, regardless of procedure use. For example, within ninety days of AMI, total DRG reimbursement for patients not undergoing procedures was 89% of reported costs ($6,177 versus $6.953 on average), for patients undergoing catheterization only was 82% of costs ($8,749 versus $10,697), for patients undergoing catheterization and PTCA was 89% of reported costs ($12,146 versus $13,614), and for patients undergoing catheterization and CABG was 94% of reported costs ($23,708 versus $25,221). Thus, for these intensive procedures used in the management of AMI patients, PPS reimbursement tracks costs quite closely on average.[15]

In contrast, table 5.4 reports summary statistics for patients grouped on the basis of whether or not they stayed in a specialized CCU or ICU bed for more than two days during their acute AMI treatment. Patients with acute CCU/ICU stays of two days or less had costs significantly lower than those of patients with stays over two days. Reimbursement differences between the two groups were much more modest, however, so that patients receiving two days or less of CCU/ICU treatment had average reimbursement levels 29% higher than average hospital costs within ninety days of AMI ($5,133 versus $3,971), while average reimbursement was only 77% of costs for patients receiving more than two days' CCU/ICU treatment ($6,802 versus $8,855). Such a pattern is expected under prospective payment—patients requiring more CCU/ICU days are sicker than those who do not—but the pattern differs markedly from that observed for intensive technologies with separate DRGs. Because DRG payments are largely independent of CCU/ICU use, additional CCU days lead to relatively little additional hospital reimbursement.

The association between reimbursement and treatment intensity for AMI illustrates the financial incentives for hospital investments in medical technologies. Implementing the capacity to perform CABG, PTCA, or cardiac catheterization, as well as choosing the quantity of CCU beds to support, are all hospital investment decisions that may have substantial effects on physician decisions for AMI treatment. As noted in section 5.1, analogous intensive hospital treatments and treatment-based DRGs exist for many other health problems, providing similar relatively low-powered incentives for investments in the capacity to perform treatments.

Table 5.5 suggests that these incentives have had a fundamental effect on the

15. The slightly lower proportion of cost sharing for patients undergoing catheterization only is likely to be a reflection of DRG structure as well. For patients admitted with AMI, only if catheterization occurs during a subsequent admission does it provide additional reimbursement for the hospital. Table 5.2 demonstrates that some patients are readmitted soon after their initial AMI admission to undergo catheterization, which then constitutes a non-AMI DRG.

Table 5.4 **Hospital Utilization, Costs, and Reimbursement for Elderly Patients with Acute Myocardial Infarction, 1987**

		Acute Coronary Care/Intensive Care Unit	
Time Interval after AMI	All Patients (N = 205,021)	Two Days or Less (N = 60,707)	More Than Two Days (N = 95,103)
Hospital admissions			
Within 30 days	1.22	1.06	1.16
	(0.48)	(0.26)	(0.40)
Within 90 days	1.46	1.19	1.39
	(0.77)	(0.52)	(0.68)
Within 1 year	1.95	1.58	1.91
	(1.39)	(1.16)	(1.37)
Hospital days			
Within 30 days	12.06	7.08	13.30
	(10.59)	(7.97)	(10.29)
Within 90 days	14.13	8.24	15.13
	(13.70)	(10.96)	(13.10)
Within 1 year	18.38	11.67	20.01
	(19.69)	(17.24)	(19.66)
CCU/ICU days			
Within 30 days	5.28	0.87	6.75
	(6.09)	(0.83)	(5.38)
Within 90 days	5.83	1.03	7.21
	(6.92)	(1.60)	(5.99)
Within 1 year	6.77	1.53	8.40
	(8.27)	(3.22)	(7.66)
Total hospital costs ($)			
Within 30 days	7,702	3,379	7,791
	(7,529)	(2,789)	(6,090)
Within 90 days	9,096	3,971	8,855
	(9,188)	(4,025)	(7,332)
Within 1 year	11,364	5,673	11,446
	(7,529)	(7,452)	(10,555)
Total hospital reimbursement ($)			
Within 30 days	6,810	4,637	5,915
	(5,704)	(2,314)	(4,002)
Within 90 days	8,098	5,133	6,802
	(7,066)	(3,245)	(4,960)
Within 1 year	10,078	6,695	9,023
	(9,196)	(6,120)	(7,885)

Note: Standard errors in parentheses.

nature of hospital expenditure growth since the adoption of PPS. The table summarizes changes in admissions and hospital treatment intensity between 1983 and 1988 for some common health problems for which alternative diagnosis- and treatment-based DRGs exist. The table includes all health problems that are indications for the principal inpatient surgical procedures reim-

Table 5.5 **Changes in Hospital Treatment Patterns for Common Health Problems in the Elderly**

Surgical Treatment Diagnosis-Related Groups				Alternative Nonsurgical Diagnosis-Related Groups			
Diagnosis-Related Group	Description	Relative Weight	1988 Discharges (% change from 1983 base)	Diagnosis-Related Group	Description	Relative Weight	1988 Admissions (% change from 1983 base)
			Spinal Nerve Compression				
4	Decompression of spinal canal	2.59	4,970 (+46.2)	9	Spinal disorder/injury	1.24	2,570 (−36.9)
					Medical treatment outside hospital		
			Cerebrovascular Disease				
5	Carotid endarterectomy[a]	1.57	47,530 (+8.3)	14	Specific cerebrovascular disorder (except transient ischemic attack)	1.24	328,900 (+9.8)
				15	Transient ischemic attack	0.62	150,395 (−5.3)
				16	Nonspecific cerebrovascular disorder (except transient ischemic attack) with complicating conditions	1.03	
				17	Nonspecific cerebrovascular disorder (except transient ischemic attack) without complicating conditions	0.63	18,470 (−63.9)
					Medical treatment outside hospital	—	

			Coronary Heart Disease				
106	Coronary artery bypass graft with catheterization	5.54	61,815 (+1,260.1)	121	Acute myocardial infarction with complicating conditions	1.72	
107	Coronary artery bypass graft without catheterization	3.99	44,790 (−11.0)	122	Acute myocardial infarction without complicating conditions	1.20	333,540 (+16.3)
112	Percutaneous transluminal coronary angioplasty[b]	1.89	109,830 (+211.5)	123	Acute myocardial infarction, expired	1.40	
124	Non–acute myocardial infarction cardiac catheterization with complicating conditions	1.18	197,175 (+414.5)	140	Angina pectoris (cardiac-related chest pain)	0.67	367,585 (+34.5)
125	Non–acute myocardial infarction cardiac catheterization without complicating conditions	0.69		132	Atherosclerosis with complicating conditions	0.80	23,420 (−92.0)
				133	Atherosclerosis without complicating conditions	0.60	

(*continued*)

Table 5.5 (continued)

Surgical Treatment Diagnosis-Related Groups				Alternative Nonsurgical Diagnosis-Related Groups			
Diagnosis-Related Group	Description	Relative Weight	1988 Discharges (% change from 1983 base)	Diagnosis-Related Group	Description	Relative Weight	1988 Admissions (% change from 1983 base)
			Cardiac Rhythm Irregularity				
115	Pacemaker implantation with complicating conditions[a]	4.05	58,300 (+6.5)	138	Cardiac arrhythmia with complicating conditions	0.85	254,705 (+11.2)
116	Pacemaker implantation without complicating conditions[a]	2.77		139	Cardiac arrhythmia without complicating conditions	0.59	
			Cardiac Valvular Disease				
104	Cardiac valve procedure with pump and cardiac catheterization	7.34	12,015 (+1,869.5)	135	Cardiac congenital and valvular disorders with complicating conditions	0.92	9,140 (−67.1)
105	Cardiac valve procedure with pump, without catheterization	5.78	12,010 (+18.0)	136	Cardiac congenital and valvular disorders without complicating conditions	0.61	
			Small and Large Intestinal Disorders				
148	Major small and large bowel procedures with complicating conditions	3.24	142,750 (+53.1)		None		
149	Major small and large bowel procedures without complicating conditions	1.83					

			Gall Bladder and Biliary Disorders				
195	Total cholecystectomy with contrast dye enhancement with complicating conditions[c]	2.39	28,285 (+284.3)	207	Biliary tract disorder with complicating conditions	0.92	52,475 (−22.9)
196	Total cholecystectomy with contrast dye enhancement without complicating conditions[c]	1.69		208	Biliary tract disorder without complicating conditions	0.58	
197	Total cholecystectomy without contrast dye enhancement with complicating conditions[c]	1.88	102,860 (+9.7)		Medical treatment outside hospital	—	
198	Total cholecystectomy without contrast dye enhancement without complicating conditions[c]	1.12					
			Inguinal/Femoral Hernia				
161	Inguinal/femoral hernia repair with complicating conditions	0.75	78,150 (−31.2)		None		
162	Inguinal/femoral hernia repair without complicating conditions	0.50					

(*continued*)

Table 5.5 (continued)

Surgical Treatment Diagnosis-Related Groups				Alternative Nonsurgical Diagnosis-Related Groups			
Diagnosis-Related Group	Description	Relative Weight	1988 Discharges (% change from 1983 base)	Diagnosis-Related Group	Description	Relative Weight	1988 Admissions (% change from 1983 base)
				Arthritis of Hip			
209	Hip replacement	2.42	209,080 (+78.0)		Medical treatment outside hospital	—	
				Hip Fracture			
210	Open reduction of hip/ femur fracture with complicating conditions	2.18	139,310 (+29.3)	235	Fracture of femur	1.21	47,110 (−23.0)
211	Open reduction of hip/ femur fracture without complicating conditions	1.61		236	Fracture of hip	0.90	
				Back Pain			
214	Intervertebral diskectomy with complicating conditions	2.14	62,550 (+87.1)	243	Back problem	0.67	129,050 (−38.2)
215	Intervertebral diskectomy without complicating conditions	1.38			Medical treatment outside hospital	—	

			Breast Cancer				
257	Total mastectomy with complicating conditions	1.04	59,155 (+44.1)		None		
258	Total mastectomy without complicating conditions	0.85					
259	Subtotal mastectomy with complicating conditions	1.00	7,700 (−28.1)				
260	Subtotal mastectomy without complicating conditions	0.60					
			Benign Prostatic Hypertrophy				
336	Transurethral prostatectomy with complicating conditions	1.08	207,570 (+35.5)	348	Benign prostatic hypertrophy with complicating conditions	0.66	8,685 (−82.3)
347	Transurethral prostatectomy without complicating conditions	0.75		349	Benign prostatic hypertrophy without complicating conditions	0.40	
					Medical treatment outside hospital	—	

(*continued*)

Table 5.5 (continued)

Surgical Treatment Diagnosis-Related Groups				Alternative Nonsurgical Diagnosis-Related Groups			
Diagnosis-Related Group	Description	Relative Weight	1988 Discharges (% change from 1983 base)	Diagnosis-Related Group	Description	Relative Weight	1988 Admissions (% change from 1983 base)
			Cancer (nonsurgical treatments)				
409	Radiotherapy[d]	1.08	8,530 (+85.8)		None		
410	Chemotherapy[d]	0.47	137,890 (+248.3)				
	Surgical discharge rate per 1,000 elderly		191 (+33.6)		Nonsurgical discharge rate per 1,000 elderly		121 (−50.0)

Note: Standard errors in parentheses.

[a]During the time period, new clinical evidence suggested these procedures are often ineffective.

[b]Some other surgical procedures for coronary heart disease account for a small proportion of admissions in this DRG.

[c]Contrast dye enhancement is an intraoperative radiological procedure.

[d]Though these cancer therapies comprise treatment-based DRGs, they are not surgical procedures.

bursed separately under PPS, as well as the alternative nonsurgical DRGs.[16] In general, the use of technologies that constituted treatment-based DRGs for each of these health problems increased substantially between 1983 and 1988. For example, the use of CABG and PTCA as intensive treatments for heart disease doubled between 1983 and 1988, reflecting the increased investment in cardiac surgery capacity and the complete cost sharing documented above. Other intensive DRG-based treatments for health problems that showed substantial increases in utilization rates included valve-replacement procedures as a treatment for heart valve diseases, cholecystectomy as a treatment for gall bladder disease, open reductions in the treatment of hip fracture, diskectomy in the treatment of back pain, and prostatectomy in the treatment of benign prostatic hypertrophy. Hospitalizations in treatment-based DRGs for diseases without alternative hospital treatments—such as joint replacement for arthritis and chemotherapy for cancer—also increased substantially. Altogether, the surgical admission rate for elderly Medicare beneficiaries increased by a third between 1983 and 1988, while the nonsurgical admission rate fell by one-half.

Table 5.6 summarizes changes in admissions and intensity between 1983 and 1988 for some common "single DRG" diseases, those that do not have distinct DRGs for more intensive treatments. Three of these health problems—chronic obstructive pulmonary disease, pneumonia and pleurisy, and bronchitis and asthma—show intensity increases, though the large changes in admission rates for these respiratory diseases imply that underlying coding changes make it difficult to compare the 1983 and 1988 populations. For most of these health problems, however, average intensity was virtually unchanged or fell, and total admissions declined. Treatment intensity growth thus appears much more modest for health problems where hospitals face single prices regardless of intensity decisions.

McClellan (1993) reviews more comprehensive evidence on these intensity trends, confirming that actual hospital expenditure growth appears to reflect the reimbursement incentives illustrated here. In particular, more use of specific intensive procedures (especially surgical procedures) appears to account for all Medicare hospital expenditure growth. These trends do not appear to be part of a general technological imperative toward more intensive treatment. For example, the reimbursement incentives for AMI illustrated that ICU and CCU bed use is associated with relatively little additional reimbursement, and in fact the total number of ICU and CCU beds has been relatively flat since PPS was adopted. These findings, based entirely on aggregate statistical data

16. For many DRGs, separate groups exist for cases with and without "complicating conditions." The DRGs with complications are reimbursed at somewhat higher rates than those without, with the goal of providing "fair" reimbursement for sicker patients within the diagnosis group as well as incentives for upcoding or "DRG creep." Both prices influence hospital investments in intensive technologies. To abstract from coding instability that severely affected admission patterns in DRGs with versus without complicating conditions during this period, I group DRGs with and without complicating conditions together in the descriptive tables.

Table 5.6 Trends in Intensity and Admissions for Health Problems without High-Intensity Treatment Diagnosis-Related Groups

Diagnosis-Related Group	Description	1983 Intensity (DRG weight)	1988 Intensity (DRG weight)	Change in Intensity, 1983–88	1988 Admissions (% change from 1983)
11	Multiple sclerosis/cerebellar ataxia	1.01	0.93	−.08	4,785 (−23.5)
12	Degenerative nervous system disorders	1.11	0.95	−.16	28,840 (−54.9)
22	Hypertensive encephalopathy	0.79	0.70	−.09	11,925 (+15.6)
87	Pulmonary edema and respiratory failure	1.55	1.57	+.02	68,265 (+7.4)
88	Chronic obstructive pulmonary disease	1.04	1.13	+.09	92,275 (−66.2)
89–90	Simple pneumonia/pleurisy	1.08	1.21	+.13	405,760 (+26.7)
96–97	Bronchitis/asthma	0.78	0.93	+.15	266,450 (+31.2)
127	Heart failure/shock	1.04	1.02	−.02	537,875 (+17.6)
128	Deep venous thrombophlebitis	0.86	0.85	−.01	30,725 (−19.0)
174–75	Gastrointestinal hemorrhage	0.91	0.92	+.01	168,810 (+22.0)
176–78	Peptic ulcer disease	0.79	0.80	+.01	38,610 (−34.2)
316	Renal failure	1.33	1.28	−.05	39,935 (−27.7)

on Medicare hospital utilization, are obviously preliminary. But they suggest that hospitals have responded quite dramatically to actual PPS incentives for treatment intensity.

5.5 Conclusions

Exogenous technological progress is viewed as a principal cause of growth in health care costs (Newhouse 1992), and the adoption and diffusion of new technologies desired by patients clearly represents a major component of Medicare hospital expenditure growth. In this paper, however, I have argued that technological change may in fact be endogenous to reimbursement incentives. This argument required a review of the details of hospital production and of a model for capturing static reimbursement incentives adequately, as well as the development of arguments for why the current structure of PPS reimbursement rules may lead to changes in equilibrium intensity choices over time. While the development of a formal model of hospital technology adoption decisions in the context of the model presented here awaits further work,[17] some important implications for technology diffusion are evident. For health problems with multiple DRGs for different levels of treatment intensity, innovations are favored that shift the net benefit distribution to permit "marginal" patients to be treated in more intensive DRGs. For example, innovations that reduce operative mortality and morbidity for specifically reimbursed surgical procedures will be adopted and will lead to dynamic cost increases through more intensive treatment of the disease. For diseases without intensive treatment-based DRGs, such technologies are less likely to be adopted. In these cases, the adoption of cost-reducing technologies may complement the dynamic decline in costs outlined previously.

The theoretical and empirical results presented here suggest that current PPS incentives are unlikely to achieve stated policy goals. In particular, they probably will not limit the use of intensive treatments with small marginal expected benefits. As the previous section and McClellan (1993) suggested, improvements in these incentives appear feasible through simple reforms in the DRG payment structure. First, for health problems for which production is competitive or for which physician incentives are likely to determine equilibrium intensity choices, the regulator's basic goal should be to set a single price for all alternative treatments for the health problem. Indeed, the regulator need only get the aggregate price level right for the set of health problems that rely on similar hospital technologies. Concerns about hospitals engaging in more patient-selection behavior if incentives really were high-powered, or about hospitals going out of business because of the high resulting variance in payments, are legitimate reasons for providing some cost sharing through DRG structure. Such concerns can be addressed by allowing partial but not complete price differentials for DRGs that lead to payment variation based on intensity within a diagnosis. Alternatively, they might be addressed through other features of the payment system. The present system uses "outlier" payments,

17. Most biomedical research is conducted not by the clinical divisions of hospitals but by their biomedical research divisions and separate research organizations. Funding for biomedical research, primarily from the federal government, is largely independent of Medicare funding for patient care. Consequently, technology diffusion rather than innovation itself has greater consequences for medical costs. In a standard model of diffusion, the firm compares its expected profit stream from investing in the innovation during this period to the expected profit stream associated with waiting until the next period (Reinganum 1989). Firm heterogeneity leads to heterogeneity in observed adoption times and frequency of use. Incentives for technology adoption include the reimbursement rules and demand responses associated with the technology. Thus the intuitive prediction of the model developed here is that technological diffusion will tend to be more rapid for the types of technologies favored by the DRG system.

which are essentially supplemental cost-sharing payments for very expensive admissions, to address these issues.

Second, if hospitals have more control over equilibrium intensity for a health problem *and* they face upward-sloping demand curves, then separate intensity-based DRGs may be optimal. However, these DRGs should not be designed to reflect completely the cost differential of producing the more intensive treatment, unless demand is completely inelastic (which seems improbable). Instead, the price differential should partially reflect the cost differential, and the price of the low-intensity treatment can be increased as needed to assure that hospitals will not find it more profitable to forgo treating patients altogether.

References

Ellis, Randall, and Thomas McGuire. 1986. Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply. *Journal of Health Economics* 5:129–52.

Grossman, Sanford J., and Oliver D. Hart. 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy* 94 (4): 691–719.

Harris, Jeffrey. 1977. The Internal Organization of Hospitals: Some Economic Implications. *Bell Journal of Economics* 8:467–82.

McClellan, Mark B. 1993. The Effects of Hospital Characteristics on Medical Treatment Intensity. Massachusetts Institute of Technology. Mimeo.

McClellan, Mark B., and Robert H. Brook. 1992. Appropriateness of Care: A Comparison of Global and Outcome Methods to Set Standards. *Medical Care* 30:565–86.

McClellan, Mark B., and Joseph P. Newhouse. 1993. The Marginal Benefits of Medical Treatment Intensity: Acute Myocardial Infarction in the Elderly. NBER Working Paper. Cambridge, MA: National Bureau of Economic Research.

Newhouse, Joseph P. 1992. The Basis of Reimbursing Health Providers. Harvard University. Mimeo.

Park, Rolla Edward, et al. 1986. Physician Ratings of Appropriate Indications for Six Medical and Surgical Procedures. *American Journal of Public Health* 76:766–72.

Prospective Payment Assessment Commission. 1993. *Report and Recommendations to the Congress.* Washington, DC: Prospective Payment Assessment Commission.

Reinganum, Jennifer F. 1989. The Timing of Innovation: Research, Development, and Diffusion. In *Handbook of Industrial Organization,* ed. Richard Schmalensee and Robert Willig. Amsterdam: North-Holland.

Rogowski, Jeannette R., and Daniel J. Byrne. 1991. Comparison of Alternative Weight Recalibration Methods for Diagnosis-Related Groups. *Health Care Financing Review* 12:87–101.

Starr, Paul. 1984. *The Social Transformation of American Medicine.* New York: Basic Books.

Comment Thomas E. MaCurdy

Reduced to its essence, McClellan's paper represents the beginnings of a research effort designed to address the following issue: viewing the prospective payment system (PPS) as regulatory structure, how should one define diagnostic groups (DRGs) to provide the prerequisite amount of health care while simultaneously controlling costs? To highlight the paper's achievements in confronting this issue, my comment first outlines the basic economic model proposed to describe behavior in the health care industry. I next explain how the paper interprets PPS as a regulatory system. Finally, I offer an annotated list of major enhancements needed to develop a richer framework that offers a basis for understanding the critical factors relevant in fashioning PPS to govern the provision of health care.

Basic Behavioral Model

McClellan's basic economic model operates as follows. Patients develop health conditions of unknown severity. For each illness, two medical procedures are available as treatment: intensive and nonintensive. Patients choose hospitals based on the likelihood of receiving intensive treatment. The intensity of treatment offered by a hospital is determined solely by its investment in technology, with higher investment raising the likelihood of intensive treatment. Hospitals act to maximize profits in a way that balances the expense of investment with the extent to which investment attracts patients. Physicians play a purely passive role. Assigned to hospitals in an unspecified manner, they observe the severity of the health condition afflicting each patient. They mechanically act to provide intensive treatment when a patient's illness reaches a sufficiently high level of severity.

To describe the formal framework in greater detail, initially consider the behavior of the patients. For their health conditions, they can receive one of two treatments: nonintensive, indicated by the discrete variable r taking the value $r = 0$; and intensive, indicated by $r = 1$. The outcome from treatment is a random variable y (my notation) generated by the following rule: $y = y^*$ when $r = 0$; $y = y^* + \theta > y^*$ when $r = 1$, where y^* is a latent variable with a mean equal to μ, and θ is a random variable whose cumulative distribution function is $F(\theta)$, with $F(0) = 0$ (i.e., $\text{Prob}(\theta > 0) = 1$). Both y^* and θ are unobserved by the patient. The outcome expected by the patient when he or she enters a hospital for treatment is

$$E(y) = \mu + E(\theta \mid r = 1)\, p(r = 1).$$

The patient chooses that hospital offering the highest expected value.

Physicians play a routine role. In effect, they act to maximize patients' wel-

Thomas E. MaCurdy is professor of economics at Stanford University and a research associate of the National Bureau of Economic Research.

fare subject to a hospital's treatment capacity. They observe θ and implement $r = 1$ whenever $\theta > \Delta$, where $\Delta = \Delta(I)$ is a threshold defined by a hospital's level of investment, I, in capital and technology. With this mechanism, the expected value of an outcome perceived by a patient becomes

$$E(y) = \mu + E(\theta \mid \theta > \Delta)\,\text{Prob}(\theta > \Delta).$$

The function $\Delta = \Delta(I)$ is a declining function of I. So a hospital makes itself more attractive to patients by raising the value of I and thereby lowering the value of Δ.

Finally, the hospital's behavior receives extensive consideration covering several circumstances. In all of these circumstances, a hospital's objective is to choose a level of investment I (or equivalently the value of Δ) to maximize its net revenue. Consider the two types of procedures. The nonintensive treatment receives revenue equal to p_0 at a cost equal to c_0, and the intensive treatment receives revenue p_1 at a cost c_1. The costs c_0 and c_1 may be nonincreasing functions of I. Higher levels of investment make the use of the intensive treatment more likely by lowering the value of Δ. McClellan considers four categories of investment or choices of technology.

1. In the case of "specific fixed capacity," the hospital chooses I to maximize profits given by

$$\Pi(I) = [F(\Delta)(p_0 - c_0) + (1 - F(\Delta))(p_1 - c_1)]q(\Delta) - bI,$$

where $\Delta = \Delta(I)$, $F(\Delta)$ is the fraction of patients who receive the nonintensive treatment, q is demand for the hospital's services and equals the number of patients, and b is the cost of a unit of investment. In optimizing this problem, the number of treatments cannot exceed a fixed capacity; satisfaction of the constraint $(1 - F(\Delta))q(\Delta) \leq I$ captures this condition. The demand q is a declining function of Δ, and therefore it is an increasing function of I.

2. In the case of "nonspecific fixed capacity," investment is a joint input into treating several classifications of health conditions. Using index k to designate these classifications, the hospital chooses I to maximize

$$\textstyle\sum_k \Pi_k(I),$$

where each Π_k takes a form analogous to Π in the previous expression. In this optimization problem, satisfaction of the constraint $\sum_k (1 - F_k(\Delta))q_k(\Delta) \leq I$ ensures that the number of procedures does not exceed a fixed capacity for all available treatments.

3. In the case of the "line of business," investment is lumpy and is either undertaken at the level $\bar{I}$ or not at all. Thus, the hospital chooses $I = \bar{I}$ if $\Pi(\bar{I}) > \Pi(0)$, and $I = 0$ otherwise, where Π once again looks like the specification in category 1.

4. Finally, the "intermediate technology" case is an unspecified variant of cases 1 and 3.

The first-order conditions associated with the optimization of categories 1–3 provide the primary motivation for the behavioral claims made in the paper.

Regulatory Structure

The basic regulatory problem in PPS is to define DRGs in a way that balances the provisions of health care with the costs of supplying this care. Within the framework outlined so far, there are two options for specifying DRGs. One can combine the intensive and the nonintensive procedures into a single DRG. This means that the hospital receives the same revenue regardless of which treatment prevails, and therefore, $p_0 = p_1$. Alternatively, one can categorize the two treatments as distinct DRGs, in which case $p_1 > p_0$.

Many health-policy decision makers and researchers believe that DRGs associated with a particular illness are primarily covered by the first option, where compensation does not depend on the intensity of treatment. However, McClellan usefully points out that in fact a large fraction of DRGs represent applications of the second option. This is especially true when considering surgical treatments. A surgical procedure often is defined as its own DRG. McClellan uses heart attacks as an illustration. A patient experiencing heart attacks can be categorized into any one of many DRGs: three represent varying degrees of intrusive treatments involving cardiac surgery and an even greater number represent nonintensive classifications such as drug treatments. The resulting practice looks very much like a cost-based reimbursement system, which PPS was designed to replace.

McClellan provides some intriguing evidence suggesting that this feature of the current system is a major factor in explaining the growth in health costs since the introduction of PPS. A related paper (McClellan 1996) provides stark evidence supporting this argument. Around the time of the introduction of PPS in the early 1980s, the trend in nonsurgical discharges for the elderly reversed direction from a continuous increase to a steady decline. At the same time, surgical discharges rose steadily and at a higher rate after the introduction of PPS. Because nonsurgical treatments tend to receive lower reimbursements on average, such evidence supports the view that medical practitioners treat patients using intensive procedures in order to justify a more lucrative DRG classification. McClellan further notes that the number of DRGs assigned to each admittee has steadily increased, with index per beneficiary rising by 20% in recent years.

The problem of creating an optimal classification scheme for DRGs is quite complicated, and it receives only peripheral discussion in the paper. The tradeoffs are clear. Grouping various medical procedures into one DRG means that the more expensive treatment will only be applied if there are substantial advantages to its implementation compared to other alternatives in the DRG for the patient in question. On the other hand, such a grouping lowers the incentives of hospitals to invest in the technology needed for more intensive procedures, which lessens their use for marginal patients who could benefit. The model provided in McClellan's paper does not offer a rich enough structure to solve the problem of defining DRGs from a regulatory perspective. In particular, the model needs to specify the willingness of informed patients to pay for more intensive treatments.

Proposed Development

McClellan's research to date provides a preliminary structure to assess how DRG definitions balance the trade-offs between costs and the provision of health care. However, there are several obvious dimensions for augmenting the current model to better address the basic regulatory problem of reforming PPS. I briefly mention four such enhancements.

1. McClellan's existing formulation is a partial equilibrium model. It is incomplete in two important respects: first, it does not specify the factors determining the demand for health care q; and, second, it is silent about the industrial organization underlying the behavior of medical service providers.

In McClellan's model, the demand for medical care is determined by the function q, which in turn depends only on the likelihood of receiving intensive care. Such a formulation may have some appeal in describing Medicare demand in the current system, because in theory price does not influence the demand for medical care by Medicare patients. However, even in defining DRGs in the current system, it is necessary to introduce a demand for health care that specifies patients' willingness to pay for various procedures. Without it, one has no basis for comparing the merits of alternative structures. The need to elaborate the demand side becomes even more obvious when one considers additional aspects of the regulatory problem such as copayments.

Adding an industrial-organizational component to the model specifies hospitals' assumptions about the actions of other health care providers. There are several standard formulations, including perfect competition or oligopoly models (e.g., Cournot). A broader range of formulations may be appropriate for describing the behavior of nonprofit hospitals. Industrial organization will play a critical role in any consideration of the regulatory problem of the health care industry. It will not only determine the tendency of hospitals to overinvest in technology from a societal perspective—as happens in patent races, a well-known example in the industrial-organizational literature—but it will also play a key role in determining the specification of the demand curves (q) faced by individual health care providers.

2. McClellan's current model does not adequately specify the behavior of physicians. Physicians are not robotic agents acting to maximize patients' welfare subject to hospitals' costs conditions. Instead, physicians act to optimize their own benefits, and indeed their actions may sacrifice the welfare of both the patients and the hospitals. One often suspects that the use of particular medical procedures may primarily reflect the compensation received by doc-

tors, with only a peripheral link to patients' welfare or to hospitals' costs. Critics of such behavior often cite unnecessary surgery as a prime example.

Introducing physicians' incentives into McClellan's model is not straightforward and will undoubtedly result in significant changes in behavioral implications. A conflict arises among patients', physicians', and hospitals' incentives. Since hospitals do not observe θ or aspects of Δ, they face a principal-agent problem in their interactions with physicians, with the hospital playing the role of the principal. Similarly, patients face a principal-agent problem in their interactions with physicians and hospitals. These are important behavioral relationships that are central to understanding the effectiveness of PPS regulations.

3. Accounting for the multiproduct nature of hospitals would provide another important enhancement to McClellan's model. Most medical providers, and certainly hospitals, produce an array of products simultaneously. Such considerations introduce significant complications in the analysis when properly incorporated into an economic framework. The section of the paper dealing with nonspecific fixed-capacity investment highlights the added complications of considering multiple products produced by hospitals. Because many technologies affect the availability of a spectrum of medical procedures, not only is it true that investment enriches the capability of a hospital to perform a group of sophisticated procedures, but it is also true that a substantial degree of substitutability may arise among the procedures making up this group. Even more important, the multiproduct nature of hospitals makes it exceedingly difficult to assign or measure costs of particular treatments, a critical assumption maintained throughout McClellan's formulation.

4. Finally, introducing an explicit formulation for the economic problem faced by regulators is vital if one wants to study the consequences of alternative PPS schemes. This involves specifying the overall objectives that regulators, patients, physicians, and hospitals act to optimize, as well as the information sets available to all agents. The resulting problem faced by the regulator is quite complicated, with several game-theoretic and principal-agent structures possible to describe behavior in the health care industry. Analyses of alternative PPS schemes will undoubtedly depend on the particular formulation for the regulatory problem. The optimal structure for DRG categories under the assumption that physicians act to maximize patients' welfare and hospitals act to maximize profits will look quite different than under the assumption that agents act to satisfy other objectives. A vast literature exists on the regulation of public utilities (e.g., *A Theory of Incentives in Procurement and Regulation* by Jean Laffont). This body of research offers a rich source for analyzing alternative PPS structures.

Conclusion

McClellan's paper is an enlightening contribution to a topic that is central to the upcoming debate on the financing of health care. I know that McClellan is planning to continue with this research project, and I view it as some of the most promising work in the health-economics area. In my comments, I have taken a time-honored approach of suggesting generalizations of the approach proposed in the paper. Introducing these generalizations will undoubtedly require significant effort and a comprehensive long-term research plan, but I expect large payoffs from this activity.

Reference

McClellan, Mark B. 1996. Why do hospital costs keep rising? A model of hospital production and optimal payment regulation. NBER Working Paper. Cambridge, MA: National Bureau of Economic Research. Forthcoming.