

# Why Is There A Quality Chasm?

The barriers to good performance are more fundamental than simply the lack of organized systems.

**by Joseph P. Newhouse**

**ABSTRACT:** Medical care seems to obtain less value from the resources it uses than other industries do, a phenomenon not limited to the United States. I explore several reasons for this, including consumers' ignorance, the rate of technological change, the widespread use of administered pricing, the difficulty of appraising a given provider's quality, and the role of the public sector with objectives other than efficiency. Although these causes suggest that the performance of medical care may always lag behind that of other industries, greater use of information technology and improved financial incentives will help to reduce the size of the quality chasm.

**I**F ONE WAKED A TYPICAL PROFESSIONAL ECONOMIST in the middle of the night and asked whether industries generally produce at something approximating minimum cost, most would probably answer yes. If the industry is competitive, a standard starting assumption, high-cost firms will be driven from the market. Even if the market is not competitive, a firm can always increase profits by reducing its costs.<sup>1</sup> And when managers in a noncompetitive industry seek to lead the quiet or high life rather than relentlessly decreasing costs, their firm may face new entrants or a takeover attempt.

Economists—and noneconomists as well—also assume that if a better mousetrap is built, consumers will stop buying the old one unless the additional price for the new one is unjustified. Products whose quality is not worth their price go the way of Soviet-era automobiles in the post-Soviet era. This is simply an extension of the standard competitive assumption into competition among products. In short, the competitive assumption, whether with respect to price or product, is one of the pillars upon which economists build their belief that most industries, most of the time, get as much as they can for the quantity and quality of inputs that they use.

Because of the pervasiveness with which competition winnows out firms whose products do not justify their price, economists have devoted only modest

.....  
*Joe Newhouse is the John D. MacArthur Professor of Health Policy and Management at Harvard University. He also is professor of health care policy at Harvard Medical School and vice-chair of the Medicare Payment Advisory Commission (MedPAC).*

*“Despite the lack of a summary measure of its efficiency, many seem convinced that the industry’s performance falls short.”*

.....

efforts to quantifying the degree to which industries and firms do not get as much as possible from the resources they employ.<sup>2</sup> By contrast, much more of the health services research literature has gone into documenting the shortcomings of the medical care industry in producing health. The well-known recent report from the Institute of Medicine (IOM) termed the gap between the actual and potential performance of the U.S. health care system a “quality chasm.”<sup>3</sup> I believe that most economists would not describe most industries in this fashion. That judgment may be wrong, but for this paper I assume that medical care is inefficient and that it fails to produce as much health as it might with the resources it uses, and I ask why.

One can raise two immediate and related important questions about this indictment of medical care. First, perhaps health is the wrong output. One could ask only whether physicians and hospitals are efficient at producing narrowly defined medical services such as office visits and hospital stays rather than health. For example, the efficiency of computer makers is judged by how cheaply they make computers of varying power; one usually does not ask whether users need the additional power. But if one assumes that consumers who choose to pay more for a more powerful computer value the additional capacity more than its cost, the relevant issue is whether the computer was made in a least-cost fashion. Because of the consumer’s ignorance and the resulting agency relationship with physicians, as well as widespread insurance coverage, the same deference is not as readily given to consumers’ preferences in medical care.

Second, one may grant that medical care is not performing well, but is it really performing markedly worse than other industries perform? Although high rates of negligent error and repeated tests, as well as long waiting times, paint a picture of poor performance, they scarcely constitute a summary measure of efficiency, let alone a measure that can be compared against other industries.<sup>4</sup>

Finding measures with which to compare the efficiency of industries is difficult. Comparisons are somewhat easier if medical care services rather than health are the relevant outputs. Then one might employ a variant on the notion of best practice: Ascertain the quantity of output (for example, hospital stays) across various firms and the quantity of inputs each firm uses, and determine how much more could be produced from the same total inputs if all firms produced at the level of the best-performing firms. Then determine if medical care in the aggregate falls short of best practice by more than other industries do. But such comparisons require accounting for differences in the quality of different firms’ output and inputs, a hard task.<sup>5</sup> Further, the market for most medical services is local; inherent differences in scale and modes of treatment complicate comparing the efficiency of a small rural hospital with that of a large a teaching hospital, not to mention a

solo general practitioner with a subspecialist in a large group.

If health rather than medical services is the output, problems are magnified. Health cannot be measured in units that are commensurate with outputs of other industries. And one wants to know the value added of medical care, something not likely to be known with much precision.

Despite the lack of a summary measure of its efficiency, many seem convinced that the medical industry's performance falls short. I begin by briefly reviewing some of the better-known evidence supporting this conviction. I then turn to my main purpose, explaining this poor performance. The explanations I discuss seem largely inherent in the delivery of medical services, which implies that substantial improvement will not be easily achieved.

### **Inefficiency In Translating Medical Care Services Into Health**

One of the first indications of inefficiency came from the vast literature on geographic variations in use of treatment, with an implied assumption of little or no variation in outcomes. U.S. variations are now presented in great and colorful detail in the well-known *Dartmouth Atlas*.<sup>6</sup> And such variations are not confined to the United States.<sup>7</sup>

Also, these variations cannot be accounted for by differences in health status.<sup>8</sup> Although variations in patients' preferences and factor prices have been less studied, it seems implausible that they could differ by enough to explain the magnitude of differences in utilization across areas. As a result, the usual interpretation of this variation is that many, perhaps all, areas are producing health inefficiently.

The early health services research literature tended to assume that the low-rate regions had it right. In the mid-1980s, however, Mark Chassin and others showed that this presumption was incorrect and in so doing brought forth more compelling evidence of inefficiency than had hitherto existed. Chassin and his colleagues defined procedures to be appropriate if the "expected health benefits of a procedure exceed its expected negative consequences by a sufficiently wide margin that the procedure is worth doing."<sup>9</sup> Conversely, inappropriate procedures had little or no expected benefit—or even a negative benefit. Physician panels assigned appropriateness ratings for patients with varying indications, and then information from a sample of charts was used to ascertain each patient's indications.

The initial studies showed that a sixth to a third of the procedures performed were inappropriate. An additional number were equivocal. These magnitudes certainly suggest a substantial problem. And the bad news was not limited to the United States. In the Trent region of the United Kingdom, for example, the rate of inappropriate coronary angiography was 51 percent, and the rate of inappropriate coronary artery bypass graft (CABG) was 42 percent. In four Israeli hospitals the rate of inappropriate or equivocal cholecystectomy was 29 percent.<sup>10</sup> High rates of inappropriateness have not been found in all studies, but they do predominate. Moreover, an economist would find even higher rates of inefficiency, since the

economist would consider medical care whose marginal benefit was positive but less than marginal cost to be inefficient, whereas the Chassin definition considers such care to be appropriate.

By setting a normative standard rather than simply observing that every region could not have it right, the studies of appropriateness strengthened the inference of inefficiency from the variations literature. Moreover, the studies by Chassin and others tended to find similar rates of inappropriate care among areas with widely varying overall procedure rates, implying underuse in low-rate areas and overuse in high-rate areas.

Later and larger studies produced more evidence of poor quality, namely variation across states in measures of proper care for a given condition. Not only did states at the tenth and ninetieth percentiles exhibit considerable spread, the median state was distressingly far from 100 percent (Exhibit 1). Moreover, these data, as well as the variation shown in the *Dartmouth Atlas*, almost certainly understate the problem, because they do not account for within-area variation.

### Why The Poor Performance?

I point to five possible causes of poor performance: consumers' ignorance; the rate of technological change; the role of administered prices; the difficulty in assessing the performance of a given provider; and the role of the public sector and objectives other than efficiency.

■ **Consumers' ignorance.** Although consumers will not necessarily understand the technical details around production of other goods or services any better than they understand the technical details of medical care, they can judge how well they like the ride and performance of a car, for example, or the quality of sound from a sound system. By contrast, they may not be able to distinguish whether a bad medical outcome is attributable to poor-quality care or to the underlying disease. Furthermore, for many acute medical problems there is little or no repeat buying, so

---

**EXHIBIT 1**  
**Variation In Process Measures Of Quality Care Across States, U.S. Medicare Population, 1997-1999**

| Quality indicator  | 10th percentile | Median | 90th percentile |
|--|-----------------|--------|-----------------|
| Beta-blocker after AMI discharge   | 53%             | 64%    | 76%             |
| Aspirin within 24 hours of AMI admission   | 78              | 84     | 88              |
| ACE inhibitor at discharge if congestive heart failure and ejection fraction < 40% | 59              | 71     | 81              |
| Warfarin at discharge for atrial fibrillation                                      | 46              | 55     | 61              |
| Hemoglobin A1c every year if diabetic  | 58              | 71     | 82              |

**SOURCE:** S.F. Jencks et al., "Quality of Care Delivered to Medicare Beneficiaries: A Profile at State and National Levels," *Journal of the American Medical Association* 284, no. 13 (2000): 1670-1676.

**NOTES:** AMI is acute myocardial infarction. ACE is angiotensin converting enzyme.

consumers may have little experience with their specific problem or provider. For both reasons consumers may continue to use providers or delivery systems that give inferior results rather than gravitating toward those with better results and leaving others to fail.

Three types of evidence suggest that the Darwinian process found in most markets does not operate as ruthlessly in medical care. First, publishing the results of substantial variation in cardiac surgery mortality among New York and Pennsylvania hospitals did not provoke patients to change hospitals.<sup>11</sup> Thus, cardiac surgery patients did not behave like 1970s U.S. automobile buyers, who deserted GM, Ford, and Chrysler for Toyota, Honda, and Nissan. Poorly performing hospitals and their medical staffs often did respond constructively to the information, but their responses appeared more motivated by professional ethics than by any actual loss of business, contrasting sharply with the U.S. automobile industry's near-death experience in the late 1970s.

Second, malpractice rarely causes a claim. Although the popular impression is a nation awash in malpractice litigation, only a small proportion, perhaps 1–2 percent, result in a claim.<sup>12</sup> Only one-third of cases with much at stake—patients under seventy years of age who either died or had a disability lasting longer than six months as a result of negligence—brought a claim.<sup>13</sup> These patients, or their heirs, were likely leaving substantial monies on the table. Although there could be several explanations, such as unwillingness to sue their physician, many patients may have simply been unaware that their care fell below professional standards and attributed the adverse outcome to the underlying disease.

Third, patients with higher cost sharing, including large deductibles, do not search for lower-price physicians to any greater degree than do patients with no cost sharing.<sup>14</sup> Because the patient with a large deductible keeps any money he or she might save by using a lower-price physician, this is contrary to the expected behavior in a more standard market.

Consumers' ignorance, however, cannot be the sole explanation of poor performance. As Kenneth Arrow emphasized in his seminal paper four decades ago, because of their ignorance, patients rely on physicians to act as their agent, so the issue becomes why some physicians apparently do not carry out this role in exemplary fashion.<sup>15</sup> In fact, there are several reasons.

■ **Technological change.** Technological change in medicine takes many forms, including the development of new devices, drugs, or procedures, as well as the adaptation of existing procedures or drugs for new patients. Many illustrations can be found. Coronary stents exemplify a new device. They came onto the market in the 1990s to help prevent restenosis (occlusion) of the coronary artery after angioplasty. Although rarely used in 1994, by 1998 around half the angioplasties in seven of nine areas in various countries used stents.<sup>16</sup>

The pace of new drug development and new procedures is similarly swift. Between 1990 and 2000 nearly 1,000 new drugs were introduced into the U.S. mar-

*“Rapid change makes knowledge quickly obsolete and places a heavy burden on mechanisms that enable physicians to keep up.”*

.....

ket, and the number of new molecular entities introduced exceeded 300.<sup>17</sup> Just since 1990 the number of cancer drugs in the pipeline has increased from 28 to 402; in 1990 there were six cancer agents in Phase I trials, and today there are 150–200.<sup>18</sup> The use of catheterization to treat elderly heart attack patients in the U.S. Medicare population increased from 11 percent in 1984 to 41 percent in 1991.

New drugs, devices, and procedures are easily recognized as change. A less well recognized form is the learning that physicians, especially surgeons, acquire as they employ new procedures. Proficiency rises with familiarity, and physicians become more willing to perform the procedure on clinically riskier patients.<sup>19</sup> Discovering effective off-label uses of drugs represents analogous learning.

Another indicator of the increase in clinical knowledge (and its perceived value) is the increased resources devoted to clinical trials. In constant 2000 dollars, spending on trials by the National Institutes of Health (NIH) more than doubled between 1990 and 2000, from \$875 million to \$1.9 billion.<sup>20</sup>

Rapid change makes knowledge quickly obsolete and places a heavy burden on mechanisms that enable physicians and other health professionals to keep up. The profession's main formal instrument for keeping current is continuing medical education (CME). However, the usual CME conference has little effect, and more-promising strategies are seldom used.<sup>21</sup> The IOM *Quality Chasm* report, which also emphasizes the rate of change in knowledge as a cause of poor performance, points toward a more systems-oriented approach and greater use of information technology to help practitioners cope and to make knowledge diffusion more rapid and more uniform.

Numerous well-known obstacles loom to the successful implementation of such a strategy, obstacles that I do not dwell on here. Even ignoring those issues, the rate of change causes a more fundamental problem, as emphasized by Barbara McNeil in her recent Shattuck lecture.<sup>22</sup> The problem is illustrated by work of Edward Guadagnoli and his colleagues, who measured compliance with the guidelines for angiography following an acute myocardial infarction (AMI). Their initial study simply gave more evidence of poor performance. When patients had characteristics for which the guidelines held angiography to be “useful and effective” (Class I), only 46 percent actually received the procedure in traditional fee-for-service (FFS) Medicare. At the other extreme, when patients had characteristics for which the guidelines held angiography to be “ineffective” (Class III), 13 percent of cases received angiography.<sup>23</sup> Greater use of systems and information technology could bring these rates more in line with evidence-based medicine.

However, these remedies cannot address the difficulty posed by the high rate of change for evaluating medical capabilities at any point in time; evidence-based

medicine labors under the onslaught of new knowledge. In a subsequent study Guadagnoli and his colleagues divided the United States into ninety-five regions and examined the degree to which various categories of AMI patients accounted for variation across the regions in angiography rates. Most of the variation was not in the Class I and Class III categories just described, but rather in two more ambiguous categories in which the guidelines judged angiography to be either “appropriate, but not necessary” or “uncertain.”<sup>24</sup> But it is precisely for the patients in the latter two categories that clinical trial data on the efficacy of angiography are unavailable or unconvincing.

In short, it is not just that some health care providers fail to keep up; at any point in time, some procedures are sufficiently new that their efficacy has not been established for substantial numbers of patients. Larger trials, of course, would permit better measurement of efficacy in subclasses of patients; underrepresentation of the elderly, women, and children in existing trials are well-known examples. Larger trials are a problematic solution, however, because accruing sufficient numbers of patients in various subclasses may delay knowledge for those classes of patients where results are more definitive.<sup>25</sup> Furthermore, by the time any trial is complete, a better procedure or drug may have appeared. Or physicians may have become more proficient at the procedure being tested. Either way, the results of the larger, more expensive clinical trial would be obsolete. And any delay to accrue more patients simply increases the chance that the results from the trial will be out of date when they appear.

A more mundane problem that greater use of information technology does not address, although greater use of systems might, is the outmoded specialist. Knowledge accretion results in greater specialization, a process that has gone on for centuries in all fields and disciplines. Specialists, however, cannot or do not always retrain to use new methods, in part because the new methods may be the province of another specialty. For example, in an earlier era some of a general surgeon's business was gastric ulcer surgery; treatment is now generally by drugs and not by general surgeons. Although surgery for gastric ulcers has essentially disappeared, in other cases the specialist may simply continue to perform an outmoded procedure because it is effective and is what he or she does to earn a living. In different settings this is termed *featherbedding*, a reminder that inefficiency is not limited to medical care.

Although a rapid rate of technological change surely has something to do with the poor performance of medical care, it cannot be the entire story either, since other industries with rapid technical change exhibit much different performance. As everyone knows, technological change in the semiconductor and computer industries has been rapid; between 1971 and 1999 the number of transistors per chip increased 10,000 times.<sup>26</sup> Between 1974 and 1996 the price of memory chips, adjusted for this phenomenal change in capabilities, decreased by a factor of 27,270 times, a staggering 41 percent per year.<sup>27</sup> Prices of logic chips, the data for which

have been available only since 1985, fell an even greater 54 percent per year in the 1985–1996 period. Although these figures do not directly show that high-defect—implying high-cost—producers have not survived, that seems likely.<sup>28</sup> Furthermore, over this period dynamic random access memory (DRAM) and metal-oxide-silicon (MOS) logic chips became commoditized, implying that quality was nearly uniform.<sup>29</sup> Why do these two industries have performance that is so different from medical care?

■ **Administered prices.** A critical difference between medical care and many other sectors of the economy, including semiconductors and computers, is the widespread use of administered prices to pay medical providers.<sup>30</sup> The need for administered pricing arises because health insurers, whether public or private, cannot agree to reimburse any price a provider names. If providers are not to be excluded from reimbursement on the basis of price, as is generally the case with public insurance and was de facto the case in U.S. private insurance for many years, administered prices must be used. In some countries the price may be implicit, as in the case of a hospital with a fixed budget, but in other countries an explicit price per unit of service is either set by the government or negotiated industrywide. In the U.S. Medicare system, for example, Congress for the most part simply legislates prices; in Canada and Germany physician fees are negotiated between the profession and a public or quasi-public entity.

Because pricing affects providers' behavior, getting the administered prices set at the right level is important.<sup>31</sup> Paying above marginal cost for a defined service offers incentives for overly intensive care, and paying less does not elicit supply.<sup>32</sup> But getting the right price is difficult for many reasons. Marginal cost must be estimated from econometric or engineering (time and motion) studies, which are likely to be imprecise. Because it is easier to compute, actual price setting tends to aim at average rather than marginal cost. Estimating average cost requires knowing only total cost and number of services, something that normal accounting practice will reveal and that can be independently audited. And there is an even more basic problem. Costs, both average and marginal, adapt to what is paid.<sup>33</sup> As a result, there is mutual causation between observed costs and reimbursement.

Not only the level but also the basis of price matters for quality. Consider hospitals that are paid on the basis of the hospital day or the hospital stay, as most U.S. hospitals are. Suppose that the hospital can purchase a medical supply that reduces the incidence of adhesions after surgery and hence the need for readmission. If the hospital purchases the supply, not only will the cost of surgical procedures be higher, but the hospital will lose the revenue it would otherwise have received from the readmission. Similarly, it will lose revenue from the prolonged stays and readmissions caused by fewer medical errors if it invests in a computerized drug order entry system.

This problem has led some to advocate paying health care providers partly on the basis of results achieved. Such methods, however, unless adequately standard-

*“The sample of patients of any provider is often too small to draw reliable inferences about that provider’s performance.”*

ized for differences in the patient mix of various providers, could promote adverse selection; the provider’s incentives, for example, would be to shun the frail or noncompliant patient if reimbursement were tied to raw outcome measures. And techniques for carrying out such standardization are still underdeveloped, another illustration of the difficulty of getting administered prices right. Moreover, if only certain outcomes were rewarded, resources could simply be diverted from areas with unmeasured results, to no overall benefit.

■ **Technological change and administered pricing interact.** The rate of technological change complicates all administered pricing methods. Costing studies of particular treatments may be quickly obsolete. Trial-and-error pricing, the usual approach, does not function well if techniques change frequently or there is learning by doing. In short, rapid technological change and administered pricing interact to produce poorer performance than either would individually.

Administered pricing also affects performance by affecting the rate of technological change. In a standard market, the manufacturer of a new product simply offers it at a given price, and the market either accepts or rejects it. In the case of medical care, however, the manufacturer of a new drug or device must obtain the approval of the relevant regulatory body to market the product at all. After obtaining approval, the manufacturer, to be paid, must convince insurers to cover the product, as must a physician employing the new procedure. In FFS systems a code for the product or procedure must sometimes be issued by another body to implement a coverage decision. All of this introduces delay relative to other markets, which in turn reduces the expected reward for an investment in development. Potentially offsetting the effect of this delay, however, is the profit to be gained if the procedure is covered, because the moral hazard from insurance coverage will induce greater use than in a standard market. The net effect on innovation is unclear, but these institutions differentiate health care from other markets.

■ **Difficulties of measuring performance.** Insurers, whether private or public, might seem at first blush to offer some help in improving performance. After all, they market directly to purchasers, whether employers or individuals. Why are they not more like hotel chains, which acquire a reputation—not always favorable—for the quality of their lodgings? Part of the reason is a reprise on the theme of consumers’ ignorance. If purchasers do not reward higher-quality health plans, we should not be surprised to see quality problems. (In the context of national health services, reward takes the form of more votes for politicians who support improved quality.) But like providers with an information problem in keeping current, health insurers have an information problem in identifying high-quality providers.

Many manufacturers can readily compute defect rates from alternative suppli-

ers because they typically purchase large quantities of a supply made to given specifications. Health plans, however, face the constraint that the sample of patients of any provider is often too small to draw reliable inferences about that provider's performance for a particular disease.<sup>34</sup> Plans could, of course, measure the performance of an organized delivery system or aggregation of providers instead, using Health Plan Employer Data and Information Set (HEDIS) or Consumer Assessment of Health Plans (CAHPS)-type measures, but this leaves the possibility of within-system variation in performance.

Implicit in the *Quality Chasm* report's call for greater use of organized systems is that management of these systems could reduce within-system variance in performance to minimal levels. To do so requires that managers not only be able to measure the performance of providers but also have the incentive to reward the better performers. Both requirements are problematic for reasons already described. It is not obvious that the organized delivery system is much better placed than the plan is to overcome the problems of small samples and difficulty of risk adjustment. Nor is it clear that the marketplace, whether economic or political, will reward plans that restrict choice to better-performing organized systems.

The existing entities with the strongest incentives and tools to minimize cost for a given quality—or equivalently maximize quality for a given cost—are firms that integrate insurance and delivery, such as U.S. group- and staff-model health maintenance organizations (HMOs). But these entities do not appear to have any pronounced superiority (or inferiority) in their quality of care, which suggests that the barriers to good performance are more fundamental than simply the lack of organized systems.<sup>35</sup> This inference is consistent with group- and staff-model HMOs' failure to thrive in the marketplace.<sup>36</sup> In short, although the creation of organized and accountable delivery systems may be a necessary condition for improvement, the existing evidence suggests that it is not sufficient.

Nonintegrated insurers in the United States, as well as in Canada, Germany, and Japan, generally defer to a physician's judgment about which services should be reimbursed. As a result, nonintegrated insurers implicitly delegate the responsibility for quality to physicians and other providers. From the nonintegrated insurer's perspective, deferring to the physician's judgment, although it may raise cost by covering inappropriate or overly intensive care, is consistent with the desire to maintain a reputation for a product that offers risk protection. As has become clear from the backlash against managed care, the insurer that seeks to inhibit inappropriate care will often incur the wrath of the patient, who generally believes that the physician, not the insurer, is his or her agent. After all, the insurer is a corporation or a public agency, not a professional, and the insurer does not take the Hippocratic oath. And the consumer cannot be certain that the care the insurer seeks to inhibit is in fact inappropriate.<sup>37</sup>

■ **Role of the public sector.** The health care financing systems of all developed countries have an important public-sector component. Because efficiency is not the

only goal of such systems, one should not expect the same performance as in standard markets.

Viewed from the local community, health care financed with federal or state taxes is an export good, as is care financed by premiums if premium payers are geographically dispersed. Local legislators will therefore seek to maximize funds coming to providers in their districts. In rural and inner-city areas health care spending may also serve community development purposes; the hospital may be among the largest employers in the local area. Although it may be possible to reduce costs and improve the quality of both medical and defense services by closing either a hospital or a defense base, both are notoriously difficult to close.

Furthermore, all developed countries regulate entry into the health care professions. The regulations necessarily specify who may perform certain tasks. Such regulations probably inhibit delegating tasks to allied health personnel in cases where delegation would improve performance.<sup>38</sup> Needless to say, the regulations are vigorously defended by those advantaged by them.

### Shrinking The Chasm

The barriers to improvement described above suggest that a medical care quality chasm will always be with us. Nonetheless, the chasm does not have to be as large as it is now. Greater use of information technology can help; if a patient's medical history and all available test and medication data were available online at the time a physician was making a diagnostic or treatment decision, quality would surely improve. Greater use of computerized decision support systems also would improve quality.

Health services research can also help. Most of the evidence of inefficiency I cited came from health services research. When the scope of the problem is not known, better performance is improbable. In some cases, simply disseminating the findings can improve matters through the goodwill, altruism, or professionalism of health care providers. And research on financial incentives could play an important role. Physicians want to practice good medicine. But there are costs to keeping up, and in many cases the rewards for using the best technique are weak or even negative. The design of better incentives thus should be a high priority.

.....  
*The author thanks the Alfred P. Sloan Foundation and the Hans Sigrist Stiftung for support and David Cutler, Victor Fuchs, Tom McGuire, Don Metz, and two referees for comments on a preliminary draft.*

#### NOTES

1. Agency problems between shareholders and management may lead management to spend potential profits as perquisites. M. Jensen and W. Meckling, "The Theory of the Firm: Managerial Incentives, Agency Costs, and Ownership Structure," *Journal of Financial Economics* 3, no. 4 (1976): 305-360.
2. Recently effort has gone into understanding productivity differences at the firm level. M. Baily and R. Solow, "International Productivity Comparisons Built from the Firm Level," *Journal of Economic Perspectives* 15, no. 3 (2001): 151-172. Economists also look at efficiencies in the context of an antitrust case involving a merger, where claimed efficiencies are often the rationale.

3. Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the Twenty-first Century* (Washington: National Academy Press, 2001).
4. *Ibid.*, and L.T. Kohn, J.M. Corrigan, and M.S. Donaldson, eds., *To Err Is Human: Building a Safer Health System* (Washington: National Academy Press, 2000).
5. For skeptical comments on the applicability of this approach to assessing variation in efficiency across health care providers, see J. Newhouse, "Frontier Estimation: How Useful a Tool for Health Economics?" *Journal of Health Economics* 13, no. 3 (1994): 317–322.
6. Center for Evaluative Clinical Sciences, Dartmouth Medical School, *The Quality of Medical Care in the United States: A Report on the Medicare Program, The Dartmouth Atlas of Health Care 1999* (Chicago: AHA Press, 1999).
7. See K. McPherson et al., "Small-Area Variations in the Use of Common Surgical Services: An International Comparison of New England, England, and Norway," *New England Journal of Medicine* 307, no. 21 (1982): 1310–1314.
8. For example, controlling for clinical characteristics, there was a much higher rate of coronary angiographies and revascularizations following a heart attack in Texas than in New York. E. Guadagnoli et al., "Variation in the Use of Cardiac Procedures after Acute Myocardial Infarction," *New England Journal of Medicine* 333, no. 9 (1995): 573–578.
9. M. Chassin et al., "Does Inappropriate Use Explain Geographic Variation in the Use of Health Care Services? A Study of Three Procedures," *Journal of the American Medical Association* 258, no. 18 (1987): 2533–2537.
10. R. Brook, "Maintaining Hospital Quality: The Need for International Cooperation," *Journal of the American Medical Association* 270, no. 8 (1993): 985–987. That the problems are worldwide suggests that for-profit and not-for-profit status is not a major cause of the problem.
11. M. Chassin et al., "Benefits and Hazards of Reporting Medical Information Publicly," *New England Journal of Medicine* 334, no. 8 (1996): 394–398; E. Schneider and A. Epstein, "Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care," *New England Journal of Medicine* 335, no. 4 (1996): 251–256; and E. Schneider and A. Epstein, "Use of Public Performance Reports: A Survey of Patients Undergoing Cardiac Surgery," *Journal of the American Medical Association* 279, no. 20 (1998): 1638–1642. Lack of response may have stemmed from patients' being mostly unaware of the information.
12. A. Localio et al., "Relation between Malpractice Claims and Adverse Events Due to Negligence: Results of the Harvard Medical Practice Study, III," *New England Journal of Medicine* 325, no. 8 (1991): 245–251.
13. This group had more at stake because damages are related to lost earnings, which would generally be larger for those under age seventy. In these cases, it is unlikely that transaction costs (for example, attorney and expert witness fees) were large enough to make a claim uneconomic.
14. M. Marquis, "Cost Sharing and Provider Choice," *Journal of Health Economics* 4, no. 2 (1985): 137–157.
15. K. Arrow, "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review* 53, no. 5 (1963): 941–973. Interestingly, Arrow pointed out a link between consumers' ignorance and physicians' desire to deliver high-quality care. Because of the need to trust the physician as agent, the physician's "safest course" is to give "the socially prescribed 'best' treatment of the day. Compromise in quality, even for the purpose of saving the patient money, is to risk an imputation of failure to live up to the social bond" (p. 966).
16. TECH Research Network, "Technological Change around the World: Evidence from Heart Attack Care," *Health Affairs* (May/June 2001): 25–42.
17. U.S. Food and Drug Administration, Center for Drug Evaluation and Research, "RDMT Reports," [www.fda.gov/cder/rdmt/CY00NDAAP.htm](http://www.fda.gov/cder/rdmt/CY00NDAAP.htm) (27 August 2001). New molecular entities are active substances not previously approved in the United States.
18. Minutes of the National Cancer Policy Board of the Institute of Medicine, 11–12 October 2001.
19. This is undoubtedly the principal explanation of the much greater increase in the rate of various procedures among those over age eighty-five than among those ages sixty-five to sixty-nine between 1987 and 1995. See V. Fuchs, "Health Care for the Elderly: How Much? Who Will Pay for It?" *Health Affairs* (Jan/Feb 1999): 11–21.
20. National Institutes of Health, "The NIH Almanac—Appropriations," [www.nih.gov/about/almanac/appropriations/part2.htm](http://www.nih.gov/about/almanac/appropriations/part2.htm) (10 October 2001). Figures for 1990 were put into year 2000 dollars using the U.S. gross domestic product (GDP) deflator. Spending on trials went up faster than the total NIH budget, which increased from \$9.6 billion in 1990 to \$17.8 billion in 2000.
21. D. Davis et al., "Changing Physician Performance," *Journal of the American Medical Association* 274, no. 9 (1995): 700–705. See also P. Cantillon and R. Jones, "Does Continuing Medical Education in General Practice

- Make a Difference?" *British Medical Journal* 318, no. 7193 (1999): 1276–1279.
22. B. McNeil, "Shattuck Lecture—Hidden Barriers to Improvement in the Quality of Care," *New England Journal of Medicine* 345, no. 22 (2001): 1612–1620.
  23. E. Guadagnoli et al., "Appropriateness of Coronary Angiography after Acute Myocardial Infarction among Medicare Beneficiaries: Managed Care versus Fee-for-Service," *New England Journal of Medicine* 343, no. 20 (2000): 1640–1646. Among the 9 percent of beneficiaries in managed care (as of 1995), the rate for Class I indications was 37 percent, below the rate for traditional Medicare, whereas the rate for Class III indications was the same.
  24. E. Guadagnoli et al., "Impact of Underuse, Overuse, and Discretionary Use on Geographic Variation in the Use of Coronary Angiography after Acute Myocardial Infarction," *Medical Care* 39, no. 5 (2001): 446–458. More specifically, the raw difference in rates between the fifth and ninety-fifth percentiles across the ninety-five regions was 33 percent. If there had been no variation across the regions in the "appropriate but not necessary" or in the "uncertain" categories, the difference in rates would have been 11 percent.
  25. It may be difficult to find enough patients with various comorbidity combinations or for diseases where multiple drug therapy is optimal, as in the case of AIDS, but proper combinations and dosing are not known.
  26. D. Jorgenson, "Information Technology and the U.S. Economy," *American Economic Review* 91, no. 1 (2001): 1–32. Moore's Law holds that chip capacity will double every eighteen to twenty-four months. This prediction, made in 1965, has held true for the subsequent thirty-five years.
  27. The GDP deflator rose 4.6 percent per year over that period.
  28. The measurement of medical care prices adjusted for quality is in its infancy. One example, which finds price declines for the treatment of acute myocardial infarction (AMI), although nowhere near the magnitude of the chip or computer industries, is D. Cutler et al., "Are Medical Prices Declining?" *Quarterly Journal of Economics* 113, no. 4 (1998): 991–1024.
  29. Baily and Solow, "International Productivity Comparisons."
  30. The issue is not moral hazard because of subsidized consumers. Although an insurance policy that subsidized my computer purchases would lead me to order a new machine each time capabilities improved, those computers should still be produced at least cost. In other words, moral hazard cannot explain a "quality chasm."
  31. See J.P. Newhouse, *Pricing the Priceless: A Health Care Conundrum* (Cambridge, Mass.: MIT Press, 2002).
  32. To be more exact, price should equal marginal cost at the desired quantity of services, assuming that paying marginal cost will cover fixed costs. Although this assumption does not necessarily hold in health care, to pursue matters would add length without changing the main conclusions. For further discussion, see Newhouse, *Pricing the Priceless*.
  33. P. Held and M. Pauly, "Competition and Efficiency in the End Stage Renal Disease Program," *Journal of Health Economics* 2, no. 2 (1983): 95–118.
  34. T. Hofer et al., "The Unreliability of Individual Physician 'Report Cards' for Assessing the Cost and Quality of Care of a Chronic Disease," *Journal of the American Medical Association* 281, no. 24 (1999): 2098–2105.
  35. R. Miller and H. Luft, "Managed Care Plan Performance since 1980," *Journal of the American Medical Association* 271, no. 19 (1994): 1512–1519, and R. Miller and H. Luft, "Does Managed Care Lead to Better or Worse Quality of Care?" *Health Affairs* (Sep/Oct 1997): 7–25.
  36. Group- and staff-model HMOs may have a small market share because consumers fear that the very incentives that reward them for minimizing cost for a given quality also reward them for underservice. From a societal point of view, such arrangements also offer the greatest incentives for adverse selection.
  37. M. Rosenthal and J. Newhouse, "Managed Care and Efficient Rationing," *Journal of Health Care Finance* (forthcoming).
  38. U.E. Reinhardt, "Health Manpower Forecasting: The Case of Physician Supply," in *Health Services Research*, ed. E. Ginzberg (Cambridge, Mass.: Harvard University Press, 1991).