Doyle, Joseph

Article

Measuring returns to healthcare

NBER Reporter Online

Provided in Cooperation with:
National Bureau of Economic Research (NBER), Cambridge, Mass.

Suggested Citation: Doyle, Joseph (2010) : Measuring returns to healthcare, NBER Reporter Online, National Bureau of Economic Research (NBER), Cambridge, MA, Iss. 1, pp. 10-13

This Version is available at:
https://hdl.handle.net/10419/61947

While the average credit quality of the loans being securitized is around B, more than 70 percent of the dollar amount of CDOs was initially rated as AAA. For this mismatch to be appropriate, it would need to be the case that a pool of assets with an average credit quality of B be able to withstand enough losses such that 70 percent of its liabilities will still remain default-risk free.

We also document a large degree of uniformity in CDO structures. The CDOs that we study have very similar liability structures and very similar collateral pools. There is little variation in the quality of the underlying collateral across different issuers; while we study around 4,000 tranches they all seem to conform to the same CDO model. What caused the uniformity in CDO structures?

One potential answer is that CDO issuers just follow market convention: if some CDO structures have been perceived as desirable, then other issuers will follow the same convention. However, while this would explain the uniformity in deal structures (that is, the amount allocated to each category), what explains the uniformity of the underlying collateral? An alternative explanation is that the issuers had access to the rating model of the credit rating agencies. According to this explanation, the rating agencies provided issuers with their model, and issuers structured their CDOs accordingly.

Anecdotal evidence suggests that the rating agencies models indeed were known to CDO issuers and were provided to them directly by the rating agencies. For example, the CDO Evaluator Manual — an optimization tool used by S&P — enabled issuers to achieve the highest possible rating at the lowest possible cost. The model, for example, would indicate to issuers when they had "excess collateral" and would advise issuers on: "the percentage of assets notional needs to be eliminated (added) in order for the transaction to provide just enough support at a given rating level."

Thus, the rating agencies may have served not just as monitors and evaluators of existing structures, but rather as architects and creators of new securities. Providing such models to issuers potentially led to the creation of CDOs with the minimum possible collateral needed to obtain an AAA credit rating. The uniformity across CDOs and the low credit ratings of the underlying collateral suggest that most issuers were using the model to target the highest possible credit rating at the lowest cost. If there were mistakes embedded in these credit rating black boxes — those were probably compounded over the trillions of dollars that were deliberately structured by CDO issuers using this model.

[1] *E. Benmelech and J. Dlugosz, "The Credit Rating Crisis," NBER Working Paper No. 15045, June 2009, and* NBER Macro Annual, 2009, *forthcoming.*

[2] *P. Bolton, X. Freixas, and J. Shapiro, "The Credit Ratings Game," NBER Working Paper No. 14712, February 2009; E. Damiano, H. Li, and W. Suen, "Credible Ratings,"* Theoretical Economics, *3 (September 2008), pp. 325–65; F. Sangiorgi, J. Sokobin, and C. Spatt, "Credit Ratings Shopping, Selection and the Equilibrium Structure of Ratings," NBER Working Paper No. 14712, November 2008.*

[3] *E. Benmelech and J. Dlugosz, "The Alchemy of CDO Credit Ratings," NBER Working Paper No. 14878, April 2009, and* Journal of Monetary Economics, *56(5) (July 2009), pp. 617–34.*

Measuring Returns to Healthcare

Joseph Doyle*

Healthcare spending in the United States comprises 16 percent of GDP — nearly 80 percent more than in the median OECD country and 45 percent above that of the second-highest spending nation, France. Across countries, and across markets within the United States, the vast disparities in spending are not associated with better measures of health-outcome.[1] However, evidence from time series and panel data suggest that higher healthcare spending has generated benefits that, when converted to dollar magnitudes in various ways, appear to exceed their costs.[2] Of course, the type of variation in treatment intensity differs across these two types of comparisons, but the question remains: are the returns to healthcare large or small?

Estimating such returns can be confounded because medical providers attempt to provide each patient with a particular level of care. With heterogeneous returns, greater care is likely provided to those with the highest returns. This would tend to bias results toward finding beneficial effects of treatment. At the same time, patients with the highest returns may be those in relatively poor health. Indeed, hospitalized patients who receive more care are much more likely to die in the hospital, even after controlling for a host of observable characteristics: more care is provided to patients in worse health. With the raw correlation between treatment and health seemingly negative,

Doyle is a Faculty Research Fellow in the NBER's Program on Aging and the Alfred Henry and Jean Morrison Hayes Career Development Associate Professor of Applied Economics at MIT's Sloan School. His profile appears later in this issue.

estimating returns is an uphill battle.

In a series of research studies, my coauthors and I have explored natural experiments that can shed some light on the returns to healthcare. Most of these papers consider conditions where selection bias associated with admission into the hospital is less of an issue: childbirth — the most common reason for hospitalization in the United States — and emergency admissions. This research summary briefly describes this work and points to future work in the area.

Evidence from At-Risk Newborns

One project, joint with Douglas Almond, Amanda Kowalski, and Heidi Williams, uses the idea that diagnostic thresholds can offer the potential to estimate returns to healthcare.[3] If physicians provide greater levels of care to patients falling just above a diagnostic criterion, then researchers can compare treatment and health outcomes for patients just above and below the threshold. The nature of the variation allows us to measure marginal returns, which are crucial for the interpretation of whether *additional* care saves lives.

Our work focuses on at-risk newborns on either side of the "very low birthweight" threshold of 1500 grams (3lbs. 5oz.) The underlying health of newborns who weigh 1499 grams is similar to those weighing 1500 grams, yet the rules of thumb used by physicians and hospital protocols call for additional attention for newborns below the threshold. By comparing newborns on either side of the threshold, we are able to avoid some of the confounding factors that usually affect measurements of the returns to health care.

We show that newborns with birthweights just below 1500 grams have discontinuously higher hospital costs than slightly heavier newborns, on the order of $10,000 each. When we study data from the census of U.S. births over twenty years, we find that newborns with birth weights just below 1500 grams have a single percentage-point lower infant mortality rate than newborns with birth weights just above this cutoff, even though mortality risk tends to decrease with birthweight. This constitutes a relatively large reduction when compared to a mortality rate of 5.5 percent just above 1500 grams. We conclude that the additional medical attention afforded to very low birthweight newborns is highly cost effective at saving lives.

The same project shows that hospitals with the most state-of-the-art neonatal intensive care units (NICUs) are less likely to use the threshold, whereas changes in treatment and mortality *are* found at those hospitals with lower-level or no NICUs.

Evidence from "Uncomplicated" Births

In another study with Douglas Almond, we test whether a longer stay in the hospital after a birth affects the health of newborns and mothers.[4] We use insurance rules that provide coverage for one or two days in the hospital after birth, and these days are counted as "the number of midnights in care." That is, a newborn delivered at 12:05 a.m will have one more night of reimbursable care than an infant born a few minutes earlier. In a dataset of California births from 1991–2002—including nearly 100,000 births within 20 minutes of midnight—we find that the discontinuous change in insurance coverage leads to significantly longer stays for those born just after midnight than for those born before midnight. We find no differences in major health problems, summarized by hospital readmissions and mortality, for either the infants or the mothers. Together with a 1997 law that mandated coverage for a minimum of two days, these results suggest that increases in the length of stay from 1-2 days or from 2-3 days impose substantial costs without apparent health benefits.

In comparison to the findings on at-risk newborns described above, this study shows the results that apply to "uncomplicated" deliveries. These newborns are representative of the typical birth. While new parents may benefit from the additional night of supervision, we conclude that in this instance the insurance mandates result in moral hazard: greater use of hospital resources with little benefit in terms of major health problems. This is consistent with efforts by insurers to reduce stays to one night in care.

Evidence from Health Emergencies

As noted earlier, there is a large amount of regional variation in healthcare spending within the United States. The *Dartmouth Atlas of Healthcare* shows that some markets spend 60 percent more than others, yet survival from a heart attack is remarkably similar across these areas. Is it possible that individuals in high-spending areas are in worse health in ways that are difficult to control for in the comparisons?

While it is not possible to randomly assign patients to different healthcare systems, I have compared the outcomes of patients who are exposed to different healthcare systems not designed for them: patients who are far from home when a health emergency strikes.[5] Patients who experience these health shocks may find themselves in an area that spends a great deal on patients or in one that tends to spend less. For example, West Palm Beach and Fort Lauderdale are neighboring cities on Florida's east coast with similar lodging prices, yet Fort Lauderdale tends to spend 30 percent more on heart attack patients. The idea is that these types of cities are close demand substitutes in terms of destinations, and they attract "close substitutes" in terms of patients.

Contrary to the literature that focuses on local patients, analyzing visitors to Florida who have a serious heart-related emergency in a high-spending market results in a 20 percent lower mortality rate than for patients in low-spending areas. These estimates are robust across different types of patients, including patient-income levels, and within groups of similar destinations. In addition, the results suggest that intensive-care unit services drive cost differences, and they appear to be cost effective.

The results apply to emergency care, and specifically to a set of patients healthy enough to travel. To the extent that the results may apply more broadly, it appears that high-spending areas may not be as wasteful as the previous cross-section results suggest.

The Consequences of Being Uninsured

An earlier study also considers health shocks by comparing patients with and without health insurance following a severe automobile accident.[6] This is a sudden health emergency when the individuals have no choice but to visit the hospital. The analysis uses a dataset originally intended for highway-safety research linking hospital discharge records with police reports. I find that the uninsured receive 20 percent less care and have a significantly higher mortality rate.

Another innovation in this project is the comparison group available in the rich data source: individuals who have health insurance but do not have automobile insurance according to the police report. These patients are quite similar to those who do not have health insurance, and the results are similar when the analysis is restricted to these two groups. This is one of the few studies to examine the potential effects of health insurance directly on health outcomes.[7] The results again suggest that greater treatment intensity yields health benefits for trauma care.

Returns to Physician Quality

Some physicians provide much more care to patients than others, which is a major source of variation within and across cities. Patients are referred or choose their physicians, however, and it is not clear how much of the variance in care can be explained by differences among the patients themselves.

A project with Todd Wagner and Steven Ewer studies a setting where over 30,000 patients in a large, urban hospital were randomly assigned to physician teams.[8] Further, the teams are affiliated with one of two academic institutions: one institution is among the top medical schools in the United States, while the other institution is ranked lower in the quality distribution. Because of the randomization, patients treated by the two teams have identical observable characteristics. Further, both teams have access to a single set of facilities and ancillary staff, because care is located in the same hospital.

We show that across common conditions, the more-prestigious teams provide care that is 10–25 percent less costly than the less-prestigious ones. Health outcomes are not related to the physician-team assignment, and the estimates are precise: they (statistically) rule out better health outcomes associated with assignment to the more-prestigious team. Further investigating the source of the treatment differences, the results are consistent with the ability of physicians in the lower-ranked institution to substitute diagnostic tests and specialist consultations for the faster judgments of physicians from the top-ranked institution.

The comparison is among only two institutions, but the results suggest a number of implications. First, local-area variation in care can be substantial, even after controlling for patient characteristics. Second, inequality in access to high-quality physicians may lead to differences in the use of specialists and testing, but not to health disparities. Third, a relaxation of accreditation standards may not adversely affect the quality of care, but it may raise operating costs. Fourth, while previous studies have found that high-cost areas are associated with lower-quality care, a greater reliance on specialists, and little difference in health outcomes, and interpreted this as evidence of wasteful spending, these results suggest the possibility of an alternative interpretation. Areas with lower-quality providers may require greater treatment intensity and the use of specialists in order to achieve outcomes on par with areas with higher-quality providers. This appears to be a fruitful area for future research.

Summary

Measuring returns to healthcare can be confounded by the nature of the delivery: more care is provided to patients in worse health. My research has investigated instances when additional care is less likely to be related to underlying patient health and found large returns for at-risk newborns and patients receiving emergency care, but small returns for longer postpartum hospital stays among typical births. Future work should continue to consider additional types of patients and treatments, begin to consider chronic conditions, and investigate the interaction between physician quality and the cost of care.

---

[1] *See, for example, E. Fisher, D. Wennberg, T. Stukel, D. Gottlieb, F. Lucas, and E. Pinder, "Implications of Regional Variations in Medicare spending, Part 2: health outcomes and satisfaction with care," in* Annals of Internal Medicine, *138(4) (2003), pp. 288–98; and A.M. Garber and J. S. Skinner, "Is American Healthcare Uniquely Inefficient?"* Journal of Economic Perspectives, *22(4) (2008), pp. 27–50.*

[2] *D. Cutler, A. Rosen, and S. Vijan, "The Value of Medical Spending in the United States, 1960–2000,"* New England Journal of Medicine, *355 (2006), pp. 920–27; and K.M. Murphy and R. Topel, "The Economic Value of Medical Research" in* Measuring the Gains from Medical Research: An Economic Approach, *K.M. Murphy and R. Topel, eds. Chicago: University of Chicago Press, 2003.*

[3] *D. Almond, J. Doyle, A. Kowalski, and H. Williams, "Estimating Marginal Returns to Medical Care: Evidence from Care for At-Risk Newborns," NBER Working Paper No. 14522, December 2008, and* Quarterly Journal of Economics, *forthcoming.*

[4] *D. Almond and J. Doyle, "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays," NBER Working Paper No. 13877, March 2008.*

[5] *J. Doyle, "Returns to Local-Area Health Care Spending: Using Health*

Shocks to Patients Far from Home," NBER Working Paper No. 13301, August 2007.

[6] *J. Doyle, "Health Insurance, Treatment, and Outcomes: Using Auto Accidents as Health Shocks," NBER Working Paper No. 11099, February 2005, and* Review of Economics and Statistics, *87(2) (2005) pp. 256–70.*

[7] *Institute of Medicine,* Hidden Costs, Value Lost: Uninsurance in America, *Washington D.C: National Academies Press, 2003, p. 141.*

[8] *J. Doyle, S. Ewer, and T. Wagner, "Returns to Physician Human Capital: Analyzing Patients Randomized to Physician Teams," NBER Working Paper No. 14174, July 2008.*

New Dynamic Public Finance

Mikhail Golosov and Aleh Tsyvinski*

Many problems in public finance and macroeconomics, such as the taxation of capital or the provision of Social Security and disability insurance, are dynamic in nature. The individuals who pay taxes or claim benefits are long-lived. The tax and benefit policies in place in one period can affect their behavior in other periods. For example, increasing retirement benefits may affect individuals' behavior and savings in earlier years.

The New Dynamic Public Finance literature extends the traditional literature on optimal income tax and optimal program design, much of which focused on settings in which individuals made decisions in a single period, to focus on such dynamic settings.[1] While the same efficiency-equity tradeoffs that apply in single-period settings also arise in dynamic settings, there are additional tradeoffs between providing insurance and preserving incentives. When individuals live for many periods, they may experience both favorable and unfavorable "shocks" as they age: unexpected increases in their wages, or the early onset of a disability, for example. Public policy can provide insurance against adverse shocks, but it may do so at some cost in incentives. Much of the research in New Dynamic Public Finance is directed at understanding how one can design social insurance or redistribution systems that achieve distributional objectives while ensuring necessary incentives to provide effort or work throughout individuals' lives.

When designing policy in dynamic settings, it is important to take account of the random shocks that confront individuals over time. These may be shocks to earnings capacity, or health status, or financial market returns. In each case, individual taxpayers or program beneficiaries are likely to have more information on their circumstances than the government does. The government cannot easily observe health status or hourly wage rates, and it cannot condition its tax or social insurance rules on them. The policy challenge is to preserve incentives for individual work and saving while still raising the necessary funds for redistribution or government revenue.

Consider a simple example of a dynamic social insurance problem: an able young worker may become disabled later in life. It may be possible to claim to be disabled even if one is able to work. For example, one can pretend to be suffering from back pain which is very difficult to verify. The fundamental challenge in designing a disability insurance system is to provide adequate transfers to truly disabled workers while discouraging fake disability applications. What a worker believes about his future decisions regarding whether to claim disability will affect his labor supply and saving choices while young. If he believes that he is very unlikely to apply for and receive disability benefits in the future, for example, then he is likely to save more in his younger years. Similarly, an individual's past saving choices may affect his willingness to fake disability and to claim disability benefits at an older age.

We highlight two sets of findings that have emerged from our own research, and that of other scholars, in the area of New Dynamic Public Finance. First, the tax treatment of saving is a key policy instrument for affecting dynamic incentives, and policymakers may need to consider its impact on a variety of labor market incentives. Second, the availability of private insurance against various risks may have an important effect on the way that government tax and transfer programs influence household behavior.

Tax Policy toward Saving Can Affect Dynamic Incentives

In our work with Narayana Kocherlakota, we develop an important insight about policy design in dynamic settings. When agents receive random shocks to their earnings capacity, one feature of government policies that achieves a Pareto-efficient allocation— one in which no one could be made better off with-

**Golosov and Tsyvinski are Research Associates in the NBER's Program on Public Economics and Professors of Economics at Yale University. Their profiles appear later in this issue.*