Journal of Health Economics 29 (2010) 866–882

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Returns to physician human capital: Evidence from patients randomized to physician teams[☆]

Joseph J. Doyle Jr.[a,*], Steven M. Ewer[b], Todd H. Wagner[c]

[a] MIT & NBER, United States
[b] University of Wisconsin-Madison, United States
[c] VA Palo Alto and Stanford, United States

ARTICLE INFO

Article history:
Received 3 February 2010
Received in revised form 12 August 2010
Accepted 16 August 2010
Available online 24 August 2010

JEL classification:
I12
J24

Keywords:
Returns to healthcare
Human capital
Medical school quality
Natural experiment

ABSTRACT

Physicians play a major role in determining the cost and quality of healthcare, yet estimates of these effects can be confounded by patient sorting. This paper considers a natural experiment where nearly 30,000 patients were randomly assigned to clinical teams from one of two academic institutions. One institution is among the top medical schools in the U.S., while the other institution is ranked lower in the distribution. Patients treated by the two programs have similar observable characteristics and have access to a single set of facilities and ancillary staff. Those treated by physicians from the higher ranked institution have 10–25% less expensive stays than patients assigned to the lower ranked institution. Health outcomes are not related to the physician team assignment. Cost differences are most pronounced for serious conditions, and they largely stem from diagnostic-testing rates: the lower ranked program tends to order more tests and takes longer to order them.

1. Introduction

A major question in healthcare is the underlying source of geographic variation in spending: high-spending areas in the U.S. incur costs that are 50% higher than low-spending ones (Fisher et al., 2003). These differences are often ascribed to divergent preferences and training among physicians (Phelps and Mooney, 1993; Eisenberg, 2002; Sirovich et al., 2008). Related evidence suggests that high-spending areas are associated with a greater number of specialists and lower quality care (Baicker and Chandra, 2004; Wennberg et al., 2009). There are also equity concerns that health disparities may result from differences in access to high-quality care (Institute of Medicine, 2002; Chandra and Skinner, 2003; Almond et al., in press).

Estimates of the effects of physicians on cost and quality of care can be confounded by omitted-variable concerns and selection issues. For example, high-risk patients may be referred to or self-select the "best" physicians (referral bias), and as a result the highest-quality physicians can have the highest mortality rates (Glance et al., 2008).[1] Indeed, public report cards that rank providers based on risk-adjusted mortality rates have been controversial due to concerns that patients differ in unobservable ways, and that the reports create incentives for providers to avoid high-risk cases (Marshall et al., 2000; Dranove et al., 2003). In addition, the environments where physicians operate differ, including differences in patient characteristics and complementary physical and human capital.

This paper considers a unique natural experiment in a large, urban Department of Veterans Affairs (VA) hospital, where nearly 30,000 patients (and over 70,000 admissions) were randomly assigned to teams comprised of clinicians from one of two academic

☆ The authors would like to thank Ann Bartel, Amitabh Chandra, Christine Durrance, Sherry Glied, Michael Greenstone, Jonathan Gruber, Lawerence Katz, Steven Levitt, Lars Lefgren, Thomas McGuire, David Meltzer, Abigail Ochberg, Ciaran Phibbs, Joseph Price, Roberto Rigobon, Jonathan Skinner, Douglas Staiger, Thomas Stoker, Tavneet Suri, and Jack Wennberg, as well as two anonymous referees and seminar participants at American Society of Health Economists Annual Meetings, Brigham Young, Carnegie Mellon, Harvard, MIT, NBER Health Care Program, San Francisco VA, Stanford, and Syracuse University, for helpful comments and insights. We also greatly benefited from the efficient programming of Andrew Siroka. All remaining errors are our own.

* Corresponding author. Tel.: +1 6174523761.
E-mail address: jjdoyle@mit.edu (J.J. Doyle Jr.).

[1] This non-random assignment of patients also plagues comparisons across hospitals. Geweke et al. (2003) find that patients with the worst unobservable severity go to high quality hospitals.

0167-6296/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2010.08.004

institutions. As described in more detail later, most VA hospitals are affiliated with one or more academic medical school. In this paper, we analyze data from a VA that has two academic affiliates. One set of physicians is affiliated with an academic institution that is among the top medical schools in the U.S.; the other set is linked with an institution that is ranked lower in the quality distribution.[2] Patient characteristics are similar across the two academic institutions due to the randomization. Meanwhile, the teams have access to the same facilities, the same nursing staff, and the same specialists for consultations. By comparing patient outcomes across these two groups, this paper aims to estimate effects of physicians on costs and health outcomes, i.e. returns to physician human capital.[3]

We find that patients assigned to physicians affiliated with the higher ranked program have 10% lower costs compared to the lower ranked program, and up to 25% lower costs for more complicated conditions. The differences largely stem from diagnostic-testing rates: the lower ranked program tends to order more tests and takes longer to order them. Meanwhile, hospital readmissions and mortality are unrelated to the physician-team assignment.

A main caveat is that the results apply directly to one hospital and two residency training programs, albeit with thousands of physicians that comprise them. The "parent hospital" of the higher ranked institution is similar in treatment intensity to other top teaching hospitals, however. This suggests that practice patterns at the top-ranked institution are similar to other highly ranked institutions as well.

The paper is organized as follows: Section 2 describes the empirical framework and defines the main parameters of interest; Section 3 provides background information on the physician teams and patient assignment, as well as a review of the previous literature; Section 4 describes the data; Section 5 reports the results; and Section 6 concludes.

2. Empirical framework

Consider a health production function that relates mortality, M, to health care inputs and a patient-level severity measure, θ:

$$M = F(H, K; \theta) \quad (1)$$

where H represents human capital of the hospital staff, and K represents physical capital.

We focus on the effects of physician human capital, H, on patient outcomes, as well as differences in treatment intensity. In our empirical application, there are two teams that differ markedly in the screening of physicians that compose each team, including different residents and attending physicians. Let P be an indicator that the patient was assigned to physicians in the lower ranked program, T be a measure of treatment, and X represent observable characteristics of the patients. The main parameters of interest can then be written as:

$$E(T|P = 1, X) - E(T|P = 0, X) \quad (2a)$$

$$E(M|P = 1, X) - E(M|P = 0, X) \quad (3a)$$

This gives rise to empirical models of the form:

$$T_i = \alpha_0 + \alpha_1 P_i + \alpha_2 X_i + \varepsilon_i \quad (2b)$$

$$M_i = \beta_0 + \beta_1 P_i + \beta_2 X_i + \upsilon_i \quad (3b)$$

where ε and υ are error terms.

A common problem when estimating α_1 or β_1 is that patients are not randomly assigned to physicians. Rather, patients choose or are referred to physicians. A patient's primary physician, who knows more about the illness severity than can be captured in typical data sets, may refer the "toughest" cases to the "best" physicians. This tends to bias against finding survival improvements for physicians with higher levels of human capital.[4] Comparisons across hospitals have the additional confounding factors of differences in technology and support staff, which may have a large impact on patient survival independent of the physician characteristics (Unruh, 2003; Evans and Kim, 2006; Bartel et al., 2009).

The main innovation in this paper is the study of a large number of patients who were randomly assigned to physician teams within the same facility. This should satisfy the identification assumptions that the physician team is mean independent of the error terms: $E(P\varepsilon) = E(P\upsilon) = 0$.

In terms of the standard errors, as in other randomized trials the individual error terms are assumed to be independently and identically distributed. The estimates reported are robust to heteroskedasticity and clustered at the patient level to account for dependence across observations for the same patients treated over time (similar results are found when we restrict the analysis to each patient's initial episode, as described below). These errors are conservative compared to alternatives considered.[5]

3. Background

3.1. Previous literature

Much of the previous work on physician human capital finds that previous test scores, such as undergraduate grade point average or Medical College Admissions Test (MCAT) scores, are positively correlated with later test scores (Case and Swanson, 1993; Glaser et al., 1992; Hojat et al., 1997; Silver and Hodgson, 1997). It is less clear whether physicians with higher scores provide higher quality care. Ferguson et al. (2002) review the literature on predictors of medical school success, and note that little has been done on post-medical school performance. There is some evidence on outcome differences by board-certification status, but it is mixed.[6]

[2] In some ways the top-ranked program's physicians are "stars". Rosen (1981) discusses star physicians, where the potential to be a superstar is limited by the extent of the market—in this case the physician's time to see patients. This time constraint inhibits the scalability of the treatment provided by top physicians.

[3] Gross returns are considered here. The residents studied earn similar wages regardless of their academic institution affiliation, and detailed data linking wages to the quality of medical education do not exist.

[4] In the case of heterogeneous treatment effects, the patients are likely referred based on the expected gain of the assignment: a correlated random coefficient model that can inflate returns to physician human capital (Bjorklund and Moffitt, 1987).

[5] One caveat is that the observations may be correlated within teams that vary over time, although we do not observe team composition. We found that clustering at the month-year level—times when team composition is likely to change—resulted in similar, and often smaller, standard errors. Similarly, when the estimates were jointly estimated using a seemingly unrelated regression, estimated standard errors were again similar and often smaller. Last, we considered correlation within each of the two groups. The outcomes considered here, however, have an intra-class correlation of close to zero (e.g. our cost measures have an intra-class correlation of less than 0.005). As in other randomized trials, these intra-class correlation coefficients imply that correcting the standard errors by clustering at the group level is unnecessary in this context (Moulton, 1986; Angrist and Pischke, 2008).

[6] Certification has been found to be associated with reductions in mortality following heart attacks (Kelly and Hellinger, 1987; Norcini et al., 2000), while other work has found differences in the use of appropriate medications but little difference in mortality (Chen et al., 2006). Licensure examination scores have been found to be related to preventive care and more appropriate prescription medicines (Tamblyn et al., 1998, 2002).