

Size Matters: The Impact of Physician Practice Size on Productivity

Douglas E. Hough

Carey Business School, Johns Hopkins University

Kai Liu

Department of Economics, Zanvyl Krieger School of Arts and Sciences

Johns Hopkins University

David N. Gans

Innovation and Research, Medical Group Management Association

Keywords: physician practice; production function; marginal product

JEL Classification Codes: I11, D24

Size Matters: The Impact of Physician Practice Size on Productivity

Abstract

Physician practices in the US are largely small-scale, independently-run enterprises, despite their potential critical role in creating a more integrated health care system. Using an approach that builds on – and extends – prior research, we estimate physician practice production functions for different types of practices (multispecialty, single-specialty, and six subspecialties within single-specialty practices). We find that these practices have distinct production functions, and that size has different implications for each. In particular, we find that the median size physician practice is well below the size suggested by the estimated production function and marginal products of physicians, for almost all practice types. These results have interesting implications for health policy.

A contradiction exists at the core of the efforts to improve the US health care system. Almost all health care reform proposals have envisioned that physicians and their practices will be integral parts of a more efficient and effective health care delivery system (American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American Osteopathic Association, March 2007; Christensen et al., 2009; Daschle, 2008; Robinson, 1999; Tollen, 2008). Yet, physicians seem reluctant to practice in the large organizations needed to effect the integration envisioned in these proposals.

We suspect that the answer to this contradiction may lie in the inherent economics of physician practices. That is, small practices may be more productive – and profitable – than large practices. On the other hand, large practices may be more productive, but physicians may be willing to sacrifice profits in order to achieve other goals more readily attainable in small practices (such as professional autonomy and control). To test these ideas, in this paper we begin the search for the “optimal” size of physician practices, by examining the structure and performance of physician practices. We will demonstrate that the production functions of physician practices (and thus the marginal products of physicians) differ by practice type and specialty. As a result, multiple optimal sizes of physician practices may exist.

In Section 1 we describe the environment of physician practices in the US and provide a brief history of recent attempts to integrate physicians into larger health care organizations. Section 2 reviews the major research to date on physician practice size and places our analytical framework in the context of this research. Section 3 describes our data and estimation strategy. In Section 4 we present descriptive statistics and the empirical estimates from our model. In Section 5 we discuss our findings and conclusions, as well as implications for health care reform.

1. The Context

Physician services currently constitute 21% of national health expenditures in the United States (Hartman et al., 2009), and many have asserted that – through the power of their referrals, hospital admissions, orders, and prescriptions – physicians have effective control over the bulk of health care spending (Sirovich et al., 2008). Paradoxically, physicians wield this influence despite the fact that, in most instances, they practice largely in small-scale, independently-run enterprises. The demise of solo practice has been heralded for decades, but nevertheless a large percentage of physicians remain solo practitioners. According to the American Medical Association (Smart, 2004), only 10% of physicians were in group practice in 1965; by 1986 that percentage had increased to 31.9%. It appeared then that group practice was becoming the predominant mode of practice. However, this trend appeared to stall out after 1986, with only 30.2% of physicians in group practices in 2003. Data from the American Medical Association also demonstrate that group practices are still relatively small, with 69% of group practices in 1994-95 and 65% of group practices in 2003 having between 3 and 6 FTE physicians. On the other hand, using the Community Tracking Study, (Liebhaber and Grossman, 2007) found a statistically significant decline in the percentage of physicians in solo or 2-physician practices, from 40.7% in 1996-97 to 32.5% in 2004-05.¹ They found the decline to be most prominent for specialists (both surgical and medical) rather than for primary care physicians.

It has been argued that physicians need to agglomerate into larger and more integrated practices. Many consider multispecialty group practices such as Kaiser Permanente and the Geisinger Health System to be the ideal structure for the physician practice of the future, with a

¹ Unpublished data from the Medical Group Management Association indicate that the pre-1986 trend continued, with 67.1% of physicians in group practices in 2003. This large discrepancy suggests underreporting of group practices to the AMA's periodic survey, but we are unable to reconcile the different estimates by the two organizations.

balance of primary care and specialty physicians, supported by a full range of clinical and ancillary services and administrative infrastructure (Shih et al., 2008). Others argue that single-specialty groups can match the quality and efficiency of multispecialty groups (Casalino et al., 2004). One recent analysis has found that small practices will be disadvantaged in increasingly popular pay-for-performance programs because they lack sufficient patient caseloads to demonstrate statistically reliable measures of superior performance (Nyweide et al., 2009).

In the past two decades, at least two major attempts have been made to consolidate physicians into large groups. The first was the rapid rise (and fall) of for-profit physician practice management companies (PPMCs). The 1990s saw the launch of dozens of PPMCs dedicated to the acquisition and management of physician practices. The PPMC business model was based on the premise that these companies could exploit cost savings through economies of scale and centralization, expanded access to capital, and bargaining clout with managed care companies. Almost all of the PPMCs foundered within a few years as the purported benefits of their business models failed to materialize, and either ceased operations or transformed their business models (Burns, 1997; Reinhardt, 2000; Robinson, 1999).

The second major attempt at physician aggregation was the acquisition of physician practices by hospitals and health systems. With the growth of managed care in the 1990s, many hospitals and health systems were concerned that they might be excluded from payer contracts. As a result, they felt the need to migrate many of their medical staff (especially primary care physicians) from the informal relationships represented by traditional hospital privileges and credentialing to more formal relationships including practice acquisition (Robinson, 1999). Even after the restrictive managed care environment receded in the early 2000s, many hospitals continued to see wholly-owned primary care physician practices as key components of their

retail supply chain (Halley, 2007). Unfortunately, many hospitals found that they were sustaining operating losses on these practices in excess of \$80,000 per physician, and responded by divesting the practices (Healthcare Advisory Board, 1999).

That both of these efforts – one well-funded and predicated on expected profit opportunities, and the other formulated by insiders and created for supposed strategic advantages – have largely failed, combined with the limited organic growth of group practices, suggests that the economics of physician practices may be more complicated than expected. Perhaps the presumed economies of scale and scope of larger physician practices are illusory. That is, the small physician practice may in fact be an economically robust enterprise; if so, this sector will remain fragmented for the foreseeable future.

2. Analytical Framework

The optimal size of a physician practice follows from – but is not coincident with – the profitability of the practice. Profitability, of course, stems from two sources – the extent of demand (as measured by volume purchased and price paid) and the nature of the practice's production function. This paper focuses on the structure of the production function for physician services as a first step in determining practice size optimality.

The structure of physician practices has generated episodic interest of economists and health service researchers over the past three decades. The research of relevance to the current study has focused on three areas: optimality as determined by survivor analysis; practices as organizations; and production function estimation. Frech and Ginsburg (1974) and Marder and Zuckerman (1985) applied survivor analysis to physician practices, under the hypothesis that the least (most) efficient size practices will fold (thrive); in general, they found that only large

multispecialty group practices may be of optimal size (although solo practices may still be efficient, at least in certain parts of the U.S.).

Other researchers have analyzed the physician practice as a firm. Madison and Konrad (1988) analyzed the growth and evolving organizational structure of large medical group practices. Using a typology based on “organizational tradition” and “market-response strategy,” they identified the expected internal conflicts within large physician practices as the practices respond to changing market demands. Pauly (1996) argued that the primary competitive advantage of multispecialty group practices is likely to be in coordinating the process of care (especially in a managed care environment), rather than any inherent economies of scale or scope. Getzen (1984) posited a “brand-name” theory of the firm applied to physician practices. Town et al. (2004) examined in considerable detail the role of incentives in physician groups, concluding that a single comprehensive theory of incentives within practices may not be possible because of the complexity of physician utility functions and practice production functions.

A considerable – and varied – body of research has focused on the estimation of physician production functions. Some have focused on input mix and interdependency, others on economies of scale or scope. The seminal work was conducted by Reinhardt (1972), who estimated a transcendental production function for physicians. He used the individual physician – rather than the practice – as the unit of observation, and measured output by total patient visits, physician office visits, and patient billings – rather than impact on the patient’s health (arguing that physician services are intermediate goods, rather than final goods).

Gaynor and Pauly (1990) used stochastic production frontier analysis to estimate both “traditional” and “behavioral” productions for primary care physicians, and confirmed the impact

of incentives on productivity. Pope and Burge (1996) used the Reinhardt production function to estimate significant increasing returns to scale for single-specialty practices. Using stochastic production frontier analysis, DeFelice and Bradford (1997) found that primary care physicians in both solo and group practices produced less than the optimal number of patient visits per week, and that the differences between the two practice modalities were statistically – but not economically – significant.

Two studies deserve particular attention because of their relevance to our work. The first, Thurston and Libby (2002), applied a generalized linear production function to the Reinhardt construct. This functional form allows for zero values for inputs, which is helpful in estimating production functions for physician services given the large number of small practices that do not employ all types of non-physician clinical providers. This function also facilitates estimation of the marginal productivity of inputs, as well as of Hicks elasticities of complementarity. Thurston and Libby used the physician as the unit of observation, and implicitly assumed that the physician services production function was homogeneous across practice types (i.e., multispecialty vs. single-specialty practices, and subspecialties within single-specialty practices). They used office visits and total patient visits as the measures of output. The model found that all of the estimated marginal productivities were positive, and that the three kinds of non-physician labor included (administrative and clerical, nurses, and technicians and aides) were complements with physician labor.

The second study of note – Rosenman and Friesner (2004) – applied data envelopment analysis to explore the efficiency of physician practices. Unlike most previous studies, they used the physician practice, rather than the individual physician within a practice, as the unit of observation. They divided their analysis into two overlapping samples: the “primary care

sample” (which included single-specialty primary care practices, multispecialty practices with primary care physicians only, and multispecialty practices with both primary and specialty care), and the “specialty care sample” (which included single-specialty specialty care practices, multispecialty practices with specialty care only, and multispecialty practices with both primary and specialty care). Using data from the 1998 Medical Group Management Association Annual Cost Survey, Rosenman and Friesner used three measures of output: “surgical and nonsurgical procedures and services done inside the practice’s facility;” “surgical and nonsurgical procedures and services done outside the practice’s facility;” and “total ancillary (diagnostic, radiological and laboratory) procedures and services performed by the practice.” The data envelopment analysis showed that single-specialty practices were more efficient than multispecialty practices, with the inefficiency due primarily to technical (rather than allocative) inefficiency. In addition, single-specialty practices had higher average scale efficiencies than multispecialty practices.

We build on – and extend – this line of research in several ways in the present paper. First, we adopt the generalized linear production function introduced into this literature by Thurston and Libby. As noted earlier, this model allows for zero inputs (such as advanced practice nurses, which many physician practices do not employ but who are becoming increasingly important in health care delivery), and allows for input complementarity and substitutability.² This function has the form:

$$Q_t = F(X_t) = F(X_{0t}, X_{1t}, \dots, X_{Kt}) = \sum_{i=0}^K \sum_{j=0}^K \alpha_{ij} \sqrt{X_{it}} \sqrt{X_{jt}}, \quad (1)$$

where $X_t = (X_{0t}, X_{1t}, \dots, X_{Kt})$ is the vector of inputs at time t (with $X_{0t} = 1$ everywhere) and $\alpha_{ij} = \alpha_{ji}$. As Thurston and Libby demonstrate, (1) can be rewritten as:

² The generalized linear production function does have the limitation of imposing constant returns to scale to the model.

$$Q_t = \sum_{i=0}^K \sum_{j=0}^K \beta_{ij} \sqrt{X_{it}} \sqrt{X_{jt}}. \quad (2)$$

In this case, $\alpha_{ij} = \beta_{ij}$ for $i = j$, and $\alpha_{ij} = \beta_{ij}/2$ otherwise. Equation (2) can be estimated using least squares regression methods.

Second, we follow Rosenman and Friesner by treating the physician practice – rather than the individual physician – as the unit of observation. Most of the previous research recognized that individual physicians worked with other clinicians, but assumed (at least implicitly) that production was driven by the individual physician. Research cited above note the critical role of organizational behavior in the production process; even though practices may be small and even atomistic, it is clear that the nature and structure of the production of medical care is determined by the firm.

Third, as has been recognized in the literature, we argue that physician practices are not homogeneous. For instance, single-specialty practices may have significantly different production functions from multispecialty physician practices. It can be argued that multispecialty practices seek scale in order to create an “ecosystem” of physicians, so that primary care physicians can refer their patients mostly to specialists within the practice, and in turn specialists in the practice can rely on their primary care colleagues for a full practice of referred patients. In addition, different kinds of single-specialty practices (primary care, medical subspecialty, surgical subspecialty) may themselves have different production functions. Primary care practices (i.e., family practice, internal medicine, and pediatrics) are usually office-based, in which they provide mostly cognitive services (also known as evaluation and management services); some (but not many) primary care practices offer revenue-generating

diagnostic and treatment procedures in their own facilities. Many medical specialties (such as cardiology and gastroenterology) and almost all surgical practices (e.g., orthopaedic surgery, neurosurgery, and urology) concentrate on major procedures; in fact, many practices own and operate their own diagnostic or surgical facilities, which provide potential for both practice efficiencies (e.g., optimizing the use of physician time) and profit (e.g., billing for technical as well as professional services). Other facility-based single-specialty practices (e.g., anesthesiology, pathology, and radiology) typically provide services in facilities owned and operated by others (e.g., hospitals) often on an exclusive contractual basis, and often have office space only for billing and collection functions.³ These substantial structural differences among physician practices suggest that both the “nature of the firm” (as elucidated by Coase (1937)) and the “theory of the business” (as proposed by Drucker (1994)) in physician practices are not uniform.

Finally, we present an alternative measure and/or interpretation of output. Many of the studies cited above use total physician visits or office visits as their measure of output. These have the disadvantage of omitting major services (e.g., surgical procedures for surgical practices and interventional procedures from medical specialists). Unfortunately, there exists no single measure of total output of physician practices, or a recognized index of practice services. However, we do have total revenue from a practice. Given the nature of the physician services market, we can assume that physician practices are price-takers. After all, the primary payer for most practices (Medicare) reimburses physicians uniformly through the Resource-Based Relative Value Scale (RBRVS) system, and many other payers use that payment methodology as well.

³ These descriptions are admittedly general and do not capture the diversity of each kind of physician practices. Readers who desire a more comprehensive description of physician practices are directed to Freidson (1988) and Wagner (2005).

We can also assume that practices of the same structure (i.e., all multispecialty groups or all single-specialty groups of a particular specialty) will have the same distribution of payer.⁴ Under these assumptions, we can use total revenue as a proxy for total output.

3. Data and Methods

3.1. Data

Like Rosenman and Friesner (2004), we use data from the annual Medical Group Management Association (MGMA) Cost Surveys. Since its founding in 1926, MGMA has collected and reported practice revenue and expense information on an ongoing basis. Since 1979, the annual MGMA Cost Survey has used a similar questionnaire and definitions to facilitate year-to-year comparisons. Questionnaires are distributed to medical practices that have an MGMA member, as well as to physician practices that register with MGMA to participate in the survey. The survey contains over 130 questions that summarize a practice's financial statements in a standardized manner to provide a comprehensive financial description of the organization's revenue and expenses. The survey response is similar on a year-to-year basis and, considering the complexity and length of the questionnaire, is considered reasonable. In 2008 10,535 questionnaires were distributed (by postal and e-mail), and 1,679 responses were received for a 15.9% overall response rate. The overlapping survey respondent population and stability in the survey questionnaire contribute to consistency in the data report. MGMA Cost Surveys are widely accepted by federal agencies and academic researchers as an accurate and comprehensive source of data on the financial performance of medical group practices (Pope et al., 2006;

⁴ For example, we can assume that all single-specialty pediatric groups will have the same payer distribution (mostly private insurance and Medicaid, with little or no Medicare), and that all single-specialty cardiology practices will have the same payer distribution (primarily, Medicare and private insurance).

Subcommittee on Health of the Committee on Energy and Commerce, 2002). These surveys are described in more detail at (Medical Group Management Association, 2007).

We use Total Gross Charges as our measure of production output. MGMA defines Total Gross Charges as the sum of gross fee-for-service charges (defined as “the full value, at the practice’s undiscounted rates, of all services provided to fee-for-service, discounted fee-for-service and non-capitated patients for all payers”⁵) and gross charges for patients covered by capitation contracts (defined as “the full value, at a practice’s undiscounted rates, of all covered services provided to patients covered by all capitation contracts, regardless of payer”⁶). Total Gross Charges is superior to collections or revenue from a theoretical perspective because charges are more indicative of practitioner performance whereas collections are influenced by the ability of the practice to collect from payers. Total Gross Charges is superior to Relative Value Units (which are used to generate the RBRVS system) from a practical perspective, because only about 1/3 of respondents to the MGMA Cost Survey report RVUs for their practice. Finally, because we use three years of data, we deflate Total Gross Charges (scaled in millions USD) by the Medical Care Services price index.

⁵ Gross fee-for-service charges includes: professional services provided by physicians, non-physician providers, and other physician extenders such as nurses and medical assistants; both the professional and technical components of laboratory, radiology, medical diagnostic and surgical procedures; contractual adjustments (such as Medicare charge restrictions, third-party payer contractual adjustments, charitable adjustments, and professional courtesy adjustments); drug charges; charges for supplies consumed during a patient encounter inside the practice’s facilities; facility fees; charges for fee-for-service services allowed under the terms of capitation contracts; charges for professional services provided on a case-rate reimbursement basis; and charges for purchased services for fee-for-service patients.

⁶ Gross charges for patients covered by capitation contracts includes: fee-for-service equivalent gross charges for all services covered under the terms of the practice’s capitation contracts, including: professional services provided by physicians, non-physician providers, and other physician extenders such as nurses and medical assistants; both the professional and technical components of laboratory, radiology, medical diagnostic and surgical procedures; drug charges; charges for supplies consumed during a patient encounter inside the practice’s facilities; and facility fees.

We use six measures of inputs: the full-time-equivalent (FTE) number of physicians involved in clinical care in the practice; the FTE number of mid-level providers⁷ involved in clinical care in the practice; the FTE number of clinical staff⁸ in the practice; the FTE number of ancillary staff⁹ in the practice; the FTE number of office staff¹⁰ in the practice; and annual capital expenditures¹¹ in the practice. In addition, we include a binary variable that identifies if physicians were the majority owners of the practice.

The combined 2005-2007 MGMA data set included 4,830 observations for the years 2004-2006. We eliminated 480 practices because of missing values of input and output variables, and 580 because of missing group identification numbers. Finally, we assumed that physician practices require non-zero FTE physicians and capital expenditures to create output, which eliminated 256 observations. This process left us with a total of 3,514 observations, of which 2,389 are unique physician practices.

3.2. Estimation Strategy

We estimate Equation (2) above as our empirical model. We are interested primarily in the coefficients for the independent variables measuring physician services, since one of our main purposes will be to evaluate the marginal productivity of physicians. We do not impose the

⁷ Mid-level providers include: non-physician providers who are specially trained and licensed providers who can provide medical care and billable services, such as Certified Registered Nurse Anesthetists, dietitians/nutritionists, midwives, nurse practitioners, occupational therapists, optometrists, physical therapists, physician assistants, psychologists, social workers, speech therapists, and surgeon's assistants.

⁸ Clinical staff include: registered nurses, licensed practical nurses, medical assistants, and nurse's aides.

⁹ Ancillary staff include: clinical laboratory, radiology and imaging, and other medical support services.

¹⁰ Office staff include: general administrative, patient accounting, general accounting, managed care administrative, information technology, housekeeping/maintenance/security, medical receptionists, medical secretaries-transcribers, medical records, and other administrative support.

¹¹ Capital expenditures are defined as annual costs for information technology (practice-wide data processing, computer, telephone and telecommunications services), building and occupancy (i.e., general operation of buildings and grounds), and furniture and equipment. Because we use three years of data, we deflate Capital expenditures (scaled in millions USD) by the U.S. City Consumer Price Index.

non-negativity of β_{ij} ex-ante; the β_{ij} are to be determined ex-post by the model. The time effects capture the macro elements affecting growth of gross charges (Q_i).

In a cross-sectional setting, the β_{ij} can be readily estimated by ordinary least squares (OLS) regression, as in Thurston and Libby (2002). However, for OLS estimates to be consistent, we need to make the assumption that the error term is uncorrelated with input variables. One potential source of bias violating this assumption may be from “managerial bias,” although part of this potential bias is controlled through the ownership type variable in our model.

The data give us a limited panel component to study medical practices’ production behavior¹². We estimate both the fixed effect and random effect model on the whole sample. The Breusch and Pagan Lagrangian multiplier test for random effects model yields a test statistic of 309.81, far exceeding the critical value ($\chi^2(1) = 3.84$), which implies that estimating a classical regression model of a production function with a single constant term is inappropriate. From a fixed effects estimation, the F statistic (p-value close to 0.00) for testing the joint significance of the firm effects is strongly in favor for a firm specific effect in data. The Hausman specification test shows that the hypothesis that the firm specific effects are uncorrelated with the other regressors cannot be rejected (p-value = 0.13), suggesting that the concern for “managerial bias”

¹² We abstract from issues related to selection bias. The MGMA Cost Survey relies on self-reporting from physicians practices and is not intended to follow a particular medical practice over time. We confirmed that the distribution of the observable variables, including numbers of physicians and total gross charges, are similar for those practices appearing once and those appearing in multiple years.

is not a serious problem in our model. We thus conclude that the random effects model is a better choice, which is also more efficient in using our highly unbalanced panel data¹³.

As noted above, we hypothesize that the production function is heterogeneous among different types of physician practices. We tested this hypothesis by estimating the production function with practice type dummy variables as well as interaction between the dummy variables and physician input variables.¹⁴ We then conducted a Wald test on the joint significance of the estimated coefficients of the interaction terms. The test statistics (distributed as $\chi^2(7)$) strongly support our hypothesis. For example, in testing the null hypothesis that the marginal product of physicians in single-practice specialties and multi-specialty practices is the same, we estimated the all-practice model included interactions involving a single-practice dummy variable. A test of the joint significance of the interaction terms revealed that the Wald statistic was 52.10, with p-value less than .01. We also tested this null hypothesis for the six subspecialties within single-specialty practice types; under almost all cases, we reject the null hypothesis at the .01 level.¹⁵ Therefore, we confirm that the marginal product of physicians is indeed different across practice types. As a result, we estimate the random effects model using generalized least squares regression, separately for all group practices, all multispecialty group practices, and all single-specialty practices, as well as six different types of single-specialty practices (family practice, pediatrics, cardiology, orthopaedic surgery, obstetrics/gynecology, and anesthesiology).

¹³ Fixed effect estimation exploits within-group variations, which is inefficient in our setting where a substantial number of practices appear only once over the three years. Estimation results from fixed effects model are available from the authors upon request.

¹⁴ The interaction variables with the practice type dummy (Practype) are: Practype*MD^{1/2}, Practype*MD, Practype*(MD^{1/2}*K^{1/2}), Practype*(MD^{1/2}*Off^{1/2}), Practype*(MD^{1/2}*Clin^{1/2}), Practype*(MD^{1/2}*Anc^{1/2}), and Practype*(MD^{1/2}*Mid^{1/2}). We interacted Practype with only the physician input variables, because our primary interest, the marginal product of physicians, does not depend on the other input variables.

¹⁵ We conducted a total of 15 joint significance tests on the coefficients of interactions. The complete results are available from authors upon request.

4. Results

4.1. Summary Statistics

We begin the empirical analysis by examining the characteristics of the 3,514 observations in our sample. Table 1 presents summary data on the size of these practices, as measured by the number of FTE physicians. For the entire sample, the mean size was 22.1 FTE physicians, with a median of 8.5; the smallest was 0.2 FTE physicians and the largest was 991.3. Data not presented in the table show that 10% of the sample had 2 or fewer FTE physicians per practice, which is notable because the source of the data (the Medical Group Management Association) is a trade association that represents group practices.

TABLE 1 GOES ABOUT HERE

Table 1 also illustrates the dramatic size differences between multispecialty and single-specialty group practices. The mean/median size of multispecialty practices in this sample was 52.1/28.0 FTE physicians, almost five times the mean/median size of single-specialty practices (11.1/6.5 FTE physicians). Note, however, that size distribution of these practice types overlaps to a considerable degree. The largest single-specialty practice (235.0 FTE physicians) exceeds the size of 75% of multispecialty practices, and the size of the 75th percentile of single-specialty practices in the sample exceeds the size of the 25th percentile of multispecialty practices. These differences suggest that the fundamental characteristics of multispecialty and single-specialty physician practices (such as business models and production functions) may be different as well; as a result, separate model estimation may be merited.

Table 1 reveals that size differences exist among single-specialty practices, as well. Primary care practices (family practice, pediatrics) are much smaller than other single-specialty

practices. As noted above, primary care practices are usually office-based, and the bulk of their services are office visits by patients, rather extensive procedures; as a result, scale may not be of economic or strategic advantage. Medical (cardiology) and surgical (orthopaedic surgery and obstetrics/gynecology) practices are larger; these practices are focused on major procedures, and often involve ownership of diagnostic or surgical facilities. The largest single-specialty practices (anesthesiology) are facility-based, often provide services on an exclusive contractual basis, and require office space only for billing and collection functions. These differences in size and practice characteristics among single-specialty physician practices suggest that it may be worthwhile to estimate separate production functions for each kind of practice.

Table 2 reveals even more differences among physician practices, in terms of the variables used in the production function model estimation. The mean physician practice among the 3,514 in the sample had \$9.4 million in total gross charges per year from its 22.1 physicians (for an average of \$424,300 per physician), and spent \$700,000 on capital expenditures. This practice employed 13.9 FTE ancillary staff, 27.2 clinical staff, 51.6 office staff, and 6.2 midlevel providers. 75% of the responding practices were owned by physicians.

TABLE 2 GOES ABOUT HERE

This mean practice masks the significant differences among practices. The differences between multispecialty and single-specialty practices are even more dramatic than simply size, as indicated in Table 1. The mean multispecialty practice generated over three times the total gross charges as the mean single-specialty practice, but with almost five times the number of FTE physicians; as a result, the mean multispecialty practice produced 29% less charges per FTE physician (\$367,000 versus \$515,000). Controlling for practice size, the mean multispecialty

practice employed more ancillary, clinical, and office staff and fewer midlevel providers than did the mean single-specialty practice. If one can assume that multispecialty and single-specialty practices have the same maximand and same level of efficiency, then it is likely that their production functions are different.

As with practice size, the subspecialties within single-specialty practices differ dramatically in terms of practice characteristics. Family practices had the lowest mean total gross charges and charges per FTE physicians, and the fewest midlevel providers; in addition, they were the least likely (by far) to be physician-owned (36%, compared to the single-specialty practice average of 80%). Cardiology practices had the highest mean total gross charges, the second-highest mean charges per FTE physician, and the highest mean level of capital expenditures; in addition, they employed the highest mean number of FTE ancillary, clinical, and office staff. Anesthesiology practices had the lowest capital expenditures, and employed the fewest FTE ancillary, clinical, and office staff; however, they employed the most midlevel providers (primarily, Certified Registered Nurse Anesthetists). Orthopaedic surgery practices generated the highest mean level of charges per FTE physician, and were the most likely (at 98%) to be owned by physicians. As with the other summary data, these differences support the conjecture that the structure and production functions of physician practices are likely to differ by specialty.

4.2 Empirical Estimates

Given these summary results, we turn to the results of the generalized least squares model estimation.¹⁶ Table 3 presents the results of the model estimation for all practices, all multispecialty practices, and all single-specialty practices in the sample. The results for each

¹⁶ The model was estimated using Stata/SE 9.2.

model are very different. We examine the direct and interaction effects separately. For the direct effects, the physician input variables (number of FTE physicians [MD] and the square root of FTE physicians [$MD^{1/2}$]) are statistically significant (at the .05 level) in both terms for only single-specialty practices; only MD is statistically significant for the all practices model; and neither variable is significant for the multispecialty practices model. For the capital input variable, the direct coefficients are statistically significant for only the all practices model. One of the office staff input variables ($Off^{1/2}$) is statistically significant with the multispecialty and single-specialty models, but not with the all practices model. The clinical staff input variables are statistically significant only with the single-specialty model. None of the ancillary staff input variables is statistically significant in any of the three models. Finally, of the mid-level provider variables, one (Mid) is statistically significant in all three models, and the other ($Mid^{1/2}$) is statistically significant in the single-specialty practice model.

TABLE 3 GOES ABOUT HERE

Examination of the interaction effects between the physician input variable ($MD^{1/2}$) and the other input variables indicates that only the interactions between the physician variable and the capital and clinical staff variables are statistically significant for the all practices model. The interactions between the physician variable and the ancillary staff and mid-level provider variables are statistically significant for the multispecialty practice model, but not for the single-specialty practice model. Conversely, the interactions between the physician variable and the office and clinical staff variables are statistically significant for the single-specialty practice model, but not for the multispecialty practice model.

Finally, the physician ownership variable is statistically significant for all three models. If we evaluate the numerical significance of this variable *ceteris paribus* (at the means of all the other explanatory variables), we find that physician ownership increases productivity (as measured by total gross charges) by a total of \$2,078,000 per year (or \$94,000 per physician) for all physician practices, a total of \$1,370,000 per year (or \$26,000 per physician) for all multispecialty physician practices, and a total of \$710,000 per year (or \$64,000 per physician) for all single-specialty practices. These increases represent a productivity premium of 22%, 7%, and 12% of total gross charges, respectively.¹⁷ The per FTE physician productivity gains from physician ownership are estimated to be \$94,000 for all practices, \$26,000 for multispecialty physician practices, and \$64,000 for single-specialty practices.

The results presented in Table 3 confirm that physician practices are not homogeneous in terms of production functions, and that multispecialty and single-specialty practices should be analyzed separately. Table 4 decomposes single-specialty practices into six subspecialties: family practice, pediatrics, cardiology, orthopaedic surgery, obstetrics/gynecology, and anesthesiology.¹⁸ In general, the results confirm that the production functions differ among subspecialties. As we did with the previous models, we examine the direct and interaction effects separately. For the direct effects, at least one of the physician input variables is statistically significant (again, at the .05 level) in four of the subspecialty models (family practice, pediatrics, orthopaedic surgery, and anesthesiology). The capital variable is significant in three models (family practice, orthopaedics, and anesthesiology); the office staff variable is significant in two models (pediatrics and anesthesiology); the clinical staff variable is significant

¹⁷ The total productivity gains from physician ownership in multispecialty practices are larger (at a .10 level of statistical significance) than that of single-specialty practices.

¹⁸ As Table 1 shows, these six subspecialties represent 1,526 (59.7%) of all single-specialty practices in the sample. The sample sizes of the remaining subspecialties were too small to permit estimation of their production functions.

in two models (pediatrics and obstetrics/gynecology); the ancillary staff variable is significant in three models (pediatrics, orthopaedic surgery, and anesthesiology); and the midlevel provider variable is not significant in any of the six models.

TABLE 4 GOES ABOUT HERE

Few of the interaction effects between the physician input variable and the other input variables are statistically significant (even at the .10 level) in the six subspecialty practice models: with capital in the cardiology and anesthesiology models; with office staff in the orthopaedic surgery and obstetrics/gynecology models; and with clinical staff in the orthopaedic surgery model. In none of the models were the interactions between the physician input variable and ancillary staff or mid-level providers statistically significant.

Finally, none of the physician ownership coefficients was statistically significant in the subspecialty single-specialty models. These results are not surprising for subspecialties such as cardiology and orthopaedic surgery, for which only 2-3% in the sample are owned by non-physicians. Nevertheless, these results are unexpected for family practice and pediatrics, given the high degree of significance in the all practice and all single-specialty practice models.

Although the generalized linear production function has a certain elegance, the multitude of coefficients makes it difficult to interpret the impact on any one input on production. (In the model tested here, each input appears in the estimated model seven times.) Since physician labor is the driver of most physician practices and since most discussions of firm size related to physician practices are couched in terms of number of physicians, we chose to analyze the models shown in Tables 3 and 4 in terms of how production is affected as the number of physicians varies. In particular, we calculated the marginal product of physicians for each of the

nine models; that is, we measured the slope of the estimated production function relative to the number of FTE physicians for each model (entering the value in the two direct and five interaction terms), holding the other inputs (capital and staffing) constant at their sample means.

The results of our calculations of the marginal product of physicians are presented in Figure 1. As is apparent visually, the marginal product of physicians varies dramatically across the different types of practices. We consider each panel in turn. In the all practices model, the marginal product of physicians is negative until the practice reaches 8 FTE physicians; that is, total product actually declines as the practice grows from 1 to 8 physicians. As Panel 1 also shows, the median size of all practices in the sample is 8 FTE physicians. Panels 2 and 3 reveal that this pattern is reversed when multispecialty and single-specialty practices are estimated separately. In the multispecialty practice model (Panel 2), marginal product is monotonically decreasing but does not approach zero. The overall single-specialty model (Panel 3) shows a very similar marginal product curve, even though the two kinds of practices have significantly different average products (see Table 2) and median sizes (28 FTE physicians for multispecialty practices and 6.5 FTE physicians for single-specialty practices).

FIGURE 1 GOES ABOUT HERE

Panels 4 through 9 show the marginal products for the 6 subspecialties that were modeled. Panel 4 indicates that the marginal product of physicians in a family practice decreases monotonically, and reaches zero at 24 FTE physicians. The median family practice size is 6.5 FTE physicians. Pediatrics, on the other hand, shows the opposite pattern (Panel 5). The slope of the marginal product is monotonically positive, and marginal product itself is negative until the practice reaches 3 FTE physicians (the median size of a pediatrics practice is 6 physicians).

Cardiology (Panel 6), the only medical subspecialty in the sample, follows the pattern of the all single-specialty practices model (Panel 3): marginal product decreases monotonically but does not approach zero. Orthopaedic surgery (Panel 7) follows pattern of the all practices model (Panel 1): The marginal product of physicians is negative until the practice reaches 9 FTE physicians (the median orthopaedic surgery practice has 8 FTE physicians); that is, total product actually declines as the practice grows from 1 to 9 physicians but grows monotonically after reaching that size. Obstetrics/gynecology (Panel 8), which is a specialty with elements of both primary care and surgery, has a marginal product curve that resembles family practice (Panel 4). Marginal product decreases monotonically, and reaches zero at 36 FTE physicians (the median obstetrics/gynecology practice has 5 FTE physicians). Finally, as Panel 9 indicates, anesthesiology follows the pattern of the orthopaedic surgery model: monotonically positive slope, with negative marginal product until the practice reaches 6 FTE physicians (the median anesthesiology practice has 20 FTE physicians).

5. Discussion and Conclusions

The results presented above demonstrate that the structure and production functions of physician practices differ substantially by type of practice. Multispecialty practices are different than single-specialty practices, and the subspecialties within single-specialty practices have different production functions. These findings suggest that each kind of physician practice has a different “theory of the business” (Drucker, 1994), and that size has different implications for different practices. For instance, multispecialty practices gain by size through the creation of an internal referral network among their physicians, and through the capture of ancillary services and control over equipment and facilities. These attributes are less of a factor overall for single-specialty practices.

Within single-specialty practices, we see a range of value of size. Family practices, for the most part, have few opportunities to generate production (and revenue) beyond direct patient care and cognitive services. As a result, family physicians have little to gain from size in terms of care or revenue. The organizational complexities of larger size (and the attendant perception of loss of control by each physician as the practice grows) will rapidly overcome any limited production advantages. These findings help explain why earlier attempts to aggregate and integrate primary care practices (by national physician practice management companies, as well as local hospitals and health systems) failed.

For other single-specialty practices, size can be an advantage. In orthopaedics, size can enable subspecialization within the practice (such as hand surgery, spine surgery, and sports medicine), as well as the acquisition of ancillary services (e.g., advanced imaging [CT, MRI], physical therapy and rehabilitation) and ambulatory surgery facilities. These services can improve the productivity of individual practitioners as well as the entire practice, and can expand sources of revenue. In a different way, single-specialty anesthesiology practices can gain from size. There is some subspecialization within the field (e.g., pediatrics, critical care medicine) and some opportunities for adding ancillary services (e.g., pain management). However, the main advantages to size for anesthesiology practices are the combination of low capital and operating costs (because the physicians perform most of their services in facilities that they do not operate) and the prospect of exclusive contracts for services with hospitals and other facilities (which value the ability to secure regularly scheduled physician services through contract rather than credentials).

It needs to be noted, however, that not all of the findings in this research have been as expected. The empirical results of the model for single-specialty pediatric practices indicate that

these practices should be like anesthesiology or orthopaedic surgery practices. However, pediatrics in reality is more similar to family practice in structure and production possibilities: Production and revenue are driven primarily by frequent office visits with established patients and consultations with parents. Similarly, we expected single-specialty cardiology practices to exhibit a production function similar to orthopaedic surgery (but perhaps on a smaller scale): Size would allow for subspecialization (e.g., interventional cardiology and electrophysiology) and some ancillary services (such as stress testing and cardiac catheterization). Yet, the estimated model found no such effects.

The results presented in this paper suggest a potential paradox: The median size of actual physician practices is considerably smaller than expected by estimated production efficiencies. For example, both the all practices and the single-specialty orthopaedic surgery practice models found that the median practice size (8 FTE physicians) was at the minimum estimated level of production. All of the other models estimated that the size of the median practice was well below that expected from the estimated marginal products of physicians. We have two hypotheses for this behavior. First, physicians may have personal preferences for smaller practices. This hypothesis combines the classic target income hypothesis (as exemplified by McGuire and Pauly (1991) and Rizzo and Blumenthal (1994)) with the hypothesis that physicians may seek to maximize non-profit-related goals such as professional autonomy and service to patients (May, 1983; Starr, 1983). That is, physicians may be willing to sacrifice more production – and more income – in return for other personal and professional interests. If this hypothesis is confirmed, those who advocate for size and integration of physician practices to facilitate health care reform will find their task more difficult.

The second hypothesis is that physician practice size may not be rewarded in the market. A production function that rewards size is, in effect, a necessary but not sufficient condition for a profit-maximizing (or even profit-satisficing) function that rewards size. A firm, after all, is assumed to be interested in maximizing profit, not production. Nevertheless, the results shown in this paper demonstrate that a physical opportunity for growth exists for some (if not all) physician practices. If this finding holds, health care reform proposals that expect that physician practices must become larger and more integrated will need to identify or create incentives for practices to expand (through internal growth or merger). Fortunately for these proposals, the analysis in this paper did not find that physician practice production functions favor small size (the solo and 2-physician practices that represent the plurality of practices in the US), which would have made the task of integration much more difficult – and costly.

To conclude, this paper has attempted to extend the analysis of the structure of physician practices in three ways. The first advance is that we analyzed the practice – rather than the individual physician – as the unit of observation. Although physicians often render their services on a one-to-one basis with their patients, increasingly the care itself is produced by a team of health professionals. In addition, with physicians practicing in increasingly formal organizations, the practice is a firm as defined by Coase (1937) and Drucker (1994), and should be analyzed as such. The second advance is that – unlike previous research on the production of physician services – we estimated total production by the firm, not individual services (such as office visits). Third, by using the annual surveys conducted by the Medical Group Management Association, we had the advantage of a large enough sample to allow analysis at a finer level of granularity than previous studies. That is, we were able to analyze multispecialty and single-

specialty practices separately. In addition, we were able to model six subspecialty practice types within single-specialty practices: family practice, pediatrics, cardiology, orthopaedic surgery, obstetrics/gynecology, and anesthesiology.

This paper does have a number of limitations. First, because the MGMA survey is conducted of its membership of organized physician practices, it contains relatively few solo and two-physician practices. As a result, our analysis may not be as robust for these smaller practices. Second, even with access to three years of data, our sample size is still small enough that some single-specialty practices of interest (i.e., general internal medicine, gastroenterology, general surgery, and radiology) cannot be included in the analysis.

Finally, as noted earlier, identifying the structure of the physician practice production function is just the first step in determining whether or not an optimal size of physician practice exists. Nevertheless, the analysis presented here provides the foundation for the next phase of the research – the analysis of “contribution margin” (i.e., total gross charges for the practice minus all costs except physician compensation) – which, after all, is the real focus of a profit-maximizing practice.

Acknowledgements

We would like to thank Maqbool Dada, Adam Litwin, and Ravi Aron for their helpful critique of our research, Angela Mancuso Lipshutz and David Sinopoli for their research assistance on earlier versions of the paper, and Jim Margolis for his tireless efforts in database management.

References

American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, & American Osteopathic Association, March 2007. *Joint principles of the patient-centered medical home* Retrieved from <http://www.medicalhomeinfo.org/joint%20Statement.pdf> , December 21, 2009.

Burns, L. R., 1997. Physician practice management companies. *Health Care Management Review*, 22(4), 32-46.

Casalino, L. P., Pham, H., Bazzoli, G., 2004. Growth of single-specialty medical groups. *Health Affairs*, 23(2), 82-90.

Christensen, C. M., Grossman, J. H., Hwang, J., 2009. *The Innovator's Prescription: A Disruptive Solution For Health Care*. New York, NY: McGraw-Hill.

Coase, R. H., 1937. The nature of the firm. *Economica*, 4(16), 386-405.

Daschle, T., 2008. *Critical: What We Can Do About The Health-Care Crisis*. New York, NY: St. Martin's Press.

DeFelice, L. C., Bradford, W. D., 1997. Relative inefficiencies in production between solo and group practice physicians. *Health Economics*, 6, 455-465.

Drucker, P. F., 1994. The theory of the business. *Harvard Business Review*, 72(5), 95-104.

Frech III, H. E., Ginsburg, P. B., 1974. Optimal scale in medical practice: A survivor analysis. *The Journal of Business*, 47(1), 23-36.

- Freidson, E., 1988. *Profession of Medicine :A Study of the Sociology of Applied Knowledge : With a New Afterword*. Chicago: University of Chicago Press.
- Gaynor, M., Pauly, M. V., 1990. Compensation and productive efficiency in partnerships: Evidence from medical group practice. *Journal of Political Economy*, 98(3), 544-573.
- Getzen, T. E., 1984. A 'brand name firm' theory of medical group practice. *The Journal of Industrial Economics*, 33(2), 199-215.
- Halley, M. D., 2007. *The Primary Care -- Market Share Connection: How Hospitals Achieve Competitive Advantage*. Chicago, IL: Health Administration Press.
- Hartman, M., Martin, A., McDonnell, P., Catlin, A., 2009. National health spending in 2007: Slower drug spending contributes to lowest rate of overall growth since 1998. *Health Affairs*, 28(1), 246-261.
- Healthcare Advisory Board, 1999. *Stopping the Bleed: Reversing Losses in Owned Practices*. Washington, D.C.: The Advisory Board.
- Liebhaber, A., Grossman, J. M., August 2007. *Physicians Moving to Mid-Sized, Single-Specialty Practices*. Tracking Report No. 18. Washington, D.C.: Center for Studying Health System Change.
- Madison, D. L., Konrad, T. R., 1988. Large medical group-practice organizations and employed physicians: A relationship in transition. *The Milbank Quarterly*, 66(2), 240-282.

- Marder, W. D., Zuckerman, S., 1985. Competition and medical groups. *Journal of Health Economics*, 4, 167-176.
- May, W. F., 1983. *The Physician's Covenant: Images of the Healer in Medical Ethics* (1st ed.). Philadelphia: Westminster Press.
- McGuire, T. G., Pauly, M. V., 1991. Physician response to fee changes with multiple payers. *Journal of Health Economics*, 10(4), 385-410.
- Medical Group Management Association, 2007. *Cost Survey: 2007 Report Based on 2006 Data*. Englewood, CO: Medical Group Management Association.
- Nyweide, D. J., Weeks, W. B., Gottlieb, D. J., Casalino, L. P., Fisher, E. S., 2009. Relationship of primary care physicians' patient caseload with measurement of quality and cost performance. *The Journal of the American Medical Association*, 302(22), 2444-2450.
- Pauly, M. V., 1996. Economics of multispecialty group practice. *Journal of Ambulatory Care Management*, 19(3), 26-33.
- Pope, G. C., Burge, R. T., 1996. Economies of scale in physician practice. *Medical Care Research and Review*, 53(4), 417-440.
- Pope, G. C., Olmsted, E., Healy, D., Zuckerman, S., McFeeters, J., 2006. *Review of Physician Practice Expense Geographic Adjustment Data, Report to Centers For Medicare & Medicaid Services by RTI International*. CMS Contract No. 500-00-0024 T.O. #16.

- Reinhardt, U. E., 2000. The rise and fall of the physician practice management industry. *Health Affairs*, 19(1), 42-55.
- Reinhardt, U., 1972. A production function for physician services. *The Review of Economics and Statistics*, 54(1), 55-66.
- Rizzo, J. A., Blumenthal, D., 1994. Physician income targets: New evidence on an old controversy. *Inquiry*, 31, 394-404.
- Robinson, J. C., 1999. *The Corporate Practice of Medicine*. Berkeley, CA: University of California Press.
- Rosenman, R., Friesner, D., 2004. Scope and scale inefficiencies in physician practices. *Health Economics*, 13(11), 1091-1116.
- Shih, A., Davis, K., Schoenbaum, S. C., Gauthier, A., Nuzum, R., McCarthy, D., 2008. *Organizing the U.S. Health Care Delivery System for High Performance*. Publication No. 1155. New York, NY: The Commonwealth Fund.
- Sirovich, B., Gallagher, P. M., Wennberg, D. E., Fisher, E. S., 2008. Discretionary decision making by primary care physicians and the cost of U.S. health care. *Health Affairs*, 27(3), 813-823.
- Smart, D. R., 2004. *Medical Group Practices in the US, 2004 Edition*. Chicago: American Medical Association.
- Starr, P., 1983. *The Social Transformation of American Medicine*. New York: Basic Books.

Medicare payment policy: Ensuring stability and access through physician payments:

Subcommittee on Health of the Committee on Energy and Commerce, U.S. House of Representatives, 107th Congress, 2nd Session, 2002.

Thurston, N. K., Libby, A. M., 2002. A production function for physician services revisited. *The Review of Economics and Statistics*, 84(1), 184-191.

Tollen, L., April 2008. *Physician Organization in Relation to Quality and Efficiency of Care: A Synthesis of Recent Literature*. New York, NY: The Commonwealth Fund.

Town, R., Wholey, D. R., Kralewski, J., Dowd, B., 2004. Assessing the influence of incentives on physicians and medical groups. *Medical Care Research and Review*, 61(3), 80S-118S.

Wagner, S. L., 2005. Organization and operations of medical group practice. In L. F. Wolper (Ed.), *Physician Practice Management, Essential Operational and Financial Knowledge* (pp. 29 – 62). Sudbury, MA: Jones and Bartlett Publishers.