M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 11, No. 3, Summer 2009, pp. 525–537
ISSN 1523-4614 | EISSN 1526-5498 | 09 | 1103 | 0525
informs®
DOI 10.1287/msom.1080.0236

Division of Labor in Medical Office Practices

Gregory Dobson, Edieal Pinker
Simon School of Business, University of Rochester, Rochester, New York 14627
{greg.dobson@simon.rochester.edu, pinker@simon.rochester.edu}

R. Lawrence Van Horn
Owen School of Management, Vanderbilt University, Nashville, Tennessee 37203,
larry.vanhorn@owen.vanderbilt.edu

This paper examines the staffing, division of labor, and resulting profitability of primary care physician practices. Division of labor is viewed as a mechanism to increase the efficiency of production processes through specialization. At the same time, division of labor also introduces coordination cost as handoffs and communication needs increase. We attempt to empirically assess the net effect in primary care physician offices. We collected data from a sample of these practices and tested two hypotheses: (H1) controlling for staff size, greater delegation through the use of more staff types will decrease the throughput of visits, and (H2) controlling for staff size, income per unit time generated by the practice is decreasing in the number of staff types. We find evidence supporting both hypotheses. We conclude that many physicians are gaining little financial benefit from delegating work to support staff. This suggests that small practices with few staff may be viable alternatives to traditional practice designs.

Key words: healthcare management; service operations; survey research; econometric analysis
History: Received: May 24, 2007; accepted: June 24, 2008. Published online in *Articles in Advance* October 24, 2008.

1. Introduction

Division of labor has long been viewed (Smith 1776) as a mechanism to increase the efficiency of production processes. Specialization is assumed to lead to higher productivity at the narrowly defined tasks. At the same time, consolidation of tasks reduces waste from communication during handoffs between process stages and set-ups, as well as blocking and starving that occurs in stochastic series systems. Therefore, there are clearly some trade-offs between these two approaches to organizing work: specialized division of labor and consolidation of tasks. In some work environments, especially professional services, there is a form of division of labor called delegation. Delegation occurs when a task that one, usually a higher skilled, worker can do, is assigned to a lower skilled worker. This division of labor has a hierarchical element to it. One purpose of delegation is to assign tasks to the lowest cost worker who can perform it. Hierarchical differentials in skills and wages create additional considerations in designing a business process, which we explore here.

A physician's office practice is an example of a work environment in which delegation is very common. Physicians are advised to "delegate all duties that don't require a physician's license," with the explanation that "even if you have to hire an extra staffer or increase a part-timer's hours to handle the extra workload, the increase in your own productivity should offset the higher personnel cost," (McKee 2004). The logic behind this advice is as follows: each encounter between a patient and a physician practice is composed of both a physician and a non-physician component. Because of legal and regulatory constraints, only the physician can deliver the physician component of the service. In the extreme, if the entire encounter consisted of only the physician component, physicians would be unable to increase visit volume through the use of support staff. The principal mechanism for increasing throughput in such an environment is the use of automation and, in the case of an office visit, this means information technology (IT). However, this extreme situation does not occur in medical services. Each patient encounter requires many nonphysician service components:[1] scheduling, simple medical care,

[1] The definition of physician vs. nonphysician components is determined by legal constraints on the care to be delivered. For example,

billing, reporting, and various administrative functions. By assigning these tasks to support staff, the physician has more time available to see patients and can potentially increase the practice's throughput. Given that the wage differential between physicians and support staff is often large, many assume that it is efficient to assign as many tasks as possible to support staff, implicitly assuming that tasks will take roughly the same time when done by support staff and that the steps required in delegation are negligible. Group practices are believed to facilitate this delegation process because multiple physicians can share support staff and, in particular, delegate administrative functions such as human resource management to a shared office manager. Some have identified a trend toward group practices and interpreted it as pro forma evidence of a survivor bias, i.e., an indication that the solo practice model is inefficient (Marder and Zuckerman 1985, Frech and Ginsburg 1974).

In recent years, this conventional wisdom has been challenged by some who are advocating for "lean" practices (e.g., Moore 2002a, b). They argue that, as discussed above, the benefits of operating a practice with a large support staff and high visit volumes are tempered by a variety of inefficiencies that we collectively term "coordination costs," and therefore do not yield significant financial benefits to physicians. At the same time, these high-volume practices have a number of drawbacks. In a 1999 Kaiser Family Foundation/Harvard School of Public Health survey of physicians and nurses, 83% said that managed care had decreased the amount of time doctors spend with their patients, and 72% replied that the quality of health care declined under managed care. When asked to reflect on their own experiences, 58% of physicians said that spending more time on administration rather than directly with patients was "a great concern." "Physicians are getting very frustrated with abuses in managed care and physicians are trying to get very creative and do anything they can to cope" (Dr. Donald Palmisano, President of the AMA Kowalczyk 2004). Dissatisfaction among physicians rose to 19% in 2001 up from 14.3% in 1997, (Landon 2003). Researchers have linked higher proportions of managed care revenue and their correlated administrative burden and reduced prices,[2] with increases in physician dissatisfaction. Other factors influencing physician satisfaction include changes in income, increasing hours worked, difficulty obtaining high-quality services, and decreasing clinical autonomy.

The goal of our research is to better inform this debate. We report on an empirical study of office practices designed to test whether increased delegation introduced inefficiencies to such a degree that the benefits of division of labor are significantly degraded. The data support the conclusion that controlling for staff size, greater delegation through the use of more staff types will both decrease the throughput of visits and the physician's income per unit time worked. These results lend credence to the arguments of those advocating for "lean," minimally staffed practices. The issues we address in this paper also apply to other professional services. We believe that a similar methodological approach could be applied to various medical specialties, as well as legal, accounting, and engineering design offices.

Our paper is organized as follows. In §2, we review the relevant literature. In §3, we develop the motivation for our study. In §4, we describe the data and survey. In §5, we present the results of regression analyses that link staff structure and practice productivity. We conclude in §6.

2. Literature Review

It has long been recognized in the operations literature that division of labor introduces a variety of costs into work processes. Kilbridge and Wester (1966) develop a framework for understanding the economic trade-offs in division of labor. One of the costs of a greater division of labor that they identify is an increase in communications and control activities that are nonproductive, caused by the interdependence of functions and handoffs between workers. They also identify setups as a source of costs. Buzacott (2004) demonstrates how division of labor in production settings can lead to inefficiency when workers are of differing productivity. Rummel et al. (2005) model the

only a physician may perform certain types of evaluations or write prescriptions. The physician is free to supply "nonphysician" components of the visit if he or she is so inclined.

[2] Under managed care, prices are often reduced relative to historical usual and customary charges and indemnity fee schedules.

optimal consolidation of tasks in arbitrary acyclic processes using mixed-integer linear programming. Their modeling effort is motivated by the fact that consolidation of tasks can eliminate handoffs and other coordination costs in multistep processes, thus speeding up processing while incurring new costs related to the effort required to implement the new process structure and lost specialization benefits. They do not model staffing nor do they directly optimize profit. In a series of papers, Seidmann and Sundararajan (1997a, b, c) investigate the factors that influence the benefits of triaging using queueing analysis. They identify factors such as knowledge intensivity, degree of task complexity and coupling, variability in processing times, and task asymmetry. Powell (2000) also models the optimal assignment of tasks to workers, considering the trade-offs between loss of specialization and reduction in stochastic interference between tasks. In his model, the objective is to minimize processing time and staffing costs, which are excluded. Pinker and Shumsky (2000) and Misra et al. (2004) explicitly model learning and job tenure processes to better understand the importance of specialization in service systems.

The work done on optimizing staff allocation in health care has, to a large extent, ignored the costs of division of labor. Ittig (1978) formulates a linear programming planning model for an ambulatory care system that determines the optimal allocation of patient types/needs to physician types, (i.e., it determines an optimal division of labor in a health care setting). However, he does not incorporate any of the potential costs of division of labor. Swisher and Jacobson (2002) analyze a simulation model of a health care clinic seeking to optimize the staff configuration. In their model, task allocations are fixed and no accounting is done for the costs of coordinating a larger support staff. Schneider and Kilpatrick (1975) use a mixed-integer planning model to determine division of labor of physicians and medical support staff, and include additional supervisory time as a cost of using support staff.

All of the studies in the operations literature listed above are predicated on the assumption that there are clear trade-offs in the optimal division of labor. Many of the studies focus on the role of specialization and various stochastic effects such as starving, blocking, and pooling. These studies are also typically focused on minimizing processing times. This research differs from those papers in that we focus on practice profits as the main performance metric and conduct an empirical investigation of the subject.

In the health care literature, prior studies have empirically examined the efficiency of group practices compared to solo practices and found that group practices were more efficient (Reinhardt 1972, 1975; Pope and Burge 1996). However, there is some evidence that as the number of physicians increases, the efficiency of the group decreases (Reinhardt et al. 1979), with the optimal number of physicians being 2.4 (Pope and Burge 1996). Researchers have also considered the efficiency with which physician practices use support staff, with conflicting results. Some studies found that physicians underuse support staff (Reinhardt 1972, 1975), whereas others have suggested overuse (Brown 1988 and Thornton 1998). This paper seeks to contribute to this stream of research by examining how the number of support staff and number of different types influence productivity and profitability.

The existing literature is notably silent on the details of the process of care delivery, focusing instead on general relations between input quantities and output levels across many markets, environments, and practice settings. These differences make it difficult to draw inferences about the optimal practice organization to the extent that such organization is dependent on the market setting in which the practice is located. Our study is different. First, we focus on a homogeneous set of internal medicine practices in a single market facing the same input prices, capital costs, and market conditions. This focus allows us to eliminate input price differences as a driver of practice organization. Second, there is a single dominant private health insurance plan with 80% of the nongovernment market, and thus the physicians in our sample are all subject to the same fee schedule. Third, our paper is framed from the perspective of an individual physician, be he a solo practitioner or a member of a group. As a result, our goal is to develop a better understanding of how practice size in terms of the number of visits and number of support staff used *per physician* are related to the operating efficiency of the practice.

3. Motivation

Almost all physicians hire support staff (secretaries, nurses, practice managers) to whom they delegate much of the work in the practice. Initially, the trade-off appears clear. The wages of the staff are low compared to what the physician expects to earn, so it would seem that allocating as many "nonphysician" components to the staff frees the time of the physician to do more "physician" work, seeing more patients and producing more revenue. For illustrative purposes, let us assume that a single patient visit requires 60 minutes of labor time. If a physician has no support staff, he can see one patient per hour. If the physician hires a nurse who can complete 30 minutes of that work, the physician may then see two patients per hour, thereby doubling his revenue while incurring the additional cost of a single nurse's salary. If specialization leads to greater efficiency, for both the nurse and physician, doing the original 30 minutes of work in 25 minutes, the physician may even do better than doubling his revenue. To the naïve observer, the more a physician can delegate work to less expensive support staff, the more he can increase his profits. In reality, however, when work is divided amongst many individuals, it becomes necessary to coordinate their activities to produce the same output. For the purposes of this paper, we define coordination time as all unproductive time that is added to the work processes as a result of delegation and division of labor.

Conceptually, we divide the work time associated with a patient visit into three categories: processing time, patient setup time, and staff interaction time. *Processing time* is defined as the minimum time it would take to deliver the patient visit independent of the way it is delegated. It can best be thought of as the time it would take for the physician to process a patient visit if he or she did it alone.[3] *Setup time* is defined as the additional labor time expended on a patient visit because the patient is interacting with multiple staff members. For example, some time is required to greet each patient or for the staff to orient themselves. *Staff interaction time* is additional labor time expended on a patient visit because the practice staff must interact with each other to coordinate service delivery. Interaction and setup time are therefore unproductive time that is added to the processing time for a patient visit because there are more staff and a greater division of labor. These are the coordination costs of division of labor for the purposes of this paper. To illustrate how interaction arises in a medical practice, we describe two common processes in primary care practice: (1) refilling a patient prescription and (2) delivering a standard office visit (see Tables 1 and 2). The following descriptions are stylized examples based on our observation of numerous primary care practice office operations.

Table 1 Process Steps for a Prescription Refill

Process step description	Staff involved
Patient calls for a refill. Record the relevant information.	Secretary
Pull the patient chart.	Secretary
Give the request(s) to the nurse and discuss any special issues.	Secretary and nurse
Review the case.	Nurse
Discuss if appointment is necessary.	Nurse and physician
If appointment is necessary, schedule it.	Nurse and secretary
Call the pharmacy and provide all relevant information.	Nurse
Record information in chart.	Nurse
Return chart to the secretary.	Nurse
Refile chart.	Secretary

Two things become clear from these descriptions. First, a substantial amount of work has been delegated to the staff and, second, as a result, a substantial amount of communication or interaction between all the parties is required. Examples of such interaction are (1) the secretary must talk to the nurse or physician to determine the visit urgency before giving that patient an appointment that day and (2) the nurse must pass on the information he or she has collected from the patient to the physician. If there are multiple staff of a particular type, there are also likely interactions among the staff of the same type. For example, one secretary may misplace a chart and involve another secretary in helping to find it, or if one secretary speaks to a patient on the phone, he or she may later ask another to call the patient back, thus requiring some hand off of information. It is important to note that when one staff member interacts with another for X minutes, they each must expend

[3] In this particularly stark example of the physician being the sole provider, some patient setup time would be incorporated into the processing time. For example, it is necessary for the physician to greet the patient and ask if they like the weather.

Table 2 **Process Steps for a Patient Visit**

Process step description	Staff involved
Patient calls to set up an appointment.	Secretary
Determine urgency of visit.	Secretary, nurse, maybe physician
Call patient back to schedule the appointment.	Secretary
Before the patient visit, pull the chart from the file cabinet.	Secretary
Patient arrives and checks in.	Secretary
Greet patient and escort him/her to the exam room.	Nurse
Weigh the patient and take blood pressure and temperature.	Nurse
Ask about the presenting problem or reason for the visit.	Nurse
Nurse and physician exchange information.	Nurse and physician
Greet patient and ask about the presenting problem or reason for the visit.	Physician
Examine patient, order tests or treatment.	Physician
Fill out encounter form.	Physician
If tests are required, explain to nurse.	Physician and nurse
Perform test.	Nurse
Communicate results of test.	Nurse and physician
Make notes on the chart and return it to the secretary.	Physician
Patient brings encounter form to the secretary and pays the copay.	Secretary
Refile the chart.	Secretary
Determine if info on encounter form is accurate enough for coding; clarify any problems.	Practice manager, nurse, maybe physician.
Code the claim (possibly sent to a third party).	Practice manager or secretary

the X minutes. Therefore, $2X$ minutes are added to the processing time consumed by all resources for the visit.

We believe that there are four forces at work in a medical practice. First, increasing support staff enables the individual physician to decrease the amount of time spent per patient, thus increasing the number of visits he or she and the practice can process per unit time. At the same time, greater division of labor enables the practice to assign tasks to the least expensive resources, therefore reducing costs. Third, the coordination costs created by division of labor influence performance. Finally, division of labor can increase efficiency by leading to specialization. In this environment, we do not expect the benefits of specialization to be significant relative to the inefficiencies introduced by division of labor. For example, we expect a nurse to be more efficient at taking blood pressure than a physician who always delegates this to a nurse. However, if a physician designs his or her practice so he or she consistently takes blood pressure, we expect the volume of activity to be sufficient for him or her to be no less efficient at the task than the nurse. In fact, one could argue that workers with broader task definitions can parallel process and gain efficiency. Given this reasoning, we hypothesize that the net effects of division of labor are as follows.

Hypothesis 1 (H1). *Controlling for staff size, greater delegation through the use of more staff types will decrease the number of visits a practice can process per day.*

If H1 holds, the question remains whether the decrease in visits and resulting revenue loss caused by coordination costs outweighs the cost savings of using lower cost staff. We hypothesize that it does.

Hypothesis 2 (H2). *Controlling for staff size, income per unit time generated by the practice is decreasing in the number of staff types.*

4. Data

Our objective in this paper is to develop a better understanding of the relationship between the organizational structure of a primary care practice and its profitability. To support this effort, we designed a survey[4] (see the appendix for the instrument) to generate data to validate our two empirical hypotheses on how the number of staff types affects the throughput and profitability of a practice. Next, we describe our sample and survey methodology.

4.1. Sample

We examined primary care practices in two ways. We first interviewed several physicians and their staff and observed the operations of several practices. Second, we surveyed internal medicine practices in the greater metropolitan area of Rochester, New York to collect detailed operational and financial information. All the physicians in our sample were subject to the same reimbursement environment dominated by fee-for-service payment arrangements with one insurer

[4] We developed the survey with input and support from community physician leaders. The survey was piloted and the results verified to ensure that we were obtaining the type of information that we desired. The survey is available on the *Manufacturing & Service Operations Management* website (http://msom.pubs.informs.org/ecompanion.html).

covering 80% of the market and one other covering the rest. The survey was mailed to 500 internal medicine and family practice physicians[5] within the region between August and September 2003. The physicians were asked to report on the practice operations of 2002. The mailing was repeated to those practices that did not respond after two weeks. From these two mailings, we obtained 105 completed surveys for a response rate of 21%. After removing outliers and practices with anomalies[6] in their responses, we were left with 61 data points.

We have a high degree of confidence in the data contained in the usable responses based on a number of efforts employed to validate them. First, our survey included questions that afforded internal consistency checks. If responses were internally inconsistent, when possible, the inconsistencies were resolved by followup telephone calls to the practice. Second, we had financial and operating data from a physician hospital organization (PHO) on a subset of respondents, which we compared with the survey responses. Finally, we compared our aggregate practice statistics with the published results of other studies of physician operations,[7] and found our data to be consistent with these other studies. We base the following discussion on our observations of these 61 practices. Descriptive statistics of the key operational measures of our sample are discussed in §4.2 and a regression analysis of the data is presented in §5.

[5] We were not able to a priori determine a list of internal and family medicine physicians who were in active practice in a primary care office setting. The surveys were physically mailed to 600 physicians, but a large number (approximately 100) were in administrative positions, working in institutions such as nursing homes or hospitals, etc. As such, we use 500 as the population size.

[6] By anomalies, we mean surveys that were incompletely or incorrectly filled out or with data that failed internal consistency checks. Outliers were, for example, members of very large group practices or physicians who worked very few hours.

[7] Table 3 is consistent with that reported in other studies in the literature. For example, Thornton (1998) reports on data from the Physicians Practice Costs and Income Survey done in 1984 for solo practitioners across multiple specialties. He reports that the average physician worked 59 hours per week, saw approximately 2.09 patients per hour worked, and employed an average of 3.12 support staff, including midlevel providers. Guglielmo (2003) reports that, median 2002 pretax, income for internists was $150,000. Terry (2004) reports that primary care physicians worked on average 50 hours per week and saw 108 patients per week.

4.2. Descriptive Statistics

The typical practice in our survey employed a secretary, a nurse, and an office manager and processed 90 visits per week, with the physician working at least 50 hours per week and earning $130 K per year. There is, however, considerable variation among these practices. A summary of key descriptive statistics appears in Table 3.

Less than 10% of the practices in our sample employed electronic medical records, meaning that paper records were the dominant method for storing, retrieving, and transmitting patient information. Although the mean visits per hour worked in the sample is 1.85, it does not reflect the pace at which the physician interacts with patients because it includes work time outside of direct patient contact. Patient appointments are typically scheduled in 15–20 minute slots with the physician spending only a portion of this allocated time in the physical presence of the patient. However, the physician may spend considerably more than 15 minutes on work related to an individual patient visit.

The physicians in our sample can be grouped into one of three types of practices, a self-owned solo practice (S-solo), a self-owned group practice (S-gp), or an employee of group practice (E-gp) owned by some health care organization. In Table 4, we compare these three types of practices with respect to some of the descriptive statistics reported for the entire sample in Table 3.

We can see from Table 4 that employed group practices are larger (have more physicians) than self-owned groups, but both types of group practices are similar in their use of office managers, support staff, hours worked, number of patients seen per hour, and labor time expended per visit. However, we also see that the E-gp practice physicians earn on average $21,400 more annually than the S-gp practices and make heavier use of midlevel providers. It is important to note that total physician annual income is essentially equivalent to total practice profit in the cases of the S-gp and S-solo practices, whereas it is not so in the E-gp practices. In fact, it is quite common for E-gp practices to lose money because the health systems that own them view them as channels for generating referrals for other medical services. For example, the Medical Group Management Association reported

Table 3 Descriptive Statistics of Sample

Measure	Mean	Median	Standard deviation	Min.	Max.
Percentage of practices with an office manager 79%					
Percentage of practices with a midlevel provider 52%					
Patient panel size	2,639	2,200	1,558	500	4,500
Number of support staff types[a]	2.45	3	0.85	0	3
Number of support staff per provider[b]	1.93	1.8	0.90	0	5
Physician hours worked per week	51.75	51.5	11.78	23.25	85
Visits processed per week	93.5	90	29.56	20	170
Physician annual net pretax income[c] ($)	146 K	130 K	51 K	70 K	290 K
Physician net pretax income per hour worked[d] ($)	60	55.56	18	29	112
Patient visits per hour of physician labor[e]	1.85	1.76	0.53	0.68	3.13
Total minutes of labor time per patient visit[f]	81.04	78.33	26.65	34.00	153.46

[a] Support staff are defined as any staff other than physicians, nurse practitioners, and physician assistants. Physicians in our survey reported as many as six distinct types of support staff, including RNs, LPNs, secretaries, office managers, medical records, and billing assistants. For the purposes of this study, we defined three possible types of support staff: nursing, secretarial, and office management. All support staff reported by physicians were classified as one of these three types.

[b] This is calculated by dividing the number of support staff by the total number of physicians and midlevel providers in the practice. Because physicians typically see more patients and work longer hours than midlevel providers, this underestimates support staff usage by physicians. For group practices, we do not know how support staff hours are allocated among individual providers, and therefore assume that they are equally divided among all providers. Ideally, one would divide the support staff according to relative visit volumes. Because we do not have the visit volumes for all members of a group practice, we cannot do this. Therefore we clearly may be introducing distortions if the responding physicians have visit volumes that differ greatly from their partners or if some physicians in the practice have chosen to independently hire more staff than the others. Since there is no reason to believe that the respondents of our survey who are in group practice have greater or lesser staff use than their colleagues who did not respond, we do not believe that we have introduced any bias into our results.

[c] Our surveys asked physicians to report their annual income within $20 K ranges to provide some level of privacy and hopefully boost response rates. For the purposes of our analysis, we infer physician incomes as the medians of the reported ranges.

[d] For example, if a physician reported working 60 hours a week, 48 weeks per year, and a pretax income of $140,000, then his income per hour worked is $48.60.

[e] Physicians reported their work hours and the number of visits they processed per week from which we derived patient visits per hour.

[f] Physicians reported the number of full-time equivalents (FTEs) employed by their practices. Using the assumption of 35 hours of work per week per FTE and the physicians reported hours, we generated a total number of labor minutes expended per week in the practice. Dividing this quantity by the number of patient visits per week yields the labor minutes per visit.

in 1999 (MGMA 1999) that the average multispecialty primary care group lost $16,662 per FTE physician. In our sample, we know that all the practices belonging to the PHO, for which we have supplemental data, had significant financial losses per FTE physician. Comparing the group practices to the S-solo practices, we see that solo practices make much less use of office managers and midlevel providers, work more hours per week, and see more patients, expending significantly less total labor time per visit than group practices ($p = 0.068$). The solo physician's income per hour worked is similar to that of the S-gp physician's and, as noted before, S-solo and S-gp practices are profitable as opposed to the E-gp practices.

In Figure 1, we plot dollars per physician hour worked versus visits processed per physician hour worked for self-employed physicians. The four series identify practices with zero, one, two, and three types of support staff. Figure 2 shows the same data but the four series identify practices with differing numbers of support staff *per physician*. We can see that there is substantial variability in these performance metrics. Because a medical practice's revenue is payments for patient visits, the number of visits worked per hour is a measure of the physician's revenue. The physician's income per hour worked is a measure of cost effectiveness because it accounts for the cost of running the practice. Looking at performance on a per hour basis is useful because it normalizes differences in the number of hours each physician works per year.

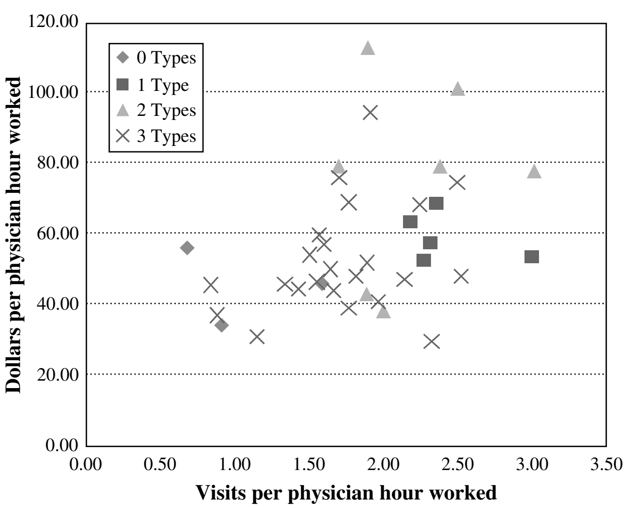
In Figures 1 and 2, we observe that physicians who are making approximately $55 per hour are processing from less than one to as many as three visits per hour. Similarly, we see that physicians processing between two and three visits per hour earn between $35 and $100 per hour. A practice in which the physician

Table 4 **Comparison of Descriptive Statistics for Three Practice Types**

	Solo ($N = 21$)		Group physician owner ($N = 17$)		Group employed ($N = 23$)	
Measure	Mean	Median	Mean	Median	Mean	Median
Number of physicians	1	1	3	3	5.47	4
Percent of office manager (absolute) (%)	52		94		91	
Percent of midlevel provider (absolute) (%)	33		47		74	
Support staff types	1.86	2	2.88	3	2.91	3
Support staff per provider	1.93	1.71	2.04	1.83	2.03	1.76
Physician hours per week	55.76	55	50.50	48	49.02	50
Physician annual net pretax income ($)	152 K	130 K	131 K	130 K	152 K	150 K
Physician net pretax income per hour ($)	57.52	51.95	54.85	51.10	65.49	60.81
Total labor per patient visit (minutes)	73.12	66.00	85.40	81.19	85.06	85.78
Visits per physician hour	1.86	1.89	1.84	1.70	1.83	1.67

processes three visits per hour is clearly different from one in which the physician processes one visit per hour. Given the homogeneity of the market environment (fee schedules and labor costs), it is likely that these differences in financial and operational performance result from cost differences linked to staffing configurations with varying efficiencies. The higher volume practices in our sample have larger panels of patients and a larger support staff. We can see that the practices with no support staff tend to see the fewest number of patients. Interestingly, practices with just one type of support staff tend to see more patients per hour than practices with three types of support staff.

Figure 1 **Dollars of Physician Pretax Take-Home Income per Hour of Physician Work as a Function of Average Number of Patients Seen per Hour (for Self-Owned Practices)**



Note. Number of support staff types identified.

Also, the practices with the most support staff are not necessarily those processing the most visits. Although an analysis of quality of care is beyond the scope of this paper, it is worth noting that the amount of time a physician spends per patient visit can potentially influence the quality of care as well.

5. Regression Results

To test H1, we estimate a visit production function using a standard Cobb-Douglas-type specification.

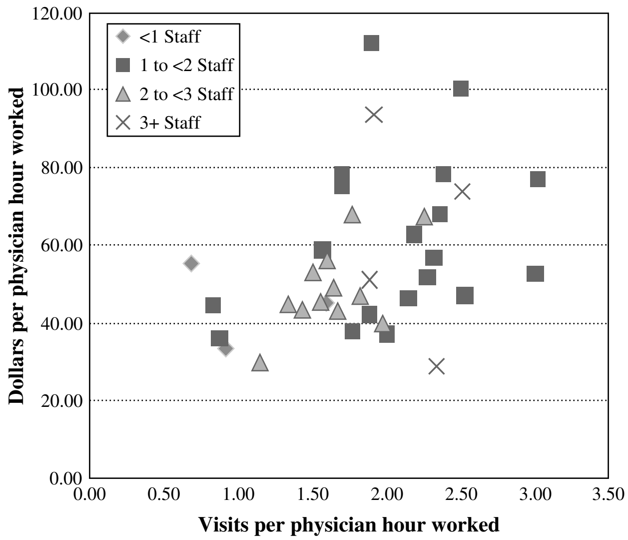
$$V_j = \theta_0 G_j^{\theta_1} A_j^{\theta_2} e^{\theta_3 i_{sj}} L_j^{\theta_4} M_j^{\theta_5} e^{Z_1 R_j} e^{Z_2 U_j} e^{\varepsilon_j}, \tag{1}$$

where

V_j = patient visits per week for physician j.
G_j = percentage of patients insured by Medicare or Medicaid programs for physician j.
A_j = average age of patients (as reported by practice) for physician j.
i_{sj} = indicator variable: 1 for self-employed physicians and 0 otherwise for physician j.
L_j = number of FTE employees for physician j.
M_j = number of different staff types producing the visits (including the physician) for physician j.
R_j = indicator variable: 1 for a practice in rural zip code and 0 otherwise for physician j.
U_j = indicator variable: 1 for a practice in an urban zip code and 0 otherwise for physician j.
ε_j = normal error term with mean 0.

Taking the log of Equation (1) linearizes the effects of the right-hand side variables. We estimate the equations in a log–log specification to aid in the interpretation because the coefficients from the log terms in

Figure 2 **Dollars of Physician Pretax Take-Home Income per Hour of Physician Work as a Function of Average Number of Patients Seen per Hour (for Self-Owned Practices)**



Note. Number of support staff per physician identified.

the model can then be interpreted as elasticities. We use ordinary least squares to estimate the following:

$$\ln(V_j) = \ln(\theta_0) + \theta_1 \ln(G_j) + \theta_2 \ln(A_j) + \theta_3 i_{sj} + \theta_4 \ln(L_j) + \theta_5 \ln(M_j) + Z_1 R_j + Z_2 U_j + \varepsilon_j. \quad (2)$$

As noted before, the physician's income has a different interpretation for a self-employed versus an employed physician. Employed physicians are reporting their personal income (or salary), whereas self-employed physicians are reporting the net profit of the practice. When making the link between visits and income, we must account for this difference and do so by including an indicator variable i_s for employment status in both equations. We also need to control for potential differences in the patient population across physicians, because differences in patient severity and case complexity may influence the number of visits or revenue per visit. Therefore we include the average patient age and the percentage of the patients in the practice that are either Medicare or Medicaid. We also include the physician's office zip-code classification to capture demographic differences among patient panels. Because our sample included 30 zip codes, we classified each as rural, urban, or suburban. We document this classification in the appendix and include dummy variables for rural and urban locations. Another possible source of performance variability may be differences in operational environments. We already control for employment status of the physician. Differences in insurance payers or additional services provided could, in theory, also affect practice staffing and performance. However, in the market, we surveyed and at the time of the survey there was one dominant payer with 80% of the non-government market with one other company covering the rest. Therefore, all practices faced the same payer environment. Regarding other services provided, we included questions about other revenue sources in our survey. For all the physicians included in our data set, these revenue sources turned out to be zero or insignificant. Table 5 presents the results of the estimation. For Model 1, both θ_4 and θ_5 are significant with opposite signs. θ_5 is significant and negative ($\theta_5 = -0.28$, $p = 0.0736$), indicating that for a given

Table 5 **Regressions of Log of Patient Visits/Hour on Practice Characteristics**

Log visit equation	Model 1	Model 2	Model 3
θ_0 intercept	**−0.5062** (0.5871)	**−0.6086** (0.5089)	**−0.8009** (0.4209)
θ_1 log percent of Medicare and Medicaid	**0.0349** (0.6130)	**0.0391** (0.5666)	**0.0340** (0.6291)
θ_2 log average patient age	**0.2539** (0.2753)	**0.2634** (0.2535)	**0.2273** (0.3597)
θ_3 self-employed (0/1)	**0.0403** (0.6576)	**0.0422** (0.6392)	**0.0426** (0.6418)
Z_1 rural	**0.0449** (0.6740)	**0.0733** (0.4892)	**0.0845** (0.4595)
Z_2 urban	**−0.1093** (0.4120)	**0.1533** (0.2633)	**−0.1506** (0.2852)
θ_4 log number of staff/physician (include physician)	**0.2905*** (0.0751)	**0.2759*** (0.0648)	**0.3074*** (0.0848)
θ_5 log number of staff types (include physician)	**−0.2812*** (0.0736)		
θ_6 number of types greater than two (include physician)		**−0.3164**** (0.0380)	
τ_1 number of types = 1			**0.3939** (0.1632)
τ_2 number of types = 2			**0.2891*** (0.0917)
τ_3 number of types = 3			**−0.0399** (0.7600)
Observations	61	**61**	**61**
R-square	0.10	**0.12**	**0.12**

Notes. Significance levels: $^*<0.1$, $^{**}<0.05$, $^{***}<0.01$.

staff size, more types of workers, or greater division of labor, reduces the number of visits that can be processed, supporting H1. For example, if a practice with two support staff of the same type (e.g., two nurses) can see X visits per hour, then we expect a practice with one nurse and one secretary (a 50% increase in the number of types—from two to three with doctors included) can see (0.28)(50%) = 14% fewer visits per hour or 0.86X visits per hour. θ_4 is significant and positive ($\theta_4 = 0.29$, $p = 0.0751$), supporting the conventional wisdom that more support staff enable the physician to see more patients. One of the challenges of estimating this production function that may be weakening the significance of our results is the fact that practices with more staff types in our sample tend to have more total staff as well. The correlation between the two quantities is 0.41, which leads to some colinearity effects.

Because of the competing and possibly nonlinear effects of staff number and number of staff types, we also consider an alternative specification (Model 2) in which instead of the number of staff types, we use an indicator variable N_j for whether or not the number of staff types exceeds two (Equation (3)).

$$\ln(V_j) = \ln(\theta_0) + \theta_1 \ln(G_j) + \theta_2 \ln(A_j) + \theta_3 i_{sj} + \theta_4 \ln(L_j) + \theta_6 N_j + Z_1 R_j + Z_2 U_j \varepsilon_j. \quad (3)$$

This model allows the possibility that there is a benefit to productivity when the number of types is low, whereas the coordination costs dominate the benefits of division of labor when it is higher. We can see in Table 2 that this, in fact, happens as θ_6 is significant and negative ($\theta_6 = -0.3164$, $p = 0.0380$).

In Model 3, we have separate dummies for one, two, or three types of staff because the number of staff types ranged from one to four. We see that the coefficients are decreasing in the number of types, which is consistent with the result in Model 2. Only τ_2, the coefficient for two types, is significant ($\tau_2 = 0.2981$, $p = 0.0917$).

To test H2, we regress income per hour worked (π_j) on staffing variables. We construct four alternative models. In all the models, we control for the average patient age, percentage of patients on Medicare or Medicaid, whether the physician is self-employed or not, and the office location type (rural, urban, or suburban). The results are displayed in Table 6. Model 1 includes the number of staff per physician, L_j, as an explanatory variable (Equation (4)). Model 2 adds the number of staff types, M_j (Equation (5a)). To allow for a possibly nonlinear relationship for the number of staff types, in Model 3, we replace M_j with an indicator variable, N_j, for whether or not the number of staff types exceeds two, (Equation (5b)) as we did for the visit regression. In Model 4, we have separate dummies for one, two, or three staff types.

$$\ln(\pi_j) = \ln(\beta_0) + \beta_1 \ln(A_j) + \beta_2 \ln(G_j) + \beta_3 i_{sj} + \beta_4 \ln(L_j) + Z_1 R_j + Z_2 U_j + \varepsilon_j \quad (4)$$

$$\ln(\pi_j) = \ln(\beta_0) + \beta_1 \ln(A_j) + \beta_2 \ln(G_j) + \beta_3 i_{sj} + \beta_4 \ln(L_j) + \beta_5 \ln(M_j) + Z_1 R_j + Z_2 U_j + \varepsilon_j \quad (5a)$$

$$\ln(\pi_j) = \ln(\beta_0) + \beta_1 \ln(A_j) + \beta_2 \ln(G_j) + \beta_3 i_{sj} + \beta_4 \ln(L_j) + \beta_6 N_j + Z_1 R_j + Z_2 U_j + \varepsilon_j. \quad (5b)$$

In Model 1, the number of staff is not significant because the correlation between staff number and staff types makes it difficult to see the positive impact of the staff number. The negative impact of staff types are mixed in when staff types are not controlled for. When we add the number of staff types in Model 2, the number of staff becomes significant and its coefficient increases. Although negative, the staff types coefficient is not significant. Because we believe the effect of staff types is complex and possibly nonlinear, in that it decreases costs while reducing revenue at the same time, and because there are only four possible values for the number of staff types, in Model 3, we try an alternative specification for the number of staff types using a dummy variable. This specification leads to both staffing variables (number of staff and dummy for number of types) becoming significant. In Model 4, we obtain a better fit from changing the way staff types are modeled and see that the coefficient on the number of staff consistently increases as more variables for staff types are added. This result is consistent with the notion that increasing staff not only has benefits, but also leads to more division of labor through types, which has some negative effects. Accounting for the effect of staff types leads to seeing more clearly the positive effect of staff increases. The coefficient on the types-greater-than-two dummy is a

Table 6 Regressions of Log of Income/Hour on Practice Characteristics

Coefficients	Model 1		Model 2		Model 3		Model 4	
Constant β_0	**3.5496*****	(0.0002)	**3.6232*****	(<0.0001)	**3.5494*****	(<0.0001)	**3.6623*****	(<0.0001)
Log average patient age β_1	**0.1423**	(0.5182)	**0.1549**	(0.4785)	**0.1680**	(0.4311)	**0.0645**	(0.7731)
Log percent of Medicare/Medicaid β_2	**−0.0294**	(0.6532)	**−0.0331**	(0.6101)	**−0.0303**	(0.6318)	**−0.0410**	(0.5212)
Self-employed β_3	**−0.1697***	(0.0531)	**−0.1651***	(0.0578)	**−0.1618***	(0.0567)	**−0.1608***	(0.0565)
Z_1 rural	**−0.0364**	(0.7191)	**−0.0438**	(0.6629)	**−0.0192**	(0.8453)	**0.0368**	(0.7216)
Z_2 urban	**−0.0899**	(0.4653)	**−0.1300**	(0.3011)	**−0.1851**	(0.1473)	**−0.1953**	(0.1283)
Log number of staff/physician β_4 (include physician)	**0.1505**	(0.2054)	**0.2842***	(0.0650)	**0.3138****	(0.0249)	**0.3505****	(0.0314)
Log number of staff types β_5 (include physician)			**−0.2025**	(0.1689)				
Number of types greater than two (include physician) β_6					**−0.3020****	(0.0332)		
τ_1 number of types = 1							**0.3819**	(0.1360)
τ_2 number of types = 2							**0.2610***	(0.0927)
τ_3 number of types = 3							**−0.1953**	(0.1033)
Observations	61		61		61		61	
R-square	0.12		0.15		0.19		0.23	

Notes. Significance levels: * <0.1, ** <0.05, *** <0.01.

large negative, indicating that more types decreases income.

Across the four models, the effect of self-employment is consistent, stable, and negative. The negative sign on β_3 ($\beta_3 = -0.1697$, $p = 0.0531$) in Model 1 indicates that self-employed physicians earn less per hour worked than employed physicians. This makes sense because self-employed physicians are not compensated by a health care system that receives additional revenues through referrals. Their income is the residual profit from their practice only. Across the four models, we see that, although consistently positive, the coefficient on the number of staff is increasing as a better fit of the staff types is achieved. This shows the interaction between the two effects. Models 3 and 4 with the indicator variables for the number of types get the best results and show that adding staff increases income, whereas additional staff types leads to a reduction in income beyond one. This result is consistent with H2.

To summarize our empirical findings, we see (Table 5) that the addition of another staff member increases output, whereas an additional staff type leads to a net decrease in the productivity of the physician's practice. These results support H1, suggesting that increased division of labor reduces productivity. We do not rule out the possibility that there are specialization benefits from division of labor, but rather that in this context, those benefits are outweighed by the costs discussed in the paper. The results in Table 6 support H2, suggesting further that whatever reduction in labor rates accrue to the practice by greater division of labor is lost from a reduction in output. In our sample, practices with larger staffs tend to have more staff types. Therefore we would expect that practices with larger scale in terms of the visits seen per hour and the number of staff will not necessarily show larger incomes per physician hour worked than smaller scale practices. This conclusion is consistent with what we observed in Figures 1 and 2. We conclude, therefore, that many practices in our survey are in a situation in which the costs of division of labor have exceeded the benefits.

6. Conclusion

It is certainly well known that, as an organization gets larger, so do its coordination costs. However, in any particular organization, it may be the case that the impact of coordination costs is larger or smaller. At the same time, the source of these costs varies across organizations. Only by focusing on specific types of organizations and analyzing their economics and work processes can one make more specific statements about the impact of coordination costs and their effect on optimal staffing and task allocation. In this paper, we have focused on primary care physician office practices with the purpose of determining if coordination costs are important in choosing the optimal practice structure. We provide empirical evidence that physician office practices, with even very few staff members, experience a reduction in productivity related to division of labor. We do not have detailed data on how the staff members spend their work time. Therefore we have no direct evidence that the reduction in productivity is from coordination activities. However, given the extensive operations literature on process design and our own observations of physician practices, we conjecture that coordination time is indeed the source of inefficiency in primary care practices. Further research using detailed process data could better help identify the precise sources of inefficiency.

Our results challenge the common advice given to physicians to delegate as much nonlicensed work to support staff as possible. We have shown that delegation can be inefficient in these settings. In fact, we show that many physicians could maintain their current income, while seeing fewer patients and managing fewer support staff. These results have potentially important implications for the quality of health care, access to health care, and the health care labor market. To the degree that more personalized attention improves care quality, smaller practices may lead to higher quality. Less delegation also means that we are substituting physician labor for support staff, implying a greater need for physicians and a reduced need for some categories of support staff. Our findings naturally lead to the question: Is there a way that delegation can be done more efficiently?

An area for further research would be to empirically identify the factors that make some practices more successful at operating with large numbers of support staff and high visit volumes than others. In particular, there is anecdotal evidence that some of the tasks performed by secretaries and nurses could be either eliminated or greatly reduced with the use of more IT. For example, a practice using electronic medical records would not need to have the charts pulled or replaced in the file cabinets. One can imagine a world, assuming security and privacy issues have been addressed, in which community-based information systems could transfer results of studies directly to the doctor's office and into the medical record. Thus, two additional avenues for further research are: (1) What percentage of the work could be reduced by IT and (2) What would be the impact of such a shift on the optimal organization and profitability of an office practice? Finally, the research approach taken here could be applied to other professional services such as law or accounting practices.

Electronic Companion

An electronic companion to this paper is available on the *Manufacturing & Service Operations Management* website (http://msom.pubs.informs.org/ecompanion.html).

Acknowledgments

The authors would like to thank Sanjog Misra for his guidance in the statistical analysis in this paper. The authors also thank the review team that provided constructive criticism that greatly improved this work.

Appendix

Zip code	Count	Classification R, U, S
14125	1	R
14411	3	R
14420	1	R
14423	1	R
14432	1	R
14476	1	R
14482	1	R
14510	1	R
14512	1	R
14589	1	R
14450	4	S
14502	1	S
14526	1	S
14580	4	S
14612	1	S
14617	2	S
14618	8	S
14619	2	S
14620	2	S
14622	1	S
14623	6	S
14624	4	S
14626	3	S
14642	1	S
14607	3	U
14608	1	U
14609	1	U
14615	1	U
14621	1	U
14625	2	U

References

Brown, D. M. 1988. Do physicians underutilize aides? *J. Human Resources* **13**(3) 342–355.

Buzacott, J. A. 2004. Modelling teams and workgroups in manufacturing. *Ann. Oper. Res.* **126** 215–230.

Frech, H. E., P. B. Ginsburg. 1974. Optimal scale in medical practice: A survivor analysis. *J. Bus.* **47**(1) 23–36.

Guglielmo, W. J. 2003. Physicians' earnings: Our exclusive survey. *Medical Econom.* **80**(18) 71.

Ittig, P. T. 1978. A model for planning ambulatory health services. *Management Sci.* **24**(10) 1001–1010.

Kilbridge, M., L. Wester. 1966. An economic model for the division of labor. *Management Sci.* **12**(6) B255–B269.

Kowalczyk, L. 2004. Rejecting health insurers doctor's decision reflects discontent within profession. *Boston Globe* (January 30) A.1.

Landon, B. E., J. Reschovsky, D. Blumenthal. 2003. Changes in career satisfaction among primary care and specialist physicians, 1997–2001. *J. Amer. Medical Assoc.* **289**(4) 442–449.

Marder, W. D., S. Zuckerman. 1985. Competition and medical groups: A survivor analysis. *J. Health Econom.* **4**(2) 167–176.

McKee, K. 2004. Ways to work smarter. *Medical Econom.* **81**(22) 36–38.

MGMA. 1999. *Performance and Practices of Successful Medical Groups—Survey 1999*. Medical Group Management Association, Englewood, CO.

Misra, S., E. Pinker, R. Shumsky. 2004. Sales force design with experience-based learning. *IIE Trans.* **36**(10) 941–952.

Moore, G. 2002a. Going solo: Making the leap. *Family Practice Management* **9**(2) 29–32.

Moore, G. 2002b. Going solo: One doc, one room, one year later. *Family Practice Management* **9**(3) 25–29.

Pinker E., R. Shumsky. 2000. The efficiency-quality trade-off of cross-trained workers. *Manufacturing Service Oper. Management* **2**(1) 32–48.

Pope, G. C., R. T. Burge. 1996. Economies of scale in physician practice. *Medical Care Res. Rev.* **53**(4) 417–440.

Powell, S. G. 2000. Specialization, teamwork, and production efficiency. *Internat. J. Production Econom.* **67** 205–218.

Reinhardt, U. 1972. A production function for physician services. *Rev. Econom. Statist.* **54** 55–66.

Reinhardt, U. 1975. *Physician Productivity and the Demand for Health Manpower*. Ballinger Publishing Co., Cambridge, MA.

Reinhardt, U., M. V. Pauly, P. Held. 1979. Analysis of economic performance in medical group practice. Project Report 79-05, Mathematica Policy Research, Princeton, NJ.

Rummel, J., Z. Walter, R. Dewan, A. Seidmann. 2005. Activity consolidation to improve responsiveness. *Eur. J. Oper. Res.* **161** 683–703.

Schneider, D. P., K. Kilpatrick. 1975. An optimum manpower utilization model for health maintenance organizations. *Oper. Res.* **23**(5) 869–889.

Seidmann, A., A. Sundararajan. 1997a. The effect of asymmetry on business process design. *Internat. J. Production Econom.* **50** 117–128.

Seidmann, A., A. Sundararajan. 1997b. Strategic and competitive implications of business process redesign: Analyzing the impact of information systems and organizational design. *Proc. 30th Hawaii Internat. Conf. Systems Sci.*, Cleveland.

Seidmann, A., A. Sundararajan. 1997c. Competing in information-intensive services: Analyzing the impact of task consolidation and employee empowerment. *J. Management Inform. Systems* **14**(2) 33–56.

Smith, A. 1776. The wealth of nations Books 1–3. *On Division of Labor*, Book I, Chapter I. Penguin Classics, New York.

Swisher, J. R., S. H. Jacobson. 2002. Evaluating the design of a family practice healthcare clinic using discrete-event simulation. *Health Care Management Sci.* **5** 75–88.

Terry, K. 2004. Exclusive survey: How hard are you working? *Medical Econom.* **81**(19) 30.

Thornton, J. 1998. Do physicians employ aides efficiently?: Some new evidence on solo practitioners. *J. Econom. Finance* **22**(2–3) 85–96.