1991

Productivity in a multispecialty medical clinic

Stephen Fraser Johnson
The University of Montana

PRODUCTIVITY IN A MULTISPECIALTY MEDICAL CLINIC

by

Stephen Fraser Johnson

B. A., Cornell University, 1967

M. D., University of Michigan, 1971

Presented in partial fulfillment of the requirements

for the degree of

Master of Business Administration

University of Montana

1991

Approved by

Chairman, Board of Examiners

Dean, Graduate School

June 3, 1991

Date

ACKNOWLEDGMENTS

Selected grapes, when crushed, filtered, fermented, and aged, yield fine wine. The winemaker gets the credit. During five years as a part time student in the University of Montana MBA program, the author has done the crushing, filtering, fermenting of ideas for this professional paper. However, other individuals deserve thanks for help and support with the project. In particular, my colleagues, every physician, surgeon, psychologist, and podiatrist at Western Montana Clinic, took time to answer detailed questions about attitudes and activities. Administrator Gary Larson educated me about the subtleties of production in a multispecialty medical clinic. Joe Pritting of the WMC accounting department implemented the "Professional Services: Productivity" software; he helped insure that the production data were reliable.

My professional paper committee, Paul Larson, Ph.D., Associate Professor of Management, Chairman; Rudyard Goode, Ph.D., Professor of Finance; and Richard Withycombe, Ph.D., Professor of Management, guided my efforts at strategic times. My friend, Lin, reviewed office schedules, listened to discussions of statistics, and kept smiling while I was complaining. Jerry Hintze of NCSS helped me drive around statistical roadblocks. Thomas R. McCarthy, Ph.D., the Vice President of National Economic Research Associates, Inc., mailed me his dissertation to read. Robert Harriman of the Pensacola Clinic and John Milor and Judy Brown of the Wichita Falls Clinic were extremely helpful in providing data for external validation.

Finally, those researchers cited repeatedly, especially Reinhardt, Fuchs, Baumol, and Pope, have my admiration for their previous hard work and insights. I have cited their contributions and those of other researchers using the format of the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" as described in the New England Journal of Medicine Special Report, volume 324, pages 424-8, on 7 February 1991. For other questions of style, I depended on the Chicago Manual of Style, published by the University of Chicago Press in 1982.

Figure II-1 was reproduced from the book, Beyond Boredom and Anxiety, by Mihaly Csikszentmihalyi with the permission of Jossey-Bass Publishers of San Francisco.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I. INTRODUCTION

GENERAL BACKGROUND

Health care services constituted 12.2% of the total economic output of the United States and 15% of its service sector in 1990; they are growing faster than the rest of the economy.[1][2] Scientific advances and technologic innovation in medicine now mean effective treatment for many diseases which, in the recent past, were fatal. However, it is difficult to find a current newspaper or magazine which does not mention the current "crisis" in health care.

Manifestations of the health care crisis include 1) rising costs to individuals, businesses, and government, 2) insufficient medical services for some segments of society, and 3) public health statistics such as life expectancy and infant mortality, which lag behind those in other industrial countries. The basic issue is one of productivity: although the input of resources into health care has increased, many perceive that the resultant output of health is inadequate.

A recent study by Fuchs and colleagues highlights the medical productivity problem at the national level. Judging from published statistics, the health of Canadians is the same or better than citizens of the United States. However, physicians in the United States generate 30% more revenue but see 20 percent fewer patients and do 20 percent fewer procedures than physicians in Canada. American physicians have substantially higher overhead costs as well. However, for the whole medical care sector, "if U.S. spending could be held to the Canadian percentage, the savings would amount to more than $100 billion a year."[3]

Health is difficult to measure.[4][5] Provision of health care services is only one of the inputs for creation of good health. Grossman, for example, argues that health is an asset which depreciates with age and neglect. Health can be increased by investment of time, education, diet, exercise, housing, and avoidance of known health risks, as well as by obtaining medical care.[6] Other independent economic analyses indicate that the marginal contribution of health care services to health is less than the marginal contribution of technologic innovation, years of education, and life style changes.[7][8][9] Nevertheless, input of health services does correlate positively with health.

Improved productivity in health services has important consequences for the economy in addition to improved health. Health care productivity gains can spare some resources for other important uses, for example, crime prevention or road building. Furthermore, a healthy population will be more productive than a sicker one. Therefore, simply looking at health services as an output may underestimate the effect of health care sector productivity on the overall standard of living.

Several observations raise questions of inefficient delivery of health care services and suggest that health care productivity is ripe for improvement. One is the very size of the health care sector and its rapid growth. A second is the rapid growth of applicable technology. Third, there has been highly individual and idiosyncratic application of health care resources in the past. If health care sector productivity can be enhanced, the effort should yield notable economic benefits for the United States.

Physicians occupy a key position regarding both input and output of health care services. They provide many services directly and they influence strongly which items patients receive from an array of diagnostic and therapeutic choices, ranging from blood pressure determinations to magnetic resonance imaging and from reassurance therapy to heart surgery. Better management of medical resources by physicians should result in enhanced outputs: prevention of some diseases, amelioration of others, and better health and function of the population. Therefore, both directly and indirectly, improved physician productivity should increase productivity in the health care sector and promote health.

Various pressures and certain advantages have drawn physicians to corporations or groups. The pressures include labyrinthine billing requirements of insurance companies and government agencies, high capital investment to open an office, and ignorance about techniques of office management. Advantages of a group practice include convenience of a common location for patients and physicians, greater stability of professional income in a group, and psychological support from other members of a group. Many physicians now practice in groups of substantial size.[10] However, the evidence about the effect of group practice on physician productivity is conflicting.

There is a clear need for short- and long-term increases in medical group practice productivity, because of increasing limitations on third-party (insurance carrier and government) reimbursements and increasing competitiveness in the medical sector. Locally, the Western Montana Clinic (WMC), a medical group practice, confronts the prospect of declining revenues to carry out its medical and business activities. Government reimbursement for Medicare and Medicaid services is scheduled to diminish. More patients must make large coinsurance payments, and as many as 20% have no health insurance at all. They will demand better service for their dollar or may avoid medical care entirely. With less revenue per patient, WMC physicians will need to see patients more efficiently, that is, increase their productivity, to meet people's medical needs and maintain clinic revenues.

STATEMENT OF THE PROBLEM

There is a lack of detailed understanding of the factors which influence health sector productivity in general and physician productivity in particular. Efficient use of existing resources and the application of new technology can augment service productivity in non-medical enterprises, so these factors may impact physician output as well. Nevertheless, for several reasons, augmenting physician productivity may be difficult. First, management of any group of professionals requires special skills.[11] Second, the determinants of physician productivity and their malleability are not completely known. Third, some technologies act as economic complements rather than substitutes for physician services. Fourth, patients and families require some irreducible amount of personal contact with the physician. Hence, some economists argue that dramatic productivity increases in medicine are impossible.[12] The problem, then, is to define which input factors are important determinants of physician services in order to improve physician output.

THE RESEARCH QUESTIONS AND THEIR IMPORTANCE

What is the production function for physicians at the Western Montana Clinic? That is the research question. A production function is a quantitative expression of the output of goods or services which can be obtained from different levels of various input factors. The question can be rephrased: are there any statistically and practically significant factors which correlate with the output of physician services?

A production function can be a guide for increasing productivity at WMC. Knowledge of this function will allow management to select, augment, or diminish controllable input factors to achieve more efficient output of medical services. In so far as the Western Montana Clinic and its management are representative of other medical clinics, the information gleaned from this study may help increase health services productivity elsewhere in the United States.

THE HYPOTHESES AND EXPECTED RESULTS

The hypothesis is that physician productivity should correlate strongly with several factors, some positively and others negatively. For example, according to the most recent comprehensive study, physician output should correlate positively with hours worked, medical specialty, years of practice experience, use of medical and administrative assistants, and use of capital equipment.[13] According to the same study, production should correlate negatively with outside income and female gender, and there should be a curvilinear correlation of production with age. Production should also increase with a positive attitude about work, number of dependents, personal debt, good health, and exercise, but there are no studies reporting such correlations for physicians. There could be a positive, negative, or curvilinear correlation between output and hobbies and recreational activities. The null hypothesis is that there are no factors which correlate significantly with physician output of medical services.

CHAPTER II. REVIEW OF PHYSICIAN PRODUCTIVITY AND MANAGEMENT LITERATURE

DEFINITION

Productivity is the ratio of outputs to inputs. Output of "health" is conceptually difficult to measure, but services provided by physicians in a medical clinic can be measured more easily.[14] Inputs for health services can be quantified. Measurement of outputs and inputs will be discussed at length below in the sections on the dependent and independent variables.

ORGANIZATIONAL INFLUENCES ON PHYSICIAN PRODUCTIVITY

Large group practices might expect economies of scale which will improve productivity, but careful economic analysis has yet to show that such economies of scale occur. Because of shirking, non-price competition within the group, and disincentives in compensation plans, multispecialty group practices exceeding eight members seem to be less productive than groups of smaller size.[15][16][17][18][19][20][21]

The method of payment for health care, for example fee-for-service versus a prepaid source of revenue, can change productivity, but the effects are complex. In a controlled study, total costs to patients in a prepaid, capitated health plan (health maintenance organization or HMO) were substantially lower than for fee-for-service insurance coverage.[22] If health outcomes were the same under capitation as under fee-for-service, the cost saving would indicate improved productivity in HMOs. However, although the average enrollee did as well under either option, low-income people who already had health problems fared worse under the prepaid plan.[23] These poor, sick patients are exactly the ones needing the most help, so the results do not augur well for current HMO patients as they grow older and sicker. The influence of method of payment on health services productivity remains unclear.

Organizational attributes other than size and method of payment also affect productivity. Such attributes include facility design, marketing, staffing, scheduling, and the corporate culture.[24][25] Local factors such as the physician to population ratio, per capita income, and extent of medical insurance coverage also impact physician productivity.

PHYSICIAN CHARACTERISTICS AFFECTING PRODUCTIVITY

Some physician characteristics correlate with productivity in published economic studies. They include time spent in seeing patients, specialty, delegation of tasks to assistants, age, experience, gender, need for money or leisure, and "physician tastes."[26][27][28] That time, specialty, delegation, age, and experience effect output of medical services seems obvious. How gender may impact physician productivity is a matter of some debate. Apparent gender effects may really be manifestations of work setting, experience, and specialty choice.[29] Behavioral researchers as well as managers have studied extensively the effects of rewards on productivity. The correlation is positive, but the mechanisms and magnitudes of the effects can be difficult to define. The collective term, "physician tastes," is economic jargon for personal factors which influence effort and output. Physician tastes have received only minor attention in prior econometric studies.

The most detailed survey of the effect of rewards on productivity in group practice is the study by Held and Reinhardt. They showed that physicians tend to pick out clinics which suit their own reward preferences. However, when this preselection bias is taken into account, a change from fixed salary to straight incentive based on output impels the physician to increase hourly patient load by 30% and results in a similar increase in the amount of surgery done.[30]

Health economists have devoted a good deal of attention to the "target income hypothesis" of physician financial behavior. The proponents of this hypothesis posit that physicians work to gain a certain level of income and that they create demand for services until they reach that level.[31] Since physicians are most often paid on an incentive basis, setting the "piecework" rate higher or lower might well influence productivity. However, the rate of payment for any medical service is reduced by office expenses, so the impact on willingness to work is less direct than the change in hourly wage for a factory worker.

Simon's book, Effort, Opportunity, and Wealth focuses on the economic determinants of individual effort with the "drive-effort hypothesis." He maintains that the strength of response to an opportunity varies directly with the rewards per hour of work and inversely with a person's accumulated wealth. From this relationship, he explains the "income effect," in which leisure time becomes more important than money at a certain level of income. He illustrates the "income effect" with the example of a physician. He also cites empirical evidence about the effects of wealth and number of dependents on effort put forth.[32]

Non-financial rewards are very important to professionals, including physicians.[33][34] Professionals rate advancement, challenge, and new skills right up with pay and benefits. Physicians bask in appreciation from patients. Professional values include defiance of authority and concern about social justice and quality of life. Overall, the association between output and rewards is positive, but the mechanisms are complex and the magnitude of the effect of rewards on physicians is not clear.

Robert H. Frank, an economist, expounds on non-verbal signals which modify behavior in ways which most economists simply skip over.[35] He considers the effects of reputation on economic behavior. Reputation, appearance, and other personal factors seem very likely to influence a physician's output.

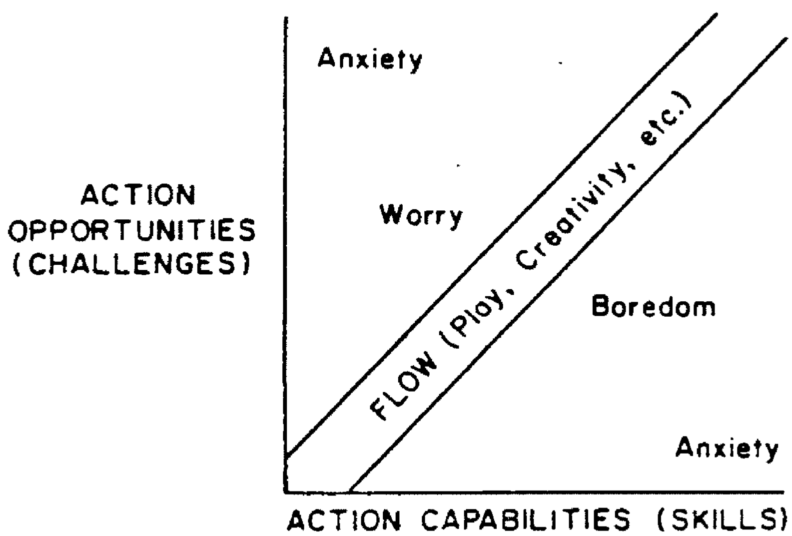
At this point, the economic literature on individual factors in productivity wanes. Information from other academic disciplines may nevertheless be of relevance. Maslow's hierarchy suggests that both self-esteem and self-actualization are central to physician productivity.[36] Enthusiasm and satisfaction relate to productivity, but most probably as consequences of high productivity rather than their cause, according to organizational theorist Charles Perrow.[37] Csikszentmihalyi includes a chapter on enjoyment of the work of surgery in his insightful book on motivation, Beyond Boredom and Anxiety. He argues that skill and challenge must be balanced to achieve a state of "flow," which results in pleasure and enthusiastic output. Figure II-1, taken from his book, summarizes his ideas about flow. Surgery, like certain other engrossing pursuits he studied, namely the game of chess and the activity of rock climbing, can result in a high flow state.[38] Presumably this applies to activities of physicians as well as surgeons.

Conversely, overwork and burnout may result in low productivity. The words of Robert Larranaga epitomize many

FIGURE II-1. EXPLANATION OF THE FLOW STATE

This figure, taken from Csikszentmihalyi's book [38] and reprinted here with permission of the publisher, was shown to each physician to explain what is meant by "flow."

Model of the Flow State. When a person believes that his action opportunities are too demanding for his capabilities, the resulting stress is experienced as anxiety; when the ratio of capabilities is higher, but the challenges are still too demanding for his skills, the experience is worry. The state of flow is felt when opportunities for action are in balance with the actor's skills; the experience is then autotelic. When skills are greater than opportunities for using them, the state of boredom results; this state again fades into anxiety when the ratio becomes too large.



physicians:

> ...high-energy people with a strong need to dominate and control. Impulsive and time sensitive, we can be demanding, aggressive, and quick tempered....We always do more than the situation requires....
>
> But no one can continue at the pace we set for ourselves. Eventually the fun goes out of work , and we become grimly determined to win at all costs....
>
> By this point, working for the sake of work consumes every available hour. It disrupts our home life, affects our health, and leaves us feeling burned out....Eventually we become cynical, bitter, and despondent.

although he is describing "workaholics" in general.[39]

PREVIOUS STUDIES OF PHYSICIAN PRODUCTION FUNCTIONS

Previous studies of physician production are listed in Table II-1. Uwe Reinhardt's publications on physician production provide a window to earlier literature and remain the benchmark for studies of physician output.[40][41][42] Lorant and Kimbell explain much more of the variation in production than other investigators, perhaps because of better quality data.[43] In his dissertation, McCarthy develops a production function for physicians in chapter 5, but he has not published his conclusions in more readily accessible form.[44] Paul Feldstein's textbook outlines the pitfalls of physician production measurements and reviews conclusions about physician output through 1988.[45] Most recently, Gaynor has published two articles containing physician production functions.[46][47]

TABLE II-1. PHYSICIAN PRODUCTION - REPORTED RESULTS[1]

AUTHOR/YEAR	DATA SOURCE	OUTPUT MEASURE	EXPLAINED VARIANCE
Reinhardt 1975	*Medical Economics* survey of solo doctors	Patient visits and billings	.42 -.69
Lorant and Kimbell 1976	AMA 1971 physician survey - solo and group	Gross revenues and patient visits	.43 -.89
Held and Reinhardt 1979	Mathemetica group practice survey	Patient visits, surgical workload, time	.40
McCarthy 1980	AMA Survey of solo physicians, 1976	Office visits in California RV units	.40 -.45
Gaynor 1989	Mathematica group practice survey	Patient visits	.42
Hurdle and Pope 1989	HCFA-NORC 1985 survey of solo primary care physicians	Office visits, gross charges, time	.28 -.50
Gaynor and Pauly 1990	Mathematica group survey	Patient visits	.40 -.44

---

1 The studies cited, which describe the data sources, are listed in the Reference section of this paper. AMA is the American Medical Association, located in Chicago. Mathematica Policy Research is based in Princeton, NJ. HCFA is the Health Care Financing Administration of the United States Government.

What are the most important factors in physician productivity reported in previous studies? They are listed in Table II-2. Physician time input must top the list. A 10% increase in time spent in patient care increases output from 2.5 to 7.0%, depending on whether the time is total time, office time, or some more refined measure and also depending on whether output is office visits or gross charges. In addition, there is variability depending on practice characteristics.

The importance of office personnel, number of offices (space), and capital equipment investment also is considerable. As with time, the reported elasticities vary because of the exact definition of the factor and the practice setting.

Prior studies of physician production are based on rough measures of physician output and highly aggregated data. Reinhardt has recognized the limitations of aggregate data.[48] Whether a study based on more precise data will yield new insights into physician production remains to be determined.

AUGMENTING PRODUCTIVITY

All of the previously published studies of physician production functions are cross-sectional observations of inputs and outputs. Alone, they cannot establish causative relationships among the variables. However, when cross-

TABLE II-2. FACTORS IN PHYSICIAN PRODUCTION - RESULTS OF PREVIOUS STUDIES[1]

FACTOR	PRODUCTION ELASTICITY	DATA SOURCE
Time	.56 to .69	Reinhardt, 1975
	.52 to .58	Held, 1979
	.31 to .65	McCarthy, 1980
	.25 to .66	Hurdle and Pope, 1989
Office personnel	.31 to .34	Reinhardt, 1975
	.42 to .59	McCarthy, 1980
	.09 to .27	Hurdle and Pope, 1989
Space	.15 to .17	Held, 1979
	.06 to .14	Hurdle and Pope, 1989
Capital equipment	.04 to .14	Reinhardt, 1975

---

1. The data listed are limited for several reasons. The investigators listed in Tables II-1 and II-2 often explored only a few independent variables. Definitions of input and output differed among studies. Often the authors did not calculate production elasticities of the input factors and did not supply sufficient data to allow others to do so. Thus, although significant correlations exist between many different input factors and the output of physicians' services, it is difficult to judge their relative importance.

sectional observations are combined with historical data (time series) and managers' practical knowledge about how changes in input factors affect production, the blend of approaches does allow determination of cause and effect.

The major conclusions from studies of the service sector are that application of new technology and efficient use of existing resources may augment productivity.[49][50] However, the effects of new technology on health outcomes and health services are by no means simple. Technology can improve health care productivity, but new technology also allows the provision of new services. New services may simply be added on top of older techniques rather than supplanting them. If the new techniques do not provide additional real output, productivity will drop. If new techniques do augment health care but if the output measures do not capture the additional output, productivity will appear to drop as judged by the crude measures. Reinhardt considered these difficulties in his extensive studies on physician productivity. He found that capital inputs often complement rather than substitute for physician time, so costs are increased and measurement of productivity is confounded.[51]

Could prospective technology assessment improve output of medical services without raising costs as much as the current unregulated approach and without squelching innovation? Is there a mechanism to identify technology which improves health service productivity, but avoid technology which is harmful or wasteful? Fuchs and Garber have assessed the prospects and some of the difficulties of assessment of technology before implementation.[52] However, an empirical study of technology assessment in Ontario points out that it is not working well even in a government controlled environment.[53] More serious yet, another article points out how the Health Care Financing Administration (HCFA) in the United States uses its reimbursement scheme to ration even well-tested new technology.[54]

In contrast, more efficient use of existing technology and resources of time, office space, non-physician labor, and capital equipment must improve health care productivity. Reinhardt concluded that substantial productivity gains for physicians can occur through supervised delegation of tasks to less highly trained individuals.[55] A review published in 1989 brings the use of "physician extenders" up to date and emphasizes that "unproductive physicians in all specialties used fewer aides than productive physicians."[56] Nevertheless, delegation of tasks by physicians remains limited.[57][58] Luft points out "implementation problems related to physician acceptance and role definition" of assistants in health maintenance organizations. He indicates that the delegation method of increasing productivity does involve management challenges.[59]

MANAGEMENT OF PHYSICIANS

Attempts to increase output by more efficient use of time, office space, and medical equipment involve management challenges, too. Controlled trials reported in medical journals indicate that feedback and education can change physician behavior, but unless such intervention is persistent, changes in behavior are not sustained.[60][61][62][63][64] However, the vast majority of sources on physician management are field observations or anecdotal reports, rather than controlled trials.[65][66] Many draw heavily on experience with other professionals or other industries.[67][68] There are no controlled trials of management interventions to enhance physician productivity.

A bibliography compiled by the Medical Group Management Association (MGMA) lists about 350 articles on improving office and professional productivity.[69][70] Culling such articles is a daunting task, but there remain a few which offer solid if unvalidated advice. Some of the better suggestions for augmenting physician productivity on an individual level involve a practice review,[71] delegation of tasks,[72][73][74][75] scheduling,[76][77][78] telephone use,[79][80] billing appropriately for time spent,[81][82] patient information and education,[83] and time management.[84][85][86]

Not surprisingly, financial incentives do alter physician behavior. For example, incentives can induce physicians to order more laboratory tests and X-rays, which will increase clinic revenues and productivity.[87] Held and Reinhardt's study of medical group practice indicates that financial incentives do increase the physician's hourly patient load and the number of patients seen, but not the total hours worked per month.[88] But an important management question is what incentives to provide for increased efficiency without sacrificing quality.[89][90][91] Considered in isolation, incentives for specific behaviors can be counterproductive. However, a broad view of productivity and quality indicates that both can improve simultaneously.

PRODUCTIVITY-QUALITY INTERACTIONS

There is a longstanding belief that too much attention to quality will lower productivity. This appears to arise from experience with quality control inspections in industry. Inspections take time and they cost money. If standards are set too high, the reject rate will rise, rework will increase, and output will diminish. Donabedian and colleagues refer to this as a "quality/cost tradeoff in medical care."[92]

According to Garvin, who has studied quality extensively, the idea of a quality/productivity tradeoff is a misconception which results from a short term perspective and flawed production measurements. Recent, more insightful analyses in manufacturing and in medicine indicate that productivity and quality are not antagonists. They change in tandem. Garvin summarizes the productivity-quality relationships in manufacturing in his book on quality:

> The argument for a positive correlation between quality and productivity is usually stated in the simplest of terms. Less rework means more time devoted to manufacturing acceptable products, and less scrap means fewer wasted materials....
>
> A better understanding of the connection between quality and productivity requires an examination of their common sources of improvement. First, however, productivity must be more carefully defined. Most analysts implicitly associate it with partial measures. Either labor productivity (output per employee or output per labor hour) or materials productivity (output per pound of input or output per dollar of material employed) is normally reported. But more comprehensive measures, reflecting total factory productivity, are also available. They track changes in output according to changes in a combination of inputs -- usually labor, material, capital, and energy -- rather than matching output to changes in a single input alone.
>
> Using these definitions, quality and productivity improvement can be traced to similar roots. For example, standardized parts and modular design simplify the assembly process, reduce opportunities for errors, involve easier-to-stock parts and materials, imply less time devoted to disruptive engineering change orders, and often require less labor for manufacturing and rework. Improved equipment and better maintenance mean less downtime, fewer machine-related errors, and less excess capacity held in anticipation of breakdowns. A more stable and trained work force means that jobs will be performed more efficiently and inadvertent mistakes will be less frequent....[93]

It seems clear that Garvin's ideas may have even more relevance in medicine, where a broad view of production and quality are essential, than in manufacturing.

Simplification of the process of medical care seems likely to improve quality, productivity, and satisfaction simultaneously. An example comes from the Framingham Heart

Study.[94] There has been a well documented 40% age-adjusted decline in mortality from heart disease and stroke in the United States since 1960. Experts debate exactly how much of the decline is due to high-technology medical interventions, e.g., thrombolytic and anti-arrhythmic drugs or heart surgery. However, at least 60% of the decline in cardiovascular mortality appears to be due to treatment of risk factors and to lifestyle changes. In this setting, which is a paradigm for other areas of medicine, attention to complication rates of open heart surgery, a traditional quality control approach, would have missed the point. Standard quality control would have a minor impact on productivity or health outcomes. Costs for surgery might even increase. In contrast, undramatic, low cost interventions which physicians frequently neglect have a substantial effect on mortality after a 20-30 year lag time.[95] More attention to risk factors improves both the quality of health (less mortality) and health service productivity (better outcome with less expensive preventive medicine input) simultaneously.

Currently, because of examples such as the one just cited, there is a great deal of enthusiasm about continuous quality improvement in health services.[96][97][98] However, empirical support for the idea that quality and productivity in medicine can improve at the same time is still meager.

OPTIMAL PHYSICIAN OUTPUT

Improving physician productivity is a realistic challenge. Optimization of physician output in economic terms is not feasible for several reasons. First, optimization must adopt one of several incompatible points of view. Government and third-party payers want low costs; they have little interest in individual amenities. Patients want personalized health care services to be available at any time. Physicians focus on high technology services for acute illnesses, but would like the illnesses to occur during regular office hours. As Reinhardt points out, the customs and expectations of physicians and patients probably would not allow consistent output at optimum levels under any circumstances.[99]

Second, as Simon emphasizes, the economist's focus on optimization requires perfect information at no cost.[100] In the health sector, there are severe constraints on information flows due to inherent complexity of the subject, ignorance and denial of consumers, and monopolization by health care providers and government agencies. More than in the business world, "satisficing" and other methods of non-optimizing decision making prevail in the health market.

Third, even if all medical services were optimum from the several vantage points, still there are many technical approaches to diagnosis and therapy of any individual's medical problems. Medical science has not defined the best therapy for most diseases. Even if the best therapy were unequivocal, the substitutability, complementarity, and constraints on input factors are not nearly well known enough to allow mathematical programming to determine optimum production.[101]

CHAPTER III. METHODOLOGY

In its most general form, a production function for physicians appears as follows:

{1} $$Q_i = f(T, L, K, P, C, U)$$

Expressing production in a mathematical framework allows a methodical analysis of all the factors involved. This may avoid subsequent confounding of conclusions. The independent variables are the inputs to the physician production function. The dependent variable, Q_i, is an individual physician's output of services. Summing Q_i for all of the WMC physicians, designated 1 to n, yields the production for the entire clinic for a period of time:

{2} $$Q_{WMC} = \sum_{1}^{n} Q_i.$$

THE INDEPENDENT VARIABLES AND THEIR MEASUREMENT

There are many possible independent variables or input factors which can determine a physician's production. Divide them into six major groups: "T" is physician time input in hours; "L" is non-physician labor time inputs; "K" is capital inputs; "P" is physician characteristics; "C" are clinic attributes, some of which are fixed in this single-clinic study; and "U" are undefined variables.

Information about the input factors in the equation were obtained from several sources, including the WMC management information system; structured, direct interviews with physicians; and interviews of hospital administrative staff. Consider first physician time input, T. Even though it is an extremely important variable in physician productivity, time input is difficult to quantify. The task is easiest for the urgent care physicians, who have standard office hours, defined vacations, no special night or weekend call, no hospital rounds, and no surgery. Other clinic physicians have complex, variable, and unique schedules. They may work at any hour of day or night. They do not punch a timecard nor otherwise log in and out. They have varying amounts of "on-call" nights and weekends. They may see patients in the office, at nursing homes, or in several different hospitals. For specialists such as cardiologists, gastroenterologists, and surgeons, significant output comes from procedures, many of which are done in the hospital rather than the office. Procedures frequently are cancelled or rescheduled; some are done as emergencies. Patients frequently miss or cancel appointments; others just show up and are worked in to the schedule.

Physicians generally overestimate the amount of time which they spend with patients. Even more than most people, they perceive themselves as "overworked and underpaid."[102] Hence, their own estimates of the time which they devote to patients may not be accurate. Short of following each physician with a stopwatch for many months, an exact accounting of time spent with patients is impossible.

Despite these obstacles, physician time input was estimated as follows. Since the number of working days and the total hours in a day are the same for everyone, time input was quantified by exception. The main WMC offices were open 252 days during 1990, as were the branch offices in Lolo and Polson. At these sites, a maximum of 9 hours of office appointments could be scheduled in a given day. Hours for the urgent care unit at the shopping center also are well-defined. Vacation, sabbatical, and meeting time, **ABSENT**, measured in working days off, should correlate negatively with productivity. Weekend days on call, **CALL**, should correlate positively with output. The fraction of available office time per day open to see patients, **APPT%**, was determined by reviewing physicians' schedules for a week in April and a week in October of 1990. Fractions of weekday time spent doing surgery was determined from operating room records. Fractions of time doing cardiac catheterization, endoscopy, and surgery were determined from scheduling logbooks in the various hospital departments. Fractions of time spent on hospital rounds or nursing home rounds were estimated by quizzing physicians directly. To estimate time spent delivering babies, the total number of deliveries per physician was determined from billing records and multiplied by the amount of time that physician estimated was needed for an average delivery.

A measure of the hours spent in patient care activities in an average month, **PATTIME**, was estimated by combining the individual time measures. To obtain PATTIME, ABSENT was subtracted from and CALL added to the total office hours available for the year. The result was multiplied by the sum of APPT% and the other fractions of daily time spent in various patient care activities. PATTIME neglects phone calls and patient visits outside of regular office hours and omits time spent assisting at surgery. These omissions are small in comparison to the total of time spent in patient care activities. Both APPT% and PATTIME should correlate positively with output.

Non-physician labor inputs, L, are closely related to the number of employees working directly with the physician. Salaries for such employees are based on skills and experience. The total pay for all employees working directly with the physician, divided by 100 for brevity, is designated **LABOR**. LABOR is a proxy for skilled help and should correlate directly with output.

Capital inputs, K, were measured several ways. One gauge is procedure orientation, **PROC**, which is a qualitative variable with a value of 1 for surgery, orthopedics, neurology, ENT, nephrology, gastroenterology, cardiology, urology, dermatology, obstetrics-gynecology, and other specialties with technical skills and 0 for general internal medicine, pediatrics, rheumatology, family practice, and other specialties which do few procedures. Procedures generally require expensive capital equipment investments, so PROC should correlate positively with output. **SPACE**, the number of offices and examination rooms occupied, is the usual measure of capital investment in large scale econometric studies of physician output. **EQUIP** is the replacement value of equipment used at least once each day, expressed in thousands of dollars. Equipment costs were estimated by the physicians themselves and checked with the purchasing agent as necessary. Costs of operating rooms, cardiac catheterization rooms, and dialysis rooms was obtained from local hospital administration sources. Details of the capital equipment estimates by specialty are presented in Table III-1. Like PROC, SPACE and EQUIP should correlate positively with output.

The physician characteristics, P, are most numerous. These include three differing measures of seniority: 1) **AGE** in years, which reportedly bears a curvilinear relationship to production; 2) experience, **EXP**, measured in total years in practice after completion of residency training; and 3) **EXPWMC**, years in practice at the Western Montana Clinic. The later two variables should have a curvilinear relation to

TABLE III-1. CAPITAL EQUIPMENT COSTS BY SPECIALTY

The capital equipment costs for every physician included the replacement cost of offices currently used. The cost for a standard 10 foot by 14 foot office was estimated by a local architect to be $72 per square foot plus $4000 in furnishings.

1. Psychologists: office furnishings and textbooks. Testing materials.

2. Family physicians, internists, pediatricians, rheumatologists, urgent care: laboratory chemistry analyzer @ $60,000 and standard X-ray unit @ $75,000.

3. Neurologists: EMG machine @ $6,000, EEG/Evoked Potential equipment @ $100,000, and CT scanner @ $530,000.

4. Obstetrician/Gynecologists: culposcope @ $3,000, ultrasound @ $80,000, and operating room @ $1.24 million.

5. Invasive cardiology: cardiac catheterization laboratory @ $2.2 million and treadmill @ $25,000 and EKG machine @ $2,500.

6. Non-invasive cardiology: EKG @ $2,500, Doeppler colored ultrasound at $200,000, treadmill @ $25,000

7. Pulmonary: bronchoscope @ $10,000, spirometry @ $7,000, oximetry @ $5,000, laboratory chemistry analyzer @ $60,000, and standard X-ray unit @ $75,000.

8. Dermatology: electrosurgery unit, operating table, lights, instruments, and microscope for total of $12,500.

9. Orthopedics: X-ray unit @ $75,000 and specially equiped surgery suite @ $ 1.39 million.

10. Podiatry: chair @ $3,000 and X-ray @ $75,000.

11. General surgery: operating room @ $1.24 million.

12. Gastroenterology: flexible sigmoidoscope @ $ 8,000, colonoscope and ERCP scope @ $12,000 each, endoscope @ $11,000, X-ray @ $75,000, and laboratory analyzer @ $60,000.

13. Nephrology: 9-bed dialysis unit @ $1,614,110, laboratory analyzer @ $60,000, and standard X-ray unit @ $75,000.

14. Urology: X-ray @ $75,000, laboratory analyzer @ $60,000, cystoscope @ $10,000, and operating room @ $1.24 million.

15. ENT surgery: microscope, camera, endoscopes, audiometry equipment, special instruments for a total of $48,000 plus operating room @ $1.24 million.

16. Oncology: X-ray unit @ $75,000, laboratory analyzer @ $60,000, and microscope @ $3,000.

productivity with a peak between 10 and 20 years.

Other physician characteristics include **GENDER**, an indicator variable with 1 for male and 0 for female. Previous reports indicate that male gender correlates positively with higher output. Two family characteristics, **MARITAL**, assigned a value of one for married and zero for single or divorced, and **DEPEND**, the number of dependents, should correlate positively with production.

Some financial measures may induce physician output. First, consider wages. A uniform fraction of the monthly pay for each professional staff member was used instead of actual pay for proprietary reasons. Total wages result predominantly from production of medical services, so any observed relationship with production is a tautology. The same argument applies to any attempt to measure a wage rate for non-salaried physicians.

If there is a working spouse or other independent income comparable to that earned by a working spouse, an indicator variable, **OUTSIDE**, will be used. OUTSIDE should vary inversely with production. **DEBT**, a qualitative variable with a value of 1 if the physician has an unpaid balance on student loans or mortgages exceeding $75,000, should correlate positively with output. **SAVINGS** is an indicator variable with a value of 1 if the physician has at least three months worth of salary saved in reserve. **PRIORITIES** is designed to measure the importance of medicine compared to other significant activities, such as raising children, recreation, and so forth, on a one to four scale, with one the highest priority. SAVINGS may correlate negatively with productivity; PRIORITIES on the scale described should correlated inversely with productivity.

HOBBIES, the number of hours per week spent on recreational or community activities, probably correlates with output in a curvilinear fashion. **OFFICE**, an indicator variable with a value of 1 if the physician has run her or his own solo practice, may or may not correlate with increased production.

FELLOW, a dummy variable with a value of 1 if a physician has attained the honor of fellowship in his or her specialty society, is a measure of professional reputation and motivation. **CME** is a dummy variable with a value of 1 if the physician has attended a national educational meeting in his or her specialty within the past year. **SELFRATE** is physician's own rating of skill in medical practice. SELFRATE may reflect self-esteem as well as skill and it may be inversely related to a tendency to undercharge for services rendered. "Undercharging" also occurs in primary care specialties in areas where competition for patients is intense.[103] A variable for undercharging, **UNDER**, is the ratio of intermediate office visits to total office visits for a given practitioner for 1990. Intermediate visits should be charged to patients who are more complicated or take longer to evaluate than usual. If a physician tends to undercharge, the ratio should be smaller than usual, since nearly all physicians might be expected to have the same number of more difficult patients. In some circumstances, e.g. psychologists and urgent care, the nature of the practice is such that this ratio is not meaningful, but for most other specialties a low ratio should correlate with low output. It might also correlate with time spent in the office, since "overtiming" is another way of looking at the problem of undercharging.

REPUTE is a 0 to 3 rating of the desirability of seeing the physician according to WMC employees. REPUTE was measured by a questionnaire given out at the beginning of two in-service education sessions for WMC employees run by the assistant administrator. Nursing staff, receptionists, and business office personnel filled out the questions. These individuals frequently see WMC physicians for their own medical care, because they receive a substantial discount. They also work with the physicians either directly or indirectly. For these reasons, they should have adequate knowledge to form an opinion about the reputations of WMC physicians. In scoring the questionnaire, which is shown in the appendix, the score of 1 for not enough information to form an opinion was not scored. The ratings of 0, 2, or 3 were averaged to give a reputation score for each physician. FELLOW, CME, SELFRATE, UNDER, and REPUTE should correlate positively with productivity.

BURNOUT on a 1 to 10 self-rating scale, should correlate negatively with production. **HEALTH** is a 1 to 10 self-rating of health status, 10 being perfect health, which should correlate positively with productivity, as should **EXERCISE**, the number of hours per week of aerobic exercise. **FATIGUE** on a 1 to 4 scale should be inversely proportional to productivity. However, **WORKAHOLISM**, a self rating of dedication to work on a 1 for "funhog" to 10 for "workaholic" scale and **FLOW**, a one to four self-rating scale for congruence of professional skills with problems for which patients are seen, ought to correlate positively with productivity. Two other ratios derived from output, **INTENSITY**, which is relative value units per patient, and **COMPLEXITY**, which is relative value units per procedure may be related to FLOW. They should correlate positively with output.

There are also a number of clinic attributes which need consideration. Some are fixed, because this is a study of a single institution with a uniform source of patients and only three associated hospitals. The invariant C factors include the demographics of WMC patients, the prepaid/fee-for-service mix, ownership status, degree of control over practice, "corporate culture," business office procedures, indirect employee inputs (e.g., administrative and janitorial services), eligibility for board certification, quality standards, and marketing activities. Some of these factors have been shown to correlate with physician productivity in studies of large samples of physicians in multiple practice settings. This study, however, can glean no insights about their importance because they are fixed for this one clinic study.

There is a single compensation plan for all of the professional staff at WMC. Hence, the effect of differing reward systems on production cannot be determined from this study, with one exception. **ANCILLARY**, a number proportionate to income derived from ownership of laboratory, radiology, and electrocardiographic departments, is a supplement to income. It might be predicted to increase proportionately to the amount of professional services rendered, but it might possibly be a disincentive to professional output, a manifestation of the so-called "income effect." **ANCILL%** is the percent of total income derived from ancillary services. Its relationship to production will be similar to **ANCILLARY**.

"U" factors constitute the unexplained variance in the data. By definition, they defy identification. The independent variables used in this study and their expected correlations with the dependent variable, productivity, are summarized in Table III-2.

THE DEPENDENT VARIABLE, PRODUCTION, AND ITS MEASUREMENT

LIMITATIONS OF PRODUCTION MEASURES. Productivity is a single concept, but its definition must be adapted for various

TABLE III-2. THE INDEPENDENT VARIABLES

VARIABLE	EXPLANATION	EXPECTED CORRELLATION
ABSENT	Vacation, sabbatical, and meeting days off	Negative
CALL	Days on call	Positive
APPT%	Fraction of office hours actually available to see patients	Positive
PATTIME	Total time avail-able to see outpatients on a regular day	Positive
LABOR	Salaries for direct non-physician help	Positive
PROC	Dummy with 1 = procedure-intensive specialty	Positive
SPACE	Rooms used	Positive
EQUIP	Replacement value of equipment used daily	Positive
AGE	Age	Curvilinear
GENDER	Male = 1	Uncertain
EXP	Years in medical practice	Curvilinear
MARITAL	Married = 1	Uncertain
DEPEND	Dependents (IRS)	Positive

TABLE III-2. THE INDEPENDENT VARIABLES

VARIABLE	EXPLANATION	EXPECTED CORRELLATION
OUTSIDE	Non-medical income	Negative
DEBT	Debts exceeding $75,000	Positive
SAVINGS	Three months income in the bank = 1	Negative
PRIORITIES	Where does medicine fit into life on 1 (top) to 4 scale	Negative
HOBBIES	Hours per week	Curvilinear
OFFICE	Previous solo practice	Uncertain
FELLOW	Honored by specialty group	Uncertain
CME	Continuing medical education	Positive
SELFRATE	1 to 10 self-rating of medical skills	Positive
UNDER	Ratio of inter-mediate to total office visits	Positive
REPUTE	Outside rating of bedside manner	Positive
BURNOUT	1 to 10 self-rating scale	Negative
HEALTH	1 to 10 self-rating of health	Positive
EXERCISE	Hours weekly aerobic exercise	Positive

TABLE III-2. THE INDEPENDENT VARIABLES

VARIABLE	EXPLANATION	EXPECTED CORRELLATION
FATIGUE	1 to 4 self-rating	Negative
WORKAHOLISM	1 = funhog to 10 for workaholic	Positive
INTENSITY	PSP Intensity	Positive
COMPLEXITY	PSP complexity	Positive
FLOW	1 to 4 self-rating	Positive
ANCILLARY	Income from laboratory, radiology, or EKG services	Positive
ANCILL%	Percentage of income from laboratory, radiology, or EKG services	Negative

applications.[104][105] Basically, productivity is the ratio of outputs to inputs. To economists, a traditional measure of productivity is a specific output, for example, tons of steel divided by a specific input, for example, labor hours.[106] However, in even the simplest management application, there are multiple inputs, multiple outputs, and many ways to measure them.

When assessing physician productivity, what output should be measured? Health care services, which in this paper can also be labelled medical services, are a better choice than measuring the actual health of patients. As already mentioned, health is difficult to quantify. Furthermore, medical services are not the only input for good health. Therefore, if health were to be used as a measure of physician output, the other input factors, namely lifestyle changes, income, and education, would have to be kept constant. That is impractical. Better focus on a direct measure of physician output and recognize that providing medical care does influence health.[107]

Previous studies of physician productivity have used patient visits, charges generated, or revenues as measures of output. Each of these measures is subject to substantial limitations.

Patient visits, the dependent variable **PATIENT** in this study, provide a useful index of production when comparing practitioners in a single specialty, but a tally of visits is inadequate for assessing productivity in a multispecialty clinic or on any regional basis. A simple tally of patients seen ignores the quality and intensity of the service performed, which could range from refilling a prescription to open-heart surgery. Even defining when to count a patient visit is difficult. If a patient has multiple outpatient encounters in a day or week for a single or for multiple problems, some visits to one physician, some to another, yet others for laboratory or radiology procedures, should the tally coincide with each trip in or out the door? Or should each procedure or each physician visit or each day with a procedure or visit be tallied? See the sections on "PSP" and on "Preliminary Data Analysis" which follow for further discussion of these limitations.

Charges are a better measure of production than revenues generated, because medical practitioners suffer substantial, non-uniform collection losses for billed services, which were performed. However, both of these financial gauges of production are subject to individual and regional variations in price for exactly the same service. Furthermore, monetary indices are subject to fee changes and inflation, which impair comparisons of data obtained at different times. The dependent variables for total charges, which includes the value of laboratory and X-ray tests done at WMC as well as professional services, will be abbreviated **$TLPROD.**

Odiorne gives a brief, insightful review of productivity in his book, The Human Side of Management.[108] He emphasizes clear definitions of objectives and use of feedback to motivate white-collar workers. At a level of individual coaching, the measures used must be directly relevant to the worker in order to achieve increased performance. Hence, for counselling physicians, a common sense definition, e.g. "productivity is using your time in the best manner, seeing your patients, and getting paid," has direct impact.[109] The motivators are time off, patient satisfaction, and pay, so productivity needs to be defined in those terms for purposes of feedback.

For purposes other than motivation, productivity must be defined and measured exactly. At the Western Montana Clinic, raw data on output, i.e. number of patients seen, type of procedures and examinations performed, and charges made, are available for each physician from a computerized management information system. However, use of this raw data has severe drawbacks: a) because of frequent changes in government and insurance company rules, allowable prices are constantly varying, so the relationship of financial data to productivity is tenuous, even in the short term, and b) there is too much detail to review.

PHYSICIAN SERVICES: PRODUCTIVITY (PSP). A computer program called "Physician Services: Productivity" (PSP) condenses the raw production data.[110][111] PSP has a felicitous effect on the output measurement. It employs a "relative value scale" (RVS) of medical production units and thus avoids the confounding effect of price changes which occurs when output is measured in dollars. A RVS weighs the average time required to perform a procedure, the physician skill and training needed, the severity of the illness for which the procedure is indicated, the risk to the patient, and the medical and legal risk to the physician in order to rank procedures and assign them proportionate values.[112] A relative value scale in medicine has similarities to the job and skill evaluation systems employed in human resources management in industry.[113] For example, a brief office service for an established patient (90040) may rate 5.5 relative value points compared to a right heart catheterization (93501) which rates 60 relative value units.

For this study, the McGraw-Hill Relative Value Scale will be used.[114] Relative values are assigned to five basic areas: medicine, surgery, anesthesiology, laboratory including pathology, and radiology. The reason for this subdivision is that the reliability of ranking the procedures is highest within the basic groups. However, total productivity can be measured by weighting the RVs for each group and then combining them. McGraw-Hill weighting suggested by PSP is medicine 1.0, surgery 10.0, anesthesiology 10.0, laboratory 2.5, and radiology 2.5. These weights were used for the production measurements in this study. Relative values can be changed into prices by using appropriate multipliers for four basic areas, medicine, surgery, laboratory, and radiology. However, if the data are left as relative value units, they remain comparable from time to time.

Now consider the inputs to PSP. The program uses input denominators of "full time equivalent" (FTE) physicians, patients, and procedures. Multiple patient visits in a short interval complicate productivity measurements. PSP therefore counts a patient only once per month for a given physician, no matter how often the patient actually visits the physician. If 3 physicians see the same patient in a month, however, that patient is counted 3 times, once for each physician. Procedures have codes (see above) which appear on the raw production data tallies and are used for billing. Procedures (CPT4 codes) are listed in <u>Physicians' Current Procedural Terminology</u>, which is updated annually by the American Medical Association.[115]

PSP produces three output/input ratios:

{3} Productivity = RVS units/FTE physician

{4} Intensity = RVS units/patient

{5} Complexity = RVS units/procedure.

Most of the analysis will use the total or combined medical/surgery relative value productivity measures, abbreviated **TTLPROD** and followed by a number ranging from 1 to 11 to signify the number of months data collected. **MEDSRGRV** stands for the medical and surgical production.

LABRV and **XRAYRV** refer to the laboratory and radiology services, expressed in relative value units, ordered by the physician, and done at WMC. Financial and relative value unit data on ancillary service production ordered by the physician at the hospital could not be obtained. The additional PSP ratios, equations {4} and {5} listed above, are designed for analysis of productivity variances. However, the amount of physician services per patient or per procedure may involve challenge and flow as described above by Csikszentmihalyi; they will be tested as independent variables.

SUBSTITUTES AND COMPLEMENTS. If one or more output or input factors can substitute for another, or if one factor requires another to be useful, measuring productivity becomes complicated. For example, if an X-ray can be substituted for a physician's examination, output per physician could rise if an X-ray unit were used, even if the physician were seeing fewer patients than before. Thus, all inputs and outputs must be considered to give an accurate picture of productivity. It follows that if a patient has a procedure such as a laboratory test done during a given month, talks to the physician or nurse by phone, but is not seen by the physician, that laboratory test does count for the physician's production and that patient is counted for that month.

Reinhardt's argument that new technology, such as diagnostic radiology procedures and complex surgical operations, complements ongoing physician activities and therefore may increase costs without improving productivity deserves some scrutiny. After citing a number of illustrative examples, he makes the following synopsis:

> "The point developed in this section is that, in practice, the infusion of capital into the production of health care also facilitates the production of new types of services for which a demand is created by the mere existence of the new, capital-using technologies. If one can judge from past experience in this area, a realistic prognosis seems to be that any future health manpower savings achieved through bona fide capital-labor substitution in the health care sector is likely to be offset -- and perhaps even more than offset -- by capital-inducted increases in the demand for health care, and hence in the demand for all types of health manpower. Such an outcome obviously does not make greater capital intensity in health care production an unattractive proposition. One merely should not expect the infusion of capital into the health care sector to permit reductions in the aggregate physician-population ratio in future years.[116]

Although the critical issue to the government or insurance company budget maker is total cost, the issue from the point of view of the patient is the quality of the service. Reinhardt well recognizes that infusion of technology into medicine can extend useful life. What is necessary is to define medical output in enough detail that improved outputs are not lumped together with less effective procedures from past years. An office visit in which lifesaving advice is given based on current research is not the same as an office visit for the same problem a year ago before the research results were known. Measuring the changes in advice given may be very difficult, but recognizing that improved procedures represent increased output should be possible within the type of relative value unit systems currently under development.

To get some insight on the relationships of physician output and technology, focus on primary care providers. Primary care physicians spend more time talking to patients and less time using high technology equipment than procedure-oriented specialists. For primary care providers, the most often used technologies are laboratory and radiology services. Besides these ancillary services, the primary care physician's major input to total production is time. The relationship is expressed in the following general production function:

$$\{6\} \qquad TP = f(T, LS, RS)$$

Cross-elasticities of the production factors in equation 6 can be determined by multiple regression analysis. If the cross-elasticity of physician time and either laboratory or radiology services has a positive sign, the services will substitute, but if negative they complement each other.

RELATIVE VALUE SCALES. Relative value scales appear to have distinct advantages over other measures of medical output, but how do they compare with one another? Lazarus reviewed this subject up to 1987. He showed that the relative rankings of procedures were similar among several scales and that the McGraw-Hill scale was the most widely applicable.[117]

Since 1987, however, there have been significant changes. The United States Department of Health and Human Services Health Care Financing Administration (HCFA) funded an intensive study of relative value scales as part of its program to set physician fees and limit costs for its Medicare program.[118] HCFA relied on research from the Harvard School of Public Health about resource-based relative values for medical procedures.[119] The Harvard RBRVS project developed a "topography of work based on four dimensions: time, mental effort and judgment, technical skill and physical effort, and stress." Total work was a product of these relatively independent dimensions.[120] The investigators also evaluated work associated with specific procedures,[121] cross-specialty linkages of work,[122] and extrapolation methods.[123] Because the Harvard RBRVS has a better theoretical basis and a superior practical level of documentation than its predecessors, it is likely to replace the McGraw-Hill scale in 1992, when it is completed and implemented. Therefore, a comparison of the McGraw-Hill and the Harvard RBRVS is appropriate.

One method of comparing the two relative value scales is to run them both against the same set of procedure codes for a group of physicians and judge the magnitude of similarities and differences. Such an empirical approach awaits completion of the Harvard RVRBS, which currently covers only about 67% of existing procedures and is undergoing expansion and revision.

Another method of comparison is to rank common procedures on both scales, determine correlations, and then examine conceptual differences. This approach is less rigorous, but is the best one available. Consider, therefore, Table III-3 on the following page. The left column is the CPT-4 procedure code, the second column a brief description of the item, the third column the percent of total WMC charges in 1990 for this procedure, the fourth column the McGraw-Hill relative value and the right column the Harvard RBRVS units. For the Harvard RVs, work, overhead, and malpractice subcategories were used, but geographic correction factors were not applied. Weighting factors for the McGraw-Hill scale were the ones recommended for use with PSP.

The level of correlation of relative values between McGraw-Hill and Harvard RBRVS for the procedures listed, which accounted for 38.4% of total WMC revenues for 1990, is .91 for the medicine procedures, .94 for radiology procedures, and .93 for surgical procedures. That two independent relative value scales have high correlation suggests that these work units do have validity. Furthermore, an appropriate linear conversion from the McGraw-Hill to the Harvard RVRBS should be possible in the future. Therefore, the current production data can be compared with future studies of physician

TABLE III-3. CORRELATION OF HARVARD AND MCGRAW-HILL RELATIVE VALUE SCALES

CPT4 CODE	DESCRIPTION	%CHARGES	MCGW RV	HRBRVS
MEDICINE				
90000	NEW BRIEF OFFICE VISIT	0.569	5.5	24.9
90010	NEW LIMITED OFFICE VISIT	0.295	7.5	31.3
90015	NEW INTERMED OFFICE VISIT	0.228	12.5	38.0
90017	NEW EXTEN OFFICE VISIT	0.222	18.0	47.8
90020	NEW COMPREH OFFICE VISIT	0.120	26.0	61.8
90030	ESTAB MIN OFFICE VISIT	0.870	3.5	12.1
90040	ESTAB BRIEF OFFICE VISIT	5.877	5.5	21.4
90050	ESTAB LIMITED OFFICE VISIT	8.029	7.5	26.5
90060	ESTAB INTERMED OFFICE VISIT	1.350	9.0	32.4
90070	ESTAB EXTEN OFFICE VISIT	0.990	13.5	43.3
90215	HOSPITAL NEW INTERMED	1.090	18.0	62.8
90220	HOSPITAL NEW COMPREHEN	0.861	26.0	96.9
90225	HOSPITAL CARE NEWBORN	0.248	22.0	77.4
90250	HOSPITAL VISIT LIMITED	2.900	8.0	31.8
90600	LIMITED CONSULTATION	0.334	13.0	43.3
90605	INTERMEDIATE CONSULTATION	0.503	18.0	57.4
90610	EXTENSIVE CONSULTATION	1.104	25.0	77.1
90620	COMPREHENSIVE CONSULTATION	1.172	32.0	99.2
90630	COMPLEX CONSULTATION	0.160	42.0	125.0
93000	COMPLETE EKG	0.635	7.8	23.7
93018	TREADMILL	0.034	21.0	49.2
93501	RIGHT HEART CATH	0.042	60.0	266.9
93536	BALLOON CATH	0.089	105.0	221.2
99160	HOUR CRITICAL CARE	0.133	42.0	104.8
RADIOLOGY				
71020	TWO VIEW CHEST	1.072	3.0	27.9
72100	LUMBAR SPINE	0.240	4.0	35.7
73030	SHOULDER	0.115	2.7	14.4
73564	KNEE	0.167	3.0	13.4
74246	UPPER GI	0.221	6.2	62.8
74280	BARIUM ENEMA	0.284	8.6	80.8
76091	MAMMOGRAPHY	1.207	6.4	67.0
76700	ABDOMINAL ECHO	0.537	12.0	96.7
76705	LIMITED ABDOMINAL ECHO	0.097	10.0	71.1
76805	ECHO PREGNANT UTERUS	0.318	10.0	93.7
76815	LIMITED ECHO UTERUS	0.218	6.0	63.3

TABLE III-3. CORRELATION OF HARVARD AND MCGRAW-HILL RELATIVE VALUE SCALES

CPT4 CODE	DESCRIPTION	%CHARGES	MCGW RV	HRBRVS
SURGERY				
11100	SKIN BIOPSY	0.140	1.0	44.3
27130	TOTAL HIP REPLACEMENT	0.087	35.5	2182.1
27244	FEMUR FRACTURE REPAIR	0.095	24.0	1229.3
29881	KNEE ARTHROSCOPIC MENISC	0.223	14.0	249.0
35081	ABDOMINAL AORTA ANEURYSM	0.198	25.0	1867.6
35301	CAROTID ENDARTERECTOMY	0.247	15.0	1216.7
31622	BRONCHOSCOPY	0.268	4.7	230.5
44950	APPENDECTOMY	0.355	10.0	440.1
45330	SIGMOIDOSCOPY	0.262	1.3	129.4
45378	COLONOSCOPY	0.325	7.0	337.2
45385	COLONOSCOPIC POLYPECTOMY	0.601	9.0	456.3
47605	CHOLECYSTECTOMY	0.390	18.0	775.9
49505	INGUINAL HERNIA	0.321	8.5	406.2
52234	CYSTOSCOPY WITH RX	0.013	5.8	431.4
52601	TRANSURETHERAL PROSTATE	0.578	20.0	990.6
55250	VASECTOMY	0.094	4.5	231.6
58120	DILITATION CURETTAGE UTERUS	0.256	4.0	219.6
58150	TOTAL HYSTERECTOMY	0.577	17.0	787.6
59510	CAESARIAN SECTION	0.695	15.0	544.5
62270	LUMBAR PUNCTURE	0.029	1.0	60.3
64721	CARPAL TUNNEL RELEASE	0.329	8.4	378.3

PERCENT OF CLINIC REVENUES	
SUBTOTAL MEDICINE	27.9
SUBTOTAL RADIOLOGY	4.5
SUBTOTAL SURGERY	6.1
TOTAL	38.4

productivity expressed in relative value units, using either the McGraw-Hill or the Harvard RBRVS.

INDEXING

If the dependent variable in a production function is indexed to avoid confounding by geographic and inflationary changes in financial measures, independent variables should be indexed, too, for the same reasons. Labor inputs, LABOR, can be tied to the local salary and benefits of a starting level registered nurse, for example. Capital equipment investment could be indexed to a particular piece of equipment. For this study, one independent index variable was used: **RVLABOR.**

RESEARCH DESIGN

This is a cross-sectional, non-interventional, descriptive study of physician productivity. The specific population for analysis in this study are the physicians, psychologists, and podiatrists employed by the Western Montana Clinic, a multispecialty group practice in Missoula, Montana.

SAMPLE FRAME: THE WESTERN MONTANA CLINIC. The subjects of this study are physicians, psychologists, and podiatrists of WMC. These professionals have completed necessary training to be eligible for certification examinations in their specialties as a condition for employment. Over 90% are board certified. Currently, there are 12 general internists, 11 medical subspecialists, 2 neurologists, a dermatologist, 5 obstetrician-gynecologists, 4 pediatricians, 4 urgent care physicians, 5 family physicians, 4 general surgeons, 3 clinical psychologists, 2 orthopedic surgeons, 2 podiatrists, a urologist, an otolaryngologist, and a radiologist, plus many non-doctorate level health care professionals, such as physical therapists, nurses, and a dietician. Except for the radiologist, all doctoral-level practitioners who have been employed at WMC for at least 9 of the 12 months of the study were included in this study of productivity. Due to recent arrival, a general surgeon, an internist, an infectious disease specialist, and a family practitioner, who had been working for 5, 4, 4, and 1 months, respectively, were excluded from the study. Productivity and compensation for radiologists differ substantially from the other directly clinical specialties, so the radiologist was also excluded. That left a total of 53 professional staff available for study in a 58 doctor clinic.

WMC performs predominantly fee-for-service medicine. Services rendered for health maintenance organizations or other pre-paid health care plans are negligible. Professional staff members become eligible for ownership in WMC after two years. 85% of professional staff compensation is based on incentive and 15% is an equal share. About 50% of each professional staff member's office costs are charged directly to that member; the remainder of overhead costs are attributed in proportion to professional revenues.

Table III-4 shows further demographic information. Using this table, the reader can compare WMC to other multispecialty groups which are members of the Medical Group Management Association (MGMA).[124] Other socio-economic data included in the table allow comparison of WMC physicians with regional and national measures.[125][126] The demographic information should allow judgment of how applicable the results of this study may be to other clinics or physician practice situations.

DETERMINATION OF THE PRODUCTION FUNCTION. The physician production function was determined from cross-sectional data collected in 1990. The data for the independent variable vectors were assembled from WMC management information and from interviews with the WMC professional staff. A checklist/questionnaire ensured systematic and uniform collection of data (see the work sheets in the appendix). The dependent variable, productivity, in relative value units for each physician, was determined from WMC production records using the PSP computer software. As an independent check on the accuracy of PSP and the relative value outputs, the gross revenue data from PSP were cross-checked against gross revenue data generated by the WMC accounting department.

The form of the production function might be strictly

TABLE III-4. DEMOGRAPHIC INFORMATION ON CLINICS AND PHYSICIANS

CATEGORY	NATIONAL	WMC
Physicians in practice	367,963	49
Psychologists	60,000	3
Podiatrists	10,000	2
Age (median)		40
Gender (female)	18%	22.6%
Specialty board certification	79%	88%
Multispecialty clinics over 50 professsional staff (AMA, 1988 data)	316	Yes
Clinics with less than 10% managed care revenue	54%	Yes
Median clinic nonphysician expense	53.8%	48.1%
Employees per full time physician (median)	4.24	3.33
Median annual outpatient visits		
Internal medicine	3310	2354*
Obstetrics/Gynecology	3403	1956*
General surgery	1653	992*
Pediatrics	5132	3603*
Mean hours in direct patient care activities per week	53.1	31.9
Mean fee for 1988 office visit	$33.91	$35.50 (90050)

---

*PSP definition of visit, which understimates the usual direct counts. See discussion of PSP in the section on the dependent variable in the text.

linear with no interactions among independent variables, but judging from previous research, that seems highly unlikely.[127][128] Therefore, besides the simple linear models for production factor inputs, Reinhardt's modification of the Cobb-Douglas production function

$$\{7\} \qquad Q = A \prod_{j=1}^{m} (X_j^{a} \; e^{-b\,X}) \; e^{g(X,Y;c)},$$

was used also. To paraphrase Reinhardt, the elements of each vector, X_j, in this production function model stand for essential inputs such as time, which must be used at positive rates for output to occur. The Y vectors, in contrast, epitomize factors not absolutely required for positive rates of output. The function g(X,Y;c) is a polynomial which is linear in its parameters, c, and contains cross-products of essential and non-essential inputs. The model allows increasing and decreasing marginal products of the inputs and specifies optimum use rates of non-essential inputs which are related to the essential inputs. Returns to scale for this model vary over the production surface.[129] Returns to scale can be estimated in a given region by substitution of relevant values for each of the variables into the production function.[130]

Taking the natural logarithm of both sides of the Cobb-Douglas-Reinhardt production function yields the following linear transformation:

$$\{8\} \qquad \ln Q = \ln A + \sum_{j=1}^{m} a_j \ln X_j - b_j X_j + g(X,Y;c)$$

The linear transformation allows calculation of the predictive variables and their coefficients by multiple linear regression techniques, which are discussed in the next section.

Production can be conceptualized as a surface in an n-dimensional hyperspace. Each of the n dimensions corresponds to one of the input factors, which were listed in the discussion of independent variables. Somewhere in the hyperspace will occur an overall maximum, the point of optimum production. The optimum can be determined from the total differential of the Reinhardt modification of the Cobb-Douglas function. Optimum levels for single inputs can be obtained using partial derivatives of the same function.

Isoquants for production factors define how one input factor can substitute for another at various levels of output. For substitution to make sense, several criteria must be met: 1) the units of the input factors must be the same, for example, dollars; 2) the input factors must be real (positive sign); and 3) the inputs must be interchangeable in practice. If these demands are satisfied, the slope of an isoquant, termed the marginal rate of technical substitution of input factors, is the negative of the ratio of partial derivatives (marginal products) of the input variables under consideration. For example, give the production function

{9} $$Q = f(X_1, X_2)$$

where X_1 and X_2 are input factors, the slope of the isoquant which indicates how X_1 and X_2 can substitute for each other is

{10} $$dX_1/dX_2 = -Q_{x2}/Q_{x1}$$

where Q_{x1} and Q_{x2} are the partial derivatives of Q with respect to X_1 or X_2.[131] The optimal ratio of the input factors at any given level of production can be obtained by dividing partial derivative of the each of the factors by its price in dollars:

{11} $$X_1/X_2 = Q_{x1}*Price_{x2}/Q_{x2}*Price_{x1}$$

as described by Seo.[132] This assumes, however, that the inputs are homogeneous goods so that a uniform price can be specified.

MULTIPLE LINEAR REGRESSION

The primary mathematical tool for fashioning a production function is the technique of multiple linear regression. Neter, Wasserman, and Kutner define regression modeling as follows in their textbook:

> A regression model is a formal means of expressing the two essential ingredients of a statistical relation: 1. A tendency of the dependent variable Y to vary with the independent variable in a systematic fashion. 2. A scattering of points around the curve of statistical relationship.
>
> These two characteristics are embodied in a regression model by postulating that: 1. There is a probability distribution of Y for each level of X. 2. The means of these probability distributions vary in some systematic fashion with X.[133]

It follows that a regression model does not prove any cause and effect relationship among the variables. Proof of causation requires additional information.

Multivariate linear regression makes the following stipulations: 1) linear relationships between the dependent and independent variables (linearity), 2) no strong relationships (collinearity) among the independent variables, 3) residual terms of the dependent variable are normally distributed with a mean of zero (normality), 4) residual terms of the dependent variable are independent of one another (no autocorrelation), and 5) residuals have constant variance (homoscedasticity).[134]

The actual calculations for multiple linear regression were done with a computer program. Number Cruncher Statistical System (NCSS) software stored, analyzed, and reported the data in tabular and graphic formats.[135] NCSS is a well-documented product; its accuracy, judging from tests on standard data and from use by 15,000 individuals over 6 years, is excellent.[136][137]

DATA ANALYSIS AND ADDITIONAL STATISTICAL METHODS

The independent variables used for analysis were identified from results previously reported by economists about physician production and also on the basis of management and motivational theory. The objective of analysis is to build a realistic model of physician production in a multispecialty clinic setting. Hence, the predictive variables ultimately used in the production function were selected not only on the basis of their correlation with output, but also on the basis of economic theory, clarity, and practicality.

After completion of the structured interviews with professional staff and collection of data from the WMC management information system, the data vector for each variable was reviewed for accuracy as it was loaded into the computer. To assess the dependent variable data vectors, the PSP report was compared with independent assessments of the same raw data from the WMC accounting department. Next, exploratory data analysis was performed. Each data vector was subjected to a batch of descriptive statistics, which included several measures of location and variance and a frequency histogram. Scatter plots of each data vector were made against the dependent variables, TTLPROD11 and MEDSRGRV using first, second, and third order curve fitting. The residuals from each fit were graphed against the independent variable and compared. These preliminary examinations detected non-linear relationships between the independent and dependent variables, inconstancy of error variance, and the presence of outlier data points.[138]

Data vectors which should be closely related, for instance variables having to do with physician time such as vacation and meeting time, weekends and holidays on call, or percentage of office time used for patient examinations or procedures, were examined for correlation with the dependent variable and for cross-correlations with one another (collinearity). Vectors were combined to give better measures of factors of interest, for example, total time spent seeing patients, PATTIME. Data vectors in related groups which had low correlations with output were eliminated from further consideration to allow concentration on as wide a range as possible of truly independent (orthogonal) variables.

With a population of only 53 professional staff, the range of possible independent variables which can be considered by multiple linear regression is at most ten.[139][140] When preliminary data analysis was completed, the remaining independent variables were run through stepwise regressions in batches of eight to see which variables were consistently related to the dependent variables TTLPROD11 and PROFPROD11.[141] When a subset of 10 independent variables was established, a robust multiple regression was run using the Tukey biweight technique, which eliminates undue influence of outlier data points.[142] Those variables with zero or very low weights were deleted from the data to obtain stable regressions. Then the regressions were rerun using the standard, least squares technique, to see if the robust weighting made any difference. This sequence established the preliminary model of the production function.

The following tests were applied to the preliminary model of the production function to assure conformation to the multiple linear regression model. Linearity was insured by checking scatter plots, especially the decomposition displays for polynomial fitting, and by checking for patterns in the residuals plots of each of the independent variables against the dependent variable.[143] Collinearity was checked by noting the cross-correlations among the independent variables, by deleting independent variables to see if there were large changes in estimated regression coefficients, by noting the signs of the regression coefficients, and by calculating variance inflation factors.[144] Contour plots of pairs of independent variables with output on the vertical axis (production isoquants) were used to check for both collinearity and interactions among independent variables.[145] Normality was studied using a normal probability plot of the dependent variable residuals. Autocorrelation, including the possibility of significant lurking variables, was assessed by inspection of the residuals plots for patterns. Homoscedasticity was insured by inspecting a plot of the dependent variable residuals versus the independent variable and also against the fitted values of the dependent variable.[146] Finally, influential data points were checked using a Studentized residuals versus Yhat diagonals plot and Cook's D statistic.[147]

RELIABILITY AND VALIDITY OF THE DATA AND CONCLUSIONS

Repeat measurements with cross-correlations were done to check reliability, which is the consistency and stability of the data vectors. Specifically, stationarity of the production data collection interval and repeatability of the WORKAHOL rating scale were assessed.

Validity, a broad concept concerning the degree of truth of the propositions under study, can be subdivided many ways. One categorization has four main divisions: 1) statistical conclusion validity refers to the power of the study to detect covariation and the quantity of any covariation present, 2) internal validity concerns the presence or absence of cause and effect relationships among variables, 3) construct validity has to do with how well the variables are related to the more abstract concepts which they are supposed to represent, and 4) external validity which deals with whether conclusions can be generalized to other settings.[148] Questions of internal validity and construct validity were raised in the review of the productivity literature; they will be considered in detail in Chapter V. Statistical conclusion validity and external validity get more specific attention now, but also later on in Chapter V.

STATISTICAL CONCLUSION VALIDITY. When the data met the prerequisites for multiple regression analysis, tests of statistical power were applied to the independent variables.[149] Power was calculated for two levels of the regression coefficients. One level is the actual simple linear regression coefficient, b, for each of the independent variables. Power (labelled POWER.2) was also determined at a production elasticity of demand of 0.2, a level deemed relevant for business significance. That is to say, for an input factor to rate corporate attention, the dependent variable, total production, needs to vary by at least 2% if the input changes by 10%. The coefficient of simple regression needed to achieve a production elasticity of 0.2 was calculated from the following formula:

{12} $$b_{0.2} = 0.2 * \text{mean } Y / \text{mean } X_i$$

Then b and $b_{0.2}$ were divided by the standard error of the regression coefficient to yield delta and delta.2. Using the deltas, the power at an alpha level of 5% and 51 degrees of freedom was determined from Table A.5 in the appendix to Neter's book.[150]

Power was used to screen variables for inclusion in the production functions, so power was determined only for simple regression coefficients. Independent variables which had power level of less than 60% for both levels of power determination were dropped from further consideration. Other tests of statistical significance, however, namely analyses of variance, were applied to the independent variables in the final production functions.

EXTERNAL VALIDITY. To validate the production functions, several approaches were used. First, the functions were compared with those from previous research described in Chapter II. The results were also compared with data from other industries.

Second, the PRESS (prediction sum of squares) criterion was used instead of large scale data splitting, since the number of subjects in the original data base is limited. For PRESS, each data point is estimated from the least squares fitted regression function determined from all of the other data points. PRESS supports external validity if it agrees well with the sum of squares error.[151]

Third, data were obtained from three medical clinics to test whether the production functions have broad applicability. The clinics are the Lexington Clinic in Kentucky, the Medical Center Clinic in Pensacola, Florida, and the Wichita Falls Clinic in Texas. The clinics which participated did so out of good will, since they could expect no direct return for the time investment in this project. They were selected from a list of PSP users provided by the MGMA. Clinic size and the degree to which salary is based on production have been shown to affect physician productivity, as already discussed. These factors were monitored, but they may confound the validation of the production function.

Furthermore, the reliability of data provided by other clinics could not be checked as closely as the WMC data. For the other clinics, estimated and observed production were compared; the mean square prediction error was calculated as a measure of the predictive capability of the model.[152]

CHAPTER IV. RESULTS

PRELIMINARY DATA ANALYSIS

DEPENDENT VARIABLES. The accuracy of the dependent variables was evaluated first. The patient visit counts appeared plausible, based on knowledge of individual physician practice patterns. Except for physicians who do predominantly hospital work, the patient counts from PSP were less than the patient tallies from office schedules. The differences were anticipated because of the PSP definition of patient visits.

Since patient counts appeared correct, attention was next given to output in relative value units. For the psychologists, production in relative value units seemed inappropriately low. Unlike physicians' and podiatrists' codes, which designate a procedure, the McGraw-Hill book lists the psychologists' codes in relative value units per minute. Multiplying the McGraw-Hill values by minutes scheduled per procedure corrected the shortfall. Relative value data for other practitioners had the right order of magnitude and appeared in plausible rank order.

Next, the gross revenues from PSP were compared with the gross revenues from the WMC accounting system. The gross charges from the two sources should correspond exactly, since both come from the same raw data. Suprisingly, there were many discrepancies. In most instances, PSP gross charges were 2.5 percent per month higher than the WMC accounting results; the largest difference was 25 percent.

To reconcile the differences, the raw data for one month were entered manually instead of downloading automatically from the mainframe computer into the microcomputer which ran the PSP program. Manual loading of data corrected the discrepancies. The correlation for gross charges from manually loaded PSP and from WMC accounting was .9997; the largest individual variation was 3.4 percent. For procedures, the correlation was .9929. Thus, the disparities were due to some error in downloading raw data from the mainframe.

A phone call to the software designers revealed that the micro-computer PSP program does not identify minus signs used to correct raw data entries for the mainframe computer. Thus, mainframe data entry errors were doubled by PSP instead of being corrected. This problem affected procedure counts, gross charges, and relative values.

Manual loading of data into PSP for an entire year was too time consuming to be practical. Therefore, a "quick fix" was devised to correct the PSP errors. WMC accounting gross charges and procedure counts were substituted for the inaccurate PSP results. Relative values were corrected by prorating from the correction in the gross charges.

For one month, the raw data correction and the "quick fix" proration method for calculating relative value units could be compared. Correlation of the two methods was .9975, but five of the 53 subjects had between 5% and 11% variation from the raw data method. The reasons for these variations were investigated at length. Large variations were related to the type of the raw data entry errors. No prospective method could be found to detect such outliers. Different physicians were outliers in the second month, so the errors appeared to be random rather than systematic. Such errors should "cancel out" as more data are added.

Because the correlation of the manually entered raw data and the "quick fix" prorating method was so good, the quicker method was applied to the rest of the production data. To be sure that errors were random and that no systematic biases were injected into the relative value data, all regressions were run both against gross charges and relative value units.

Some characteristics of the output variables, including distributions and cross-correlations, are summarized in Table IV-1 on the following page. The correlation of financial output with relative value based output for this cross-sectional study is .95. For theoretical and proprietary reasons, the use of financial measures of output was limited to the accuracy studies just described and to external validation. The cross-correlations of patient visits with other measures of output are low.

The limited correlation of total production in relative

TABLE IV-1. OUTPUT VARIABLES

I. MEASURES OF DISTRIBUTION

VARIABLE	MEAN	STANDARD DEVIATION	HISTOGRAM
TLPROD11	4.938	2.019	SKEWED
$TLPROD			SKEWED
MEDSRGRV	4.006	1.736	SKEWED
LABRV	0.564	0.518	SKEWED
XRAYRV	0.384	0.306	SKEWED
PATIENTS	208.6	162.8	SKEWED

II. CROSS-CORRELATIONS

VARIABLE	TTLPR	$TTLPRD	MSRV	LABRV	XRAYRV	PATIENT
TTLPROD11	1.00	0.95	0.93	0.48	0.45	0.25
$TLPROD	0.95	1.00	0.90	0.42	0.43	0.06
MEDSRGRV	0.93	0.90	1.00	0.15	0.15	0.28
LABRV	0.48	0.42	0.15	1.00	0.68	0.04
XRAYRV	0.45	0.43	0.15	0.68	1.00	0.05
PATIENTS	0.25	0.06	0.28	0.04	0.05	1.00

value units with patient visits is explored graphically in Figure IV-1. Each of the three panels in the figure shows a different scatter plot of PATIENTS versus TTLPROD11. The first panel shows all of the data points. For this data, the adjusted coefficient of determination is 0.04 and the probability that the relationship could occur by chance alone is 0.07. Thus, in the first panel, the relationship between production in relative value units and patients seen is very weak. The second panel shows only data points for primary care physicians. Here there is a relatively strong relationship between output in relative value units and patients seen: the adjusted coefficient of determination is 0.33 and the probablility of the null hypothesis is less than 0.000. The third scatter plot is limited to data points for specialists doing procedures. For specialists, the adjusted coefficient of determination is 0.02 and the probability of the null hypothesis being correct is 0.24.

INDEPENDENT VARIABLES. The results of preliminary and exploratory data analysis for the independent variables are shown in Table IV-2. The variables are grouped according to theoretical relationships with one another. The units for the means and standard deviations were explained in the descriptions of independent variables. The frequency distribution of each independent variable is listed in the fourth column. The results of scatter plots and residual

FIGURE IV-1

PATIENT VISITS OR RELATIVE VALUES?

Primary care versus specialty output

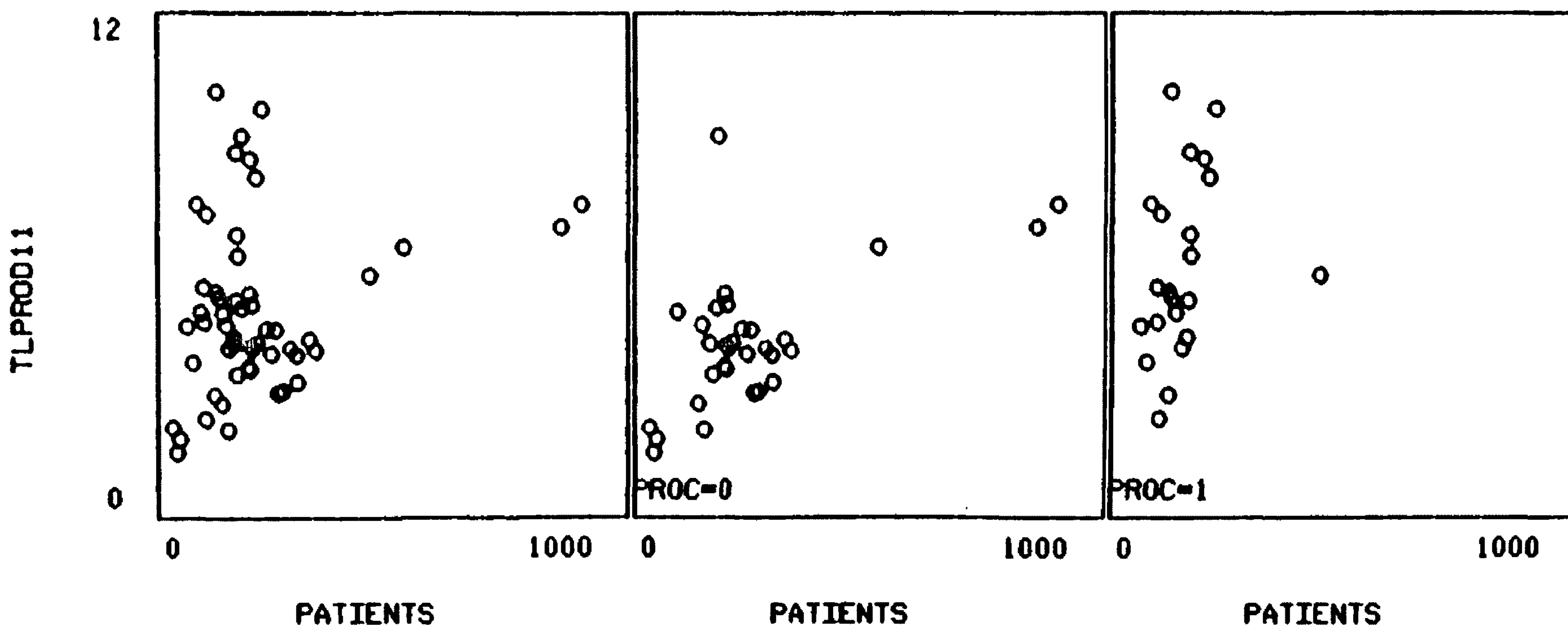


TABLE IV-2. PRELIMINARY AND EXPLORATORY DATA ANALYSIS OF INDEPENDENT VARIABLES AGAINST TOTAL PRODUCTION

VARIABLE	MEAN	STANDARD DEVIATION	DISTRIB-UTION	SCATTER-PLOT	RESID. PLOT	R STUDENT RESIDUAL OUTLIERS?
ABSENT	43.10	18.70	NORMAL	RANDOM	NORMAL	2
CALL	19.10	11.90	BIPHASIC	LINEAR	NORMAL	1
APPT%	0.59	0.18	NORMAL	RANDOM	NORMAL	2
PATTIME	138.10	37.90	SKEWED	LINEAR	NORMAL	1
LABOR	22625.00	8606.00	SKEWED	LINEAR	NORMAL	1
RVLABOR	1.08	0.41	SKEWED	LINEAR	NORMAL	1
PROC	0.42	0.50			NORMAL	1
SPECIALTY	1.35	0.54	BIPHASIC	LINEAR	NORMAL	0
SPACE	2.92	0.82	NORMAL	LINEAR	NORMAL	1
EQUIP	523.80	594.40	SKEWED	LINEAR	NORMAL	1
AGE	42.10	8.50	SKEWED	RANDOM	NORMAL	1
EXP	11.10	8.00	SKEWED	RANDOM	NORMAL	1
EXPWMC	7.90	7.90	SKEWED	RANDOM	NORMAL	1
GENDER	0.77	0.42			NORMAL	1
MARITAL	0.91	0.30			NORMAL	1
DEPEND	3.40	1.40	NORMAL	RANDOM	NORMAL	1
OUTSIDE	0.42	0.50			NORMAL	2
DEBT	0.55	0.50			NORMAL	1
SAVINGS	0.64	0.48			NORMAL	1
PRIORITY	2.00	0.67	NORMAL	LINEAR	NORMAL	0
HOBBIES	7.50	4.20	NORMAL	RANDOM	NORMAL	2
OFFICE	0.38	0.49			NORMAL	1
FELLOW	0.57	0.50			NORMAL	1
CME	0.60	0.49			NORMAL	1
REPUTAT	2.45	0.24	SKEWED	RANDOM	NORMAL	1
BURNOUT	3.60	2.40	NORMAL	RANDOM	NORMAL	1
SELFRATE	7.10	1.20	SKEWED	RANDOM	NORMAL	2
HEALTH	8.70	1.30	SKEWED	RANDOM	NORMAL	2
EXERCISE	3.90	2.90	SKEWED	RANDOM	NORMAL	1
FATIGUE	2.40	0.70	NORMAL	RANDOM	NORMAL	2
WORKAHOL	6.40	1.50	SKEWED	RANDOM	NORMAL	1
WKHOL2	6.70	1.50	NORMAL	LINEAR	NORMAL	1
COMPLEX	31.70	19.60	SKEWED	RANDOM	NORMAL	1
INTENSTY	8.80	5.50	SKEWED	LINEAR	NORMAL	1
FLOW	3.30	0.58	SKEWED	RANDOM	NORMAL	2
ANCILL	10.10	7.50	SKEWED	LINEAR	NORMAL	1
ANCILL%	0.11	0.10	SKEWED	RANDOM	NORMAL	1
UNDER	0.08	0.08	SKEWED	RANDOM	NORMAL	2

TABLE IV-2 CONTINUED. PRELIMINARY DATA ANALYSIS OF INDEPENDENT VARIABLES AGAINST TOTAL PRODUCTION

VARIABLE	GROUP CROSS CORR.	EXPECTED SIGN OF COEFF.	REGRES-SION COEFF.	STD ERR OF REGR COEFF.	T VALUE	PROB-ABILITY
ABSENT	YES	-	0.0110	0.0150	0.75	0.45
CALL	YES	+	0.0640	0.0220	2.90	0.01
APPT%	YES	+	-2.5980	1.5560	-1.67	0.10
PATTIME	YES	+	0.0156	0.0052	2.68	0.01
LABOR		+	0.0001	0.0000	2.51	0.02
RVLABOR		+	1.6250	0.6460	2.51	0.02
PROC	YES	+	2.8217	0.4196	3.28	0.00
SPECIALTY	YES	+	2.1190	0.4350	4.88	0.00
SPACE	NO	+	0.8860	0.3200	2.77	0.01
EQUIP	YES	+	0.0019	0.0004	4.86	0.00
AGE	YES	CHANGING	0.0137	0.0333	0.41	0.68
EXP	YES	CHANGING	0.0069	0.0354	0.19	0.85
EXPWMC	YES	CHANGING	0.0156	0.0356	0.44	0.66
GENDER		?	0.2608	0.6682	0.39	0.70
MARITAL		+	0.4046	0.9564	0.42	0.67
DEPEND		+	0.1232	0.2048	0.60	0.55
OUTSIDE		-	-0.6204	0.5617	-1.10	0.27
DEBT		+	0.1767	0.5621	0.31	0.75
SAVINGS		-	0.1933	0.5834	0.33	0.74
PRIORITY		-	-0.6075	0.4142	-1.47	0.15
HOBBIES		CHANGING	-0.0489	0.0664	-0.74	0.46
OFFICE		+	0.1072	0.5776	0.19	0.85
FELLOW		+	0.4054	0.5622	0.72	0.47
CME		+	0.6055	0.5692	1.06	0.29
REPUTAT		+	-0.0281	1.1947	-0.02	0.98
BURNOUT		-	-0.0095	0.1191	-0.80	0.43
SELFRATE		+	-0.2592	0.2393	-1.08	0.28
HEALTH		+	0.1942	0.2193	0.89	0.38
EXERCISE		+	-0.0025	0.0096	-0.26	0.79
FATIGUE		-	-0.3580	0.3927	-0.91	0.37
WORKAHOL		+	0.1333	0.1902	0.70	0.49
WKHOL2		+	0.1544	0.1952	0.79	0.43
COMPLEX	YES	+	0.0587	0.0511	1.15	0.26
INTENSTY	YES	+	0.0436	0.0131	3.33	0.00
FLOW	NO	+	0.5147	0.4862	1.06	0.29
ANCILL		+	0.0810	0.0361	2.24	0.03
ANCILL%		+	-2.5443	2.9574	-0.86	0.39
UNDER		-	0.4558	3.6480	0.12	0.90

TABLE IV-2 CONTINUED. PLELIMINARY DATA ANALYSIS OF INDEPENDENT VARIABLES AGAINST TOTAL PRODUCTION

VARIABLE	DELTA(b)	POWER(b)	b0.2	DELTA.2	POWER.2
ABSENT	0.73	11.0%	0.02	1.53	28.0%
CALL	2.91	78.0%	0.05	2.35	60.0%
APPT%	-1.67	36.0%	1.69	1.08	16.0%
PATTIME	2.98	83.0%	0.01	1.36	26.0%
LABOR	5.11	99.8%	0.00	1.78	40.0%
RVLABOR	2.52	68.0%	0.91	1.42	27.0%
PROC	6.73	100.0%	2.38	5.67	100.0%
SPECIALTY	4.87	99.8%	0.73	1.68	36.0%
SPACE	2.77	86.0%	0.34	1.06	16.0%
EQUIP	4.87	99.8%	0.00	4.83	99.6%
AGE	0.41	6.0%	0.02	0.70	8.0%
EXP	0.19	3.5%	0.09	2.51	66.0%
EXPWMC	0.44	6.0%	0.13	3.51	93.0%
GENDER	0.39	6.0%	1.28	1.91	46.0%
MARITAL	0.42	6.0%	1.09	1.14	17.0%
DEPEND	0.60	8.0%	0.29	1.42	27.0%
OUTSIDE	-1.10	17.0%	2.38	4.24	98.3%
DEBT	0.31	5.0%	1.80	3.21	87.0%
SAVINGS	0.33	5.0%	1.54	2.64	72.0%
PRIORITY	-1.47	28.0%	0.49	1.19	21.0%
HOBBIES	-0.74	11.0%	0.13	1.98	50.0%
OFFICE	0.19	3.5%	2.62	4.53	99.0%
FELLOW	0.72	11.0%	1.74	3.10	84.0%
CME	1.06	17.0%	1.64	2.87	78.0%
REPUTAT	-0.02	2.5%	0.40	0.34	6.0%
BURNOUT	-0.08	3.0%	0.27	2.30	59.0%
SELFRATE	-1.08	28.0%	0.14	0.58	8.0%
HEALTH	0.89	14.0%	0.11	0.52	8.0%
EXERCISE	-0.26	4.0%	0.25	26.32	100.0%
FATIGUE	-0.91	14.0%	0.41	1.05	16.0%
WORKAHOL	0.70	11.0%	0.15	0.81	11.0%
WKHOL2	0.79	11.0%	0.15	0.76	10.0%
COMPLEX	1.15	19.0%	0.03	0.61	7.0%
INTENSTY	3.33	90.0%	0.11	8.57	100.0%
FLOW	1.06	28.0%	0.30	0.62	7.0%
ANCILL	2.24	60.0%	0.10	2.71	73.0%
ANCILL%	-0.86	14.0%	9.32	3.15	85.0%
UNDER	0.12	3.0%	12.50	3.43	92.0%

plots of the independent variables against total production are listed in the fifth and sixth columns. Residual plots refers to graphs of residuals versus independent variables, residuals versus predicted values, and residuals versus expected normal quantiles. None of the residual plots showed any strong deviations from normality. Studentized residual outliers, listed in column seven, helped check for data entry errors, but never exceeded 5% of the total population, as expected.

Independent variables in certain groups have high cross-correlations as anticipated. The cross-correlations are listed in the second column of Table IV-2 on page 74. To avoid problems with collinearity, one best predictor variable from each group was chosen for use in the final production functions. The predicted and actual signs of the simple linear regression coefficients are listed in columns three and four. The T values and probabilities in the last two columns on page 74 are significant for CALL, PATTIME, LABOR, RVLABOR, PROC, SPECIALTY, SPACE, EQUIP, INTENSITY, and ANCILL. The probability level for the null hypothesis, that is, only chance relationship between the independent variable and output, was 0.03 or less for each of these independent variables.

The last five columns of Table IV-2 on page 75 are the delta and power of the simple linear regression coefficients and the regression values, delta, and power for production elasticity of 0.2 for each independent variables. Note that all of the independent variables which were significantly related to output have substantial power.

While considering Table IV-2, recall that the simple regression coefficient of a variable does not always provide a guide to the importance of that variable in multiple regression analysis.[153] For that reason, all of the independent variables with power over 60% were screened to decide on the optimal subset for the production functions even if statistical significance in simple regression was low.

SUBSTITUTION OF ANCILLARY SERVICES FOR PHYSICIAN SERVICES

The adjusted coefficient of determination for the regression model of primary care physician output (equation {6} on page 47) is .39, which is statistically significant at the .001 level by analysis of variance. Correlation analysis revealed interactions between LABRV and XRAYRV at a level of 0.65. This cross-correlation is expected, since many patients have both lab and X-ray procedures performed when they see a physician. The cross-elasticities of production of physician services, laboratory services, and radiology services are shown in Table IV-3.

IMPORTANT FACTORS IN PRODUCTION

Listed in the second column of Table IV-4 on page 79 are the independent variables which, grouped together, best

TABLE IV-3. CROSS-ELASTICITIES OF MEDICAL SERVICES

SERVICE	SUBSTIT/COMPLEMT	CROSS-ELASTICITY
PHYSICIAN TIME	LAB	0.0007
LAB	PHYSICIAN TIME	40.5923
PHYSICIAN TIME	RADIOLOGY	−0.0001
RADIOLOGY	PHYSICIAN TIME	−18.1041
LAB	RADIOLOGY	0.4209
RADIOLOGY	LAB	1.0147

TABLE IV-4. IMPORTANT FACTORS IN PHYSICIAN PRODUCTION

FUNCTION	VARIABLE	REGRESSION COEFFICIENT	STANDARD ERROR	SIMPLE R SQ	SEQUENTIAL R SQ	ELASTICITY
PLANNING	PATTIME	0.0181	0.0050	0.1231	0.1231	0.506
(TTLPROD)	RVLABOR	1.1060	0.5937	0.1102	0.1961	0.243
	SPACE	0.7110	0.2865	0.1304	0.2867	0.421
	EQUIP	0.0019	0.0003	0.3169	0.6188	0.203
	CME	0.7079	0.3730	0.0314	0.6460	0.087
PREDICT	PATTIME	0.0251	0.0050	0.1231	0.1231	0.701
(TTLPROD)	SPECLTY	2.0524	0.3375	0.3180	0.4667	0.561
	SPACE	0.9328	0.2215	0.1304	0.5980	0.552
	EXPWMC	0.0396	0.0238	0.0037	0.6198	0.063
REINHRDT	LNEQUIP	0.2147	0.0300	0.3734	0.3734	0.215
(LNTLPRD)	LNPTTIME	0.4805	0.1515	0.0715	0.4899	0.481
	RVLBSPC	0.0541	0.0160	0.1423	0.5901	0.421
	SPCME	0.0590	0.0236	0.0939	0.6313	0.231
	OUTSIDE	-0.1456	0.0748	0.0349	0.6576	-0.139
	EXERCISE	0.0165	0.0122	0.0005	0.6707	0.147
TIME	WORKAHOL	9.1366	3.3793	0.1531	0.1531	0.432
(PTTIME)	UNDER	168.5281	65.8125	0.1408	0.2608	0.097

TABLE IV-4. IMPORTANT FACTORS IN PHYSICIAN PRODUCTION

FUNCTION	ADJUSTED COEFFICIENT DETERMIN	F-RATIO	PROBABILITY	VARIABLE	T-VALUE	PROBABILITY
PLANNING	0.61	17.15	0.000	PATTIME	3.65	0.0007
(TTLPROD)				RVLABOR	1.86	0.0687
				SPACE	2.48	0.0167
				EQUIP	6.22	0.0000
				CME	1.90	0.0638
PREDICT	0.59	19.56	0.000	PATTIME	5.03	0.0000
(TTLPROD)				SPECLTY	6.08	0.0000
				SPACE	4.21	0.0001
				EXPWMC	1.66	0.1034
REINHARDT	0.63	15.61	0.000	LNEQUIP	7.16	0.0000
(LNTLPRD)				LNPTTIME	3.17	0.0027
				RVLBSPC	3.39	0.0014
				SPCME	2.50	0.0160
				OUTSIDE	-1.95	0.0575
				EXERCISE	1.35	0.1833
TIME	0.23	7.94	0.001	WORKAHOL	2.70	0.0096
(PTTIME)				UNDER	2.56	0.0139

correlate with output of medical services. The regression coefficients, standard errors of the regression coefficients, coefficients of simple determination, marginal coefficients of partial determination, and the output elasticities for the independent variables are listed in columns three through seven. In the continuation of Table IV-4 on page 80, the F-ratios in column three are for the adjusted coefficients of determination for the dependent variable from the previous column. The T-values for each of the independent variables are listed in column six. Columns four and seven give the probability that the null hypothesis is correct for the F-ratios and T-values listed, i.e. that the results are due to chance association.

For each of the four groupings shown in Table IV-4, the cross-corrleations of the independent variables are listed in Table IV-5 on page 82.

DETERMINANTS OF PHYSICIAN TIME INPUT

For reasons discussed below, preliminary data analysis was extended to include time as an output variable. PATTIME and CALL both are strongly related to production and have reasonable power, but PATTIME was selected as the dependent variable because it is a more comprehensive measure of time spent in patient care.

Relationships among the four time variables were checked by multiple regression of CALL, ABSENT, and APPT% on PATTIME.

IV-5. CROSS-CORRELATIONS OF PRODUCTION FACTORS

Planning Production Function, Equation {13}

	PATTIME	RVLABO	SPACE	EQUIP	CME
PATTIME	1.000	0.190	-0.116	0.006	-0.009
RVLABOR	0.190	1.000	0.615	-0.164	-0.257
SPACE	-0.116	0.615	1.000	0.002	-0.170
EQUIP	0.006	-0.164	0.002	1.000	0.203
CME	-0.009	-0.257	-0.170	0.203	1.000

Predictive Production Function, Equation {14}

	PATTIME	SPECLT	SPACE	EXPWMC
PATTIME	1.000	-0.060	-0.116	-0.276
SPECLTY	-0.060	1.000	0.080	0.102
SPACE	-0.116	0.080	1.000	-0.053
EXPWMC	-0.276	0.102	-0.053	1.000

Cobb-Douglas-Reinhardt Production Function, Equation {15}

	LNEQUIP	LNPTTIM	RVLBSPC	SPCME	OUTSIDE	EXERCIS
LNEQUIP	1.000	-0.117	0.028	0.203	0.132	-0.176
LNPTTIME	-0.117	1.000	0.133	-0.117	-0.279	-0.027
RVLBSPA	0.028	0.133	1.000	0.045	-0.134	-0.031
SPCME	0.203	-0.117	0.045	1.000	0.092	-0.038
OUTSIDE	0.132	-0.279	-0.134	0.092	1.000	0.012
EXERCISE	-0.176	-0.027	-0.031	-0.038	0.012	1.000

The results are shown in Table IV-6 on the following page.

The next step in the time analysis was to eliminate from consideration variables which 1) have no logical, a priori relationship to time, 2) lack construct validity, or 3) were already included in the medical services production functions because of strong correlation with output. Detailed analysis was done on the 21 remaining variables. The results are shown in Table IV-7.

Nine of the variables have power of 60 percent or less, so they were not considered further. The remaining 12 variables were screened by stepwise regression in groups of eight. Significant correlation with PATTIME was evident only with WORKAHOL (and its replicate, WKHOL2) and UNDER, as shown in Table IV-4, which also lists the time elasticities of these independent variables.

PRODUCTION FUNCTIONS

Two simple linear production functions without interaction terms were determined directly from the results of the multiple linear regression analyses:

{13} TTLPROD = -2.266 + .018 * PATTIME + 1.106 * RVLABOR + .711 * SPACE + .002 * EQUIP + .708 * CME

{14} TTLPROD = -3.538 + .021 * PATTIME + 1.969 * SPECIALTY + .672 * SPACE + .011 * EXPWMC

TABLE IV–6. CROSS–CORRELATIONS OF FACTOR GROUPS

A. TIME VARIABLES

	ABSENT	CALL	APPT%	PATTIME
ABSENT	1.00	0.42	–0.22	–0.30
CALL	0.42	1.00	–0.58	0.13
APPT%	–0.22	–0.58	1.00	0.25
PATTIME	–0.30	0.13	0.25	1.00

B. CAPITAL EQUIPMENT VARIABLES

	PROC	SPACE	EQUIP	SPECLTY
PROC	1.00	–0.09	0.73	0.76
SPACE	–0.09	1.00	0.00	0.08
EQUIP	0.73	0.00	1.00	0.69
SPECLTY	0.76	0.08	0.69	1.00

TABLE IV–7. PRELIMINARY AND EXPLORATORY DATA ANALYSIS OF INDEPENDENT VARIABLES FOR PHYSICIAN TIME OUTPUT

VARIABLE	MEAN	SCATTE PLOT	EXPECTD SIGN	REGRES COEFF	STD ERR COEFF	T VALUE
PROC	0.415	–	?	–7.30	10.62	–0.69
GENDER	0.774	–	0	5.56	12.54	0.44
MARITAL	0.906	–	–	0.81	17.99	0.04
DEPEND	3.377	Random	+	1.67	3.85	0.43
DEBT	0.547	–	+	2.22	10.56	0.21
SAVINGS	0.642	–	–	–4.47	10.95	–0.41
PRIORITY	2.009	Linear	–	–10.78	7.80	–1.38
HOBBIES	7.500	Random	–	0.57	1.25	0.46
OFFICE	0.377	–	+	0.51	10.85	0.05
FELLOW	0.566	–	+	–0.25	10.60	–0.24
REPUTE	2.448	Random	+	–9.53	22.39	–0.43
BURNOUT	3.599	Linear	–	2.36	2.23	1.06
SELFRAT	7.108	Random	+	–4.57	4.50	–1.02
HEALTH	8.689	Random	+	–0.85	4.15	–0.2
FATIGUE	2.046	Random	–	9.13	7.32	1.25
WORKAH	6.387	Linear	+	9.30	3.34	2.78
WKAHL2	6.712	Linear	+	7.40	3.58	2.07
COMPLEX	8.767	Random	+	–0.18	0.97	–0.18
FLOW	3.278	Random	+	–4.12	9.21	–0.45
ANCILL%	0.106	Random	–	–12.81	55.90	–0.23
UNDER	0.079	Linear	+	191.10	69.61	2.75

TABLE IV-7. PRELIMINARY AND EXPLORATORY DATA ANALYSIS OF INDEPENDENT VARIABLES FOR PHYSICIAN TIME OUTPUT

VARIABLE	PROB-ABILITY	DELTA	POWER	b0.2	DELTA.2	POWER.
PROC	0.495	-0.69	8.0%	65.9	6.20	100.0%
GENDER	0.660	0.44	6.0%	35.4	2.82	77.0%
MARITAL	0.964	0.04	2.0%	30.2	1.68	36.0%
DEPEND	0.667	0.43	6.0%	8.1	2.10	52.0%
DEBT	0.834	0.21	3.0%	50.0	4.73	99.5%
SAVINGS	0.685	-0.41	6.0%	42.6	3.90	96.5%
PRIORITY	0.173	-1.38	26.0%	13.6	1.75	39.0%
HOBBIES	0.648	0.46	6.0%	3.6	2.92	78.0%
OFFICE	0.963	0.05	2.0%	72.5	6.68	100.0%
FELLOW	0.812	-0.02	2.0%	48.3	4.56	99.1%
REPUTE	0.672	-0.43	6.0%	11.2	0.50	6.5%
BURNOUT	0.294	1.06	16.0%	7.6	3.42	91.0%
SELFRAT	0.315	-1.02	16.0%	3.8	0.86	12.0%
HEALTH	0.839	-0.20	3.0%	3.1	0.76	11.0%
FATIGUE	0.218	1.25	21.0%	13.4	1.83	42.0%
WORKAH	0.008	2.78	77.0%	4.3	1.28	23.0%
WKAHL2	0.044	2.07	51.0%	4.1	1.14	18.0%
COMPLEX	0.857	-0.18	3.0%	3.1	3.21	87.0%
FLOW	0.656	-0.45	6.0%	8.3	0.91	13.0%
ANCILL%	0.820	-0.23	3.0%	258.2	4.62	99.2%
UNDER	0.009	2.75	74.0%	348.4	5.00	99.8%

Equation {13} will be termed the "planning production function" and equation {14} the "predictive production function" for reasons discussed below.

Interactions among variables were considered while developing the the Reinhardt modification of the Cobb-Douglas production function. The variables and interaction terms considered are listed in Table IV-8. Because of the small size of the regression sample, not all of the 24 independent variables and interaction terms could be screened simultaneously. Screening was done against LNTLPRD in batches of 8 variables. The first batch contained EQUIP, LNEQUIP, PATTIME, LNPTTIME, RVLABOR, SQRVLAB, RVLBSPAC, RVLBCME. Stepwise regression identified LNEQUIP, LNPTTIME, RVLBSPAC, and RVLBCME as significant predictor variables. Therefore, EQUIP, PATTIME, RVLABOR, and SQRVLAB were dropped from the model and four more variables were tried. Seven runs resulted in an eight variable list of significant terms, but two of the interaction terms were eliminated because of collinearity. That yielded the final Reinhardt-Cobb-Douglas model:

$$\{15\} \qquad \text{TTLPROD} = .097 * (\text{EQUIP})^{.215} * (\text{PATTIME})^{.480} * e^{(.054 * \text{RVLBSPAC} + .059 * \text{SPCME} - .146 * \text{OUTSIDE} + .016 * \text{EXERCISE})}$$

The linear production function for time is as follows:

$$\{16\} \qquad \text{PATTIME} = 64.4 + 9.1 * \text{WORKAHOL} + 168.5 * \text{UNDER}.$$

IV-8. VARIABLES EVALUATED FOR THE COBB-DOUGLAS-REINHARDT PRODUCTION FUNCTION, EQUATION {15}

VARIABLE	EXPLANATION
DEPENDENT	
LNTLPRD	Logarithm of TTLPRD11
LN$TLPRD	Logarithm of $TTLPROD
LNMSRV	Logarithm of medical/surgical services
INDEPENDENT	
EQUIP	Replacement value, frequently used equipment
LNEQUIP	Logarithm of EQUIP
PATTIME	Monthly hours in patient care activities
LNPATTIME	Logarithm of PATTIME
RVLABOR	Indexed salaries of assistants
SQRVLAB	RVLABOR squared
RVLBSPACE	RVLABOR times SPACE
RVLBCME	RVLABOR times CME
SPACE	Offices used
SQSPACE	SPACE squared
SPCME	SPACE times CME
TWKAHOL	PATTIME time WORKAHOL
TEXPWMC	PATTIME times EXPWMC
TOUTSIDE	PATTIME times OUTSIDE
EQEXPWMC	EQUIP times EXPWMC
CME	Attended national specialty meeting in 1990
EXPWMC	Years worked at WMC
OUTSIDE	Outside income or working spouse
EXERCISE	Hours of aerobic exercise weekly
WORKAHOL	Self-rating of workaholism
EQOUTSIDE	EQUIP times OUTSIDE
EQWKAHOL	EQUIP times WORKAHOL

THE PRODUCTION SURFACE

Returns to scale for linear production functions are constant. Partial differentiation of the linear production functions results only in positive or negative constants, which means that output can be increased in the linear model by augmenting any of the input factors. Calculation of the marginal rates of technical substitution for the input factors demands conversion of units. The factors PATTIME, RVLABOR, SPACE, EQUIP, and perhaps CME in equation {13} meet the practical substitution criterion. To convert the factors to dollar units per month, multiply PATTIME by $60 per hour, which is a round estimate of net hourly pay. Substitute LABOR for the RVLABOR index. SPACE can be converted into dollars by using a monthly rental fee of $392 per month per room determined from WMC accounting data. CME is a small component of the production function. To convert CME into financial monetary units requires additional information: cost of meeting tuition, transportation, lodging, and lost production. At best, CME can be estimated roughly, so it was simply left out.

The marginal rate of technical substitution of PATTIME for LABOR is -1.75, of PATTIME for EQUIP -0.31, of PATIME for SPACE -5.82, and of LABOR for EQUIP -0.18. Because EQUIP is heterogenous, there is no single price, so optimum ratios involving capital equipment cannot be calculated for this physician production function. Likewise, both physician and non-physician labor are heterogeneous. Assuming costs of salaries and benefits of $60/hour for physicians and $15/hour for supporting staff, the optimum ratio of physician time to non-physician labor input is .14 or approximately 7 support staff per physician if both work the same number of hours.

Figure IV-2 shows the production isoquants of PATTIME and LABOR calculated using the raw data with the an exponential smoothing function provided with the NCSS software.[154]

Equation {14} is a predictive production function and was not designed for optimizing inputs by substitution, nor do the input factors in the physician time function, equation {16}, appear to be practical substitutes.

Partial differentiation of Reinhardt-Cobb-Douglas model, equation {15}, with respect to capital equipment investment or time spent seeing patients or the variables in the exponential term also yields no local optimum. All of the partial derivatives except the one with respect to SPACE are simply constants, which cannot be set to zero. Setting the partial derivative of SPACE to zero demands a negative input of either non-physician labor or continuing medical education, which makes no economic sense. Hence, useful optima from the Cobb-Douglas-Reinhardt production function cannot be determined for this application.

For the Cobb-Douglas-Reinhardt production function, increasing the current level of all input factors by 10%

PRODUCTION ISOQUANTS OF TIME AND LABOR

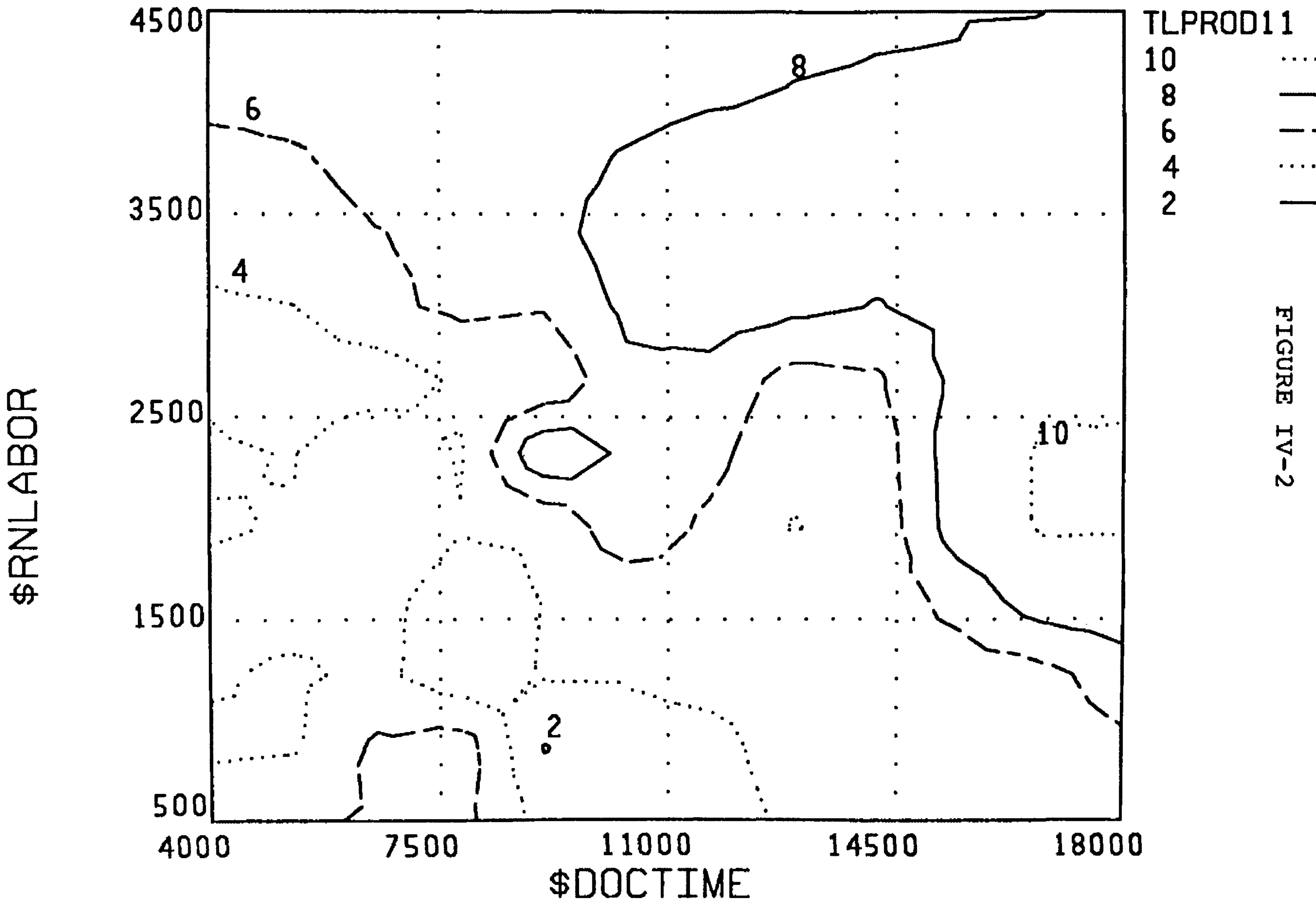


FIGURE IV-2

causes a 13.1% increase in output. This is positive returns to scale.

FACTORS FOR WHICH POWER WAS INSUFFICIENT

The following independent variables had insufficient power to be used as predictors of physician production in this study: ABSENT, APPT%, AGE, GENDER, MARITAL, DEPEND, PRIORITY, HOBBIES, REPUTAT, BURNOUT, SELFRATE, HEALTH, FATIGUE, COMPLEX, WORKAHOL, and FLOW. From this list, GENDER, HOBBIES, BURNOUT, WORKAHOL, and COMPLEX were powerful enough to include in the screening for factors in the physician time function.

FACTORS NOT CORRELATING WITH OUTPUT

In the range of this study, with a power level of better than 75% to pick up any effect with an output elasticity of 0.2, the variables DEBT, SAVINGS, OFFICE, FELLOW, and ANCILL% are irrelevant to medical services output either directly or by influencing physician time input. In addition, GENDER, HOBBIES, BURNOUT, and COMPLEX are not associated with physician time spent seeing patients.

CHAPTER V. VALIDITY

In order to draw conclusions, researchers and businessmen alike require accurate, reliable data and valid results. The accuracy of the output variables was discussed at length in the section on preliminary data analysis, beginning on page 67. The accuracy of production in relative value units is about plus or minus 3 percent. Accuracy of the independent variables was checked by screening outliers, as discussed on page 60 in the section on data analysis. Miscalculations and transcription errors were corrected.

RELIABILITY

All output for a given time period was recorded for each physician in this study. Thus, for the dependent variables, reliability is really a question of stationarity. How long must production data be collected to obtain consistent results? The stationarity of the cross-sectional production data was estimated for collection periods of 1, 3, 6, 9, and 11 months. Using the 11 month collection time as the gold standard, how reliable were data collected over shorter intervals? The results are shown in Table V-1 on the next page. A four month data collection interval is a bare

TABLE V-1. STABILITY OF SHORT INTERVAL PRODUCTION DATA COMPARED TO ANNUAL OUTPUT

TIME PERIOD	COEFF DETERM	LARGEST CHANGE	T VALUE
1 Month	0.74	71.3%	11.4
3 Months	0.86	25.7%	17.1
6 Months	0.95	25.1%	32.8
9 Months	0.99	10.2%	75.2

minimum, but data collected for six or nine month intervals seem nearly stationary.

Reliability is also an important consideration for certain independent variables, namely PATTIME, where sampling and judgment about physicians' time input took place, and the relatively subjective, self-reported variables, such as HOBBIES, BURNOUT, EXERCISE, and WORKAHOL. The remaining independent variables which have adequate power have essentially no subjective or sampling components.

Repeating the PATTIME sampling and data assembly was impractical. Initial sampling was done for two distinct one week intervals. Care was taken to be sure that the Monday through Friday schedules sampled were not distorted by unusual events, such as sickness, court testimony, out of town clinics, and so forth. If something unusual was evident, that day was checked in a different week. The methods for determining PATTIME required less subjective physician input than the estimates of time obtained by questionnaire in previous large-scale econometric studies. The attention to detail in this study suggests that the reliability of time input is better than previous studies of physician production. However, the actual reliability for PATTIME is unknown.

Of the remaining "subjective" independent variables, only WORKAHOL and EXERCISE appear in the production functions. The original WORKAHOL question was repeated two months after the first interview. The results, denoted WKHOL2, had a correlation of 0.62 with WORKAHOL. Several of the professional staff asked if the rating scale is one of "absolute workaholism" or if the rating should be scored relative to other physicians. Future investigation of the effects of self-rated workaholism on productivity need to focus this question more sharply. The reliability of EXERCISE is unknown, but it is likely to be higher than WORKAHOL.

The validity of the results is a more complex question than their accuracy or reliability. Validity also needs explict attention. To organize this discussion, validity will be divided into statistical conclusion validity, internal validity, construct validity, and external validity, as defined earlier on page 63.

STATISTICAL CONCLUSION VALIDITY

Table V-2 lists several production function validity measures. The first column lists the four production functions. Each of the three medical services production functions was done for three output variables, which also are listed in the first column. Column two lists the coefficients of determination for each of the three versions of the three medical services production functions. Note that the coefficients are very similar, except for the higher coefficient with $TTLPROD for the predicitive production function. The higher coefficient may occur because the main predictor, SPECIALTY, is an index of national mean gross

TABLE V-2. PRODUCTION FUNCTION VALIDITY MEASURES

FUNCT	AJ COEFF DETERM	COEFF VAR	STP REG FACTORS	MSE	SSE	PRESS
PLANNING						
TTLPRD1	0.6083	0.2559	5	1.60	75.1	97.7
$TLPRD1	0.6194	0.2639	4	53.07	2494.2	3211.6
MDSGRV	0.6157	0.2728	5	1.19	56.1	70.4
PREDICT						
TTLPRD1	0.5881	0.2624	4	1.68	80.6	102.5
$TLPRD1	0.6754	0.2437	4	45.26	2172.4	2787.5
MDSGRV	0.5829	0.2842	3	1.30	62.2	82.1
REINHRDT						
LNTLPRD	0.6277	0.1678	6	0.06	3.0	4.1
LN$TPRD	0.5762	0.0871	4	0.08	3.6	5.0
LNMSRV	0.5957	0.2048	5	0.08	3.3	4.4
TIME						
PATTIME	0.2280	0.2538	2	1204.52	54203.3	65764.3

revenues for various specialties rather than an index of mean production in relative value units.

Column three in Table V-2 lists the coefficients of variation for each of the output variables for each of the production functions. The coefficients of variation, defined as the standard deviation of the outputs divided by their means, allow comparison of the variations in estimation using different output variables. The results are similar within each of the production functions. The small differences in the coefficients of multiple determination and small differences among the coefficients of variation in the models support the accuracy of the "quick fix" correction procedure for relative value units, which was described earlier on page 68. Further support for the "quick fix" and the accuracy of the relative value measurements comes from the number of independent variables accepted in forward stepwise regression (column 4) and the ratio of the sum of squares error to PRESS results (columns 6 and 7).

The power of this study to detect production factors for medical services and physician time deserves comment. Power was inadequate for 16 of the 35 original variables. The reason for low power for these variables is predominantly the limited population size for this study; to a lesser extent, lack of reliability of some of the questions may have caused larger than necessary variability of responses, which could also have limited power. A clinic of 100 or 200 physicians would yield more powerful results. Nevertheless for 19 variables, power exceeded 60% for both medical services and time. For an additional 5 variables, power was adequate to regress the variables against time. Power was not determined for several of the cross-product or power terms tested for the Cobb-Douglas-Reinhardt production function. However, the cross-product and power terms were combinations of factors which did demonstate adequate power in the simple linear regressions. For the factors which were included in the production functions, power was excellent, usually above 95 percent.

The screening process for variables to include in the Reinhardt-Cobb-Douglas production function did not guarantee an optimum combination of input factors. Nevertheless, it is impossible to apply optimizing algorithms given the low ratio of subjects to the total number of variables. Compared with previously published physician production functions and the linear production functions from this study, the coefficient of determination of the Cobb-Douglas-Reinhardt function is in the right range. A study with a larger number of subjects could allow more confident variable selection.

The assumptions for the use of multiple regression techniques were met in this study. Diagnostic residual analysis for the important production factors from Table IV-4 revealed only very slight deviation from normality of the residuals. Collinearity of SPACE and RVLABOR is evident, since more personnel require more space, but the cross-correlations of other factors are acceptably low, as displayed in Table IV-5. Furthermore, even for SPACE and RVLABOR, the variance inflation factors were only 1.81 and 1.95, respectively. These were the highest for any of the variables in any of the production functions. The mean variance inflation factor for equation {13} is 1.42, for equation {14} is 1.11, and for equation {15} is 1.08, all of which are acceptably low and give no evidence of significant collinearity. The Cobb-Douglas-Reinhardt model does have lower cross-correlations than the linear models, so the elasticities of the variables can be expressed more confidently. Residual plots indicated no grouping of data points to suggest omitted variables or heteroscedasticity.

The statistical significance of the production functions was evaluated by F-ratios; T-values were used to check the significance of each of the individual inputs to the production functions. The results, listed in Table IV-4 on page 80, indicate that each of the production functions is highly statistically significant. The null hypothesis is extremely unlikey; physician production can be modelled.

The T-values indicate that most of the individual factors are significantly related to the models. The execeptions, EXPWMC and EXERCISE, nevertheless add more explanatory power to the models than would be expected from addition of a nonsense variable. Thus, judging from the statistical tests, each of the models shows significant covariation among the input and output variables. The magnitude of the covariation is expressed by the elasticities, which also are listed in Table IV-4.

INTERNAL VALIDITY

Are there cause and effect relationships among the input and output variables in this study? Multiple linear regression can only determine the presence or absence of association among variables; it cannot establish causation. Furthermore, the cross-sectional design of this study does not allow observation of changes in the independent variables over time, which is the usual method of establishing causation. However, the signs of the regression coefficients do throw some light on the interactions of the variables. As shown in Table IV-2 on page 74, most of the independent variables have regression coefficients with the expected sign, which does support the theoretical cause and effect relationships discussed in the review of production in Chapter II.

Consider, however, the discordant results of independent variables which have power over 60% shown in Table IV-2. SAVINGS correlates positively with production, rather than negatively. The following hypothesis may apply: those physicians with higher production have more disposable income, so they are more likely to save. Simon's theoretical explanation that money in the bank discourages production appears to be incorrect in the range of savings studied here.[155] EXERCISE has a weak negative correlation with output only in the simple regression function, but in the multiple regression analysis which produced the Cobb-Douglas-Reinhardt production function, EXERCISE has the expected sign.

As far as the independent variables for physician time, the signs of the regression coefficients diverge from the theoretical predictions for HOBBIES, FELLOW, BURNOUT, and COMPLEX. Hobbies, by improving sense of wellbeing, may increase time in patient care rather than diminishing it. Those who spend more time at work may well feel more burned out. Those with more complex practices may have to take more time out for relaxation. However, why more academically inclined physicians who have achieved fellowship status spend less time at work is puzzling. Overall, the signs of the regression coefficients support the expected relationships of independent and dependent variables and provide plausible support for causative relationships among the variables.

The factors which do appear in the final production functions are consistent with standard economic theory, as will be further discussed in the section on external validity. Specifically, all three production functions include physician time input, which is essential for any professional production by definition. The planning production function and the Cobb-Douglas-Reinhardt function contain capital equipment, non-physician labor, and space, all of which seem logically related to output. Increasing any or all of these three factors would be expected to increase a physician's production. It is possible that the relationships sometimes operate in the opposite direction, however. A more productive physician may want more offices or support personnel because it makes life easier. In the clinic environment, however, space is allotted primarily on how well it helps patient flow. Support personnel are a direct expense to the physician, who has more take-home pay if workers are used efficiently. Thus, alternative explanations of causation or intermediary effects seem less likely than the mechanisms described earlier. Of the remaining factors, continuing medical education brings in new ideas and enthusiam, which apparently boosts productivity more than the time off work diminishes it. The availability of significant outside income might logically diminish the desire to work. The beneficial effects of exercise on overall performance hardly need emphasis here.

SPECIALTY is the main factor in the predictive production function, equation {14}. Surgical specialists require more training and are subject to more risk than primary care physicians. It comes as no suprise that their output is different. Specialty has cross correlations of -0.21 with non-physician labor, .69 with capital equipment, and .19 with continuing medical education. Thus, the factor SPECIALTY in the predictive equation subsumes the non-physician labor, capital equipment, and continuing education factors in the planning production function.

Although the relationships of the variables may seem logical from economic theory and from experience, that does not prove that changing the inputs will actually change the level of output. For example, a workaholic lives to work by definition, but he or she can fall prey to the "activity trap," which may explain why the correlation of workaholism with production is not stronger.[156] Conversely, much time put in on the job is necessary to consider oneself a workaholic. In sum, the theoretical cause and effect relationships of variables seem to be credible; discrepancies in signs are easily explained; and the factors in the final production functions may be causally related to output.

CONSTRUCT VALIDITY

Do the dependent and independent variables measure what they are supposed to? For the dependent variable, this question was considered at length in Chapter III about methodology. In addition, some of the results from this study have bearing on construct validity. First, consider the various ways of measuring output. Patient visits are inadequate measures of output, because visits are heterogeneous. This is demonstrated graphically by Figure IV-1 and by the differences in regression results between primary care and procedure-oriented specialists.

Substitution of laboratory tests or X-rays for physician services is theoretically possible. Indeed, as shown in Table IV-3, the cross-elasticities of laboratory and physician time are positive, indicating that laboratory services can substitute for professional services to a small extent. However, the major effect of physician time input is to generate laboratory services, which is really a complementary effect, according to Reinhardt's argument. Both the sign and the size of the elasticities indicate that radiology services are complements to physician services. That physician time, laboratory services, and radiology services are almost totally complementary indicates that for primary care physicians, output can be measured either by total services or by medical services exclusive of laboratory and X-ray services. This is demonstrated also by the cross-correlation of 0.93 between MDSRGRV and TTLPRD11 in Table IV-1. It therefore seems likely that the inability to measure ancillary services ordered on hospital patients by the physicians in this study is not a major defect. From the point of view of clinic management, the total production figure is probably the most useful, so it was used.

At the Western Montana Clinic in 1990, production measured in financial units and in relative value units has high correlation, as expected. Furthermore, the production functions from this study have similar input factors, similar regression coefficients, and similar coefficients of determination no matter if financial or relative value output measures are used, as shown in Table IV-4. The independently derived McGraw-Hill and Harvard resource based relative value scales each have a strong theoretical foundation and they correlate highly with each other, as shown in Table III-3. All of these observations give empirical support that production measurements in relative value units have construct validity.

As far as the time variables, PATTIME was scrupulously and exhaustively determined, as already discussed. Table IV-6, which shows the cross-correlations of the time variables, yields additional insight on how well they represent time spent by physicians in health service activities. Time absent on vacation or at meetings correlates positively with time on call. Apparently, those doctors on call more frequently need more rest and relaxation; they can afford more vacation and meeting time. Logically, time absent correlates negatively with total time spent in health service activities and negatively with outpatient appointment density. Call time has a positive correlation with total time spent with patients. However, call has a negative relationship to outpatient appointment density. This negative relationship may indicate that those on weekend and holiday call take more time off during the work week; more likely, it reflects that surgical and procedure-intensive specialists have more call but spend relatively more time in hospital than primary care providers. Outpatient appointment density correlates positively with total hours spent rendering medical services, as expected. That all time cross-correlations have appropriate signs and magnitude supports that they do correctly measure time spent providing health services.

Now consider the construct validity of the independent variables which have significant power. As already discussed, the signs of their simple and multiple regression coefficients fit theory and logic, which lends some support to their construct validity. From the group of factors which are irrelevant to production or time at an output elasticity of 0.2, DEBT, SAVINGS, OFFICE, FELLOW, ANCILL%, and GENDER were defined unequivocally and have straighforward interpretations. HOBBIES and BURNOUT were not defined explicitly, but are commonly understood terms. The means and variances of these factors seem the right order of magnitude. COMPLEX is defined unambiguously. Whether it is a proxy for interesting medical activity which could engender "flow" may be open to discussion: undoubtedly some physicians prefer the routine to the "interesting case." However, it does appear by definition to be a measure of medical complexity in a practice.

Of the factors in the production functions, time has been discussed. A basic assumption for LABOR is that salaries and benefits correlate highly with the skills of non-physician workers. Certainly, starting salaries for receptionists, aides, nurses, and physician's assistants reflect their initial training and skills. Additional skills acquired on the job are reflected in seniority pay to a greater or lesser extent. However, the assumption that the level of non-physician labor use is similarly efficient for all of the physicians in the study is one weakness of this measure of labor. It does not appear to be a major problem, but its magnitude is unclear.

SPACE, the number of offices and examination rooms per physician, has been used by many economists as a proxy for capital investment. As shown in Table IV-6, the correlation of SPACE with other factors involving capital investment is negligible. SPACE does correlate positively with RVLABOR, as shown in Table IV-5, and it appears in all of the medical service production functions. A common sense interpretation of SPACE seems appropriate. More examination rooms allow patients to be seen more efficiently, so SPACE correlates positively with output. Additional helpers require room to work, so SPACE is highly correlated with RVLABOR.

Is EQUIP a good measure of capital investment or technology or both? The total replacement costs of equipment used daily certainly is a strong measure of capital investment. EQUIP assumes that all equipment is newly purchased, but medical equipment does depreciate rapidly, so that assumption is realistic. EQUIP is logically related to technology as well, but it does not measure changes in basic understanding of disease mechanisms. Basic understanding of diseases makes medical advice more valuable; such basic understanding does come at a cost. This dilemma, discussed in the section on substitutes and complements on page 45 in Chapter III, may be ameliorated by the use of CME variable.

Does CME include information other than high interest in keeping up to date in one's medical specialty? Probably not, since there is little cross-correlation with other production factors. Leaving to attend a national conference does take time away from medical practice, which accounts for the negative cross-correlation with PATTIME in Table IV-5. This negative effect of CME on PATTIME and hence on production is more than compensated by the benefits of CME on production. Certainly CME and its regression coefficient underestimate the importance of continuing medical education other than national meetings. Local meetings and televised courses can bring substantial new information to many physicians without loss of time from practice. The positive relationship of CME with EQUIP bespeaks the interaction of new ideas and technology; the negative correlation with SPACE in equations {13} and {16} is enigmatic.

The remaining factors in the production functions, namely EXPWMC, OUTSIDE, EXERCISE, WORKAHOL, and UNDER seem clear. That a more explicit definition of workaholism and an independent rating of its severity would yield more a reliable and valid measure of this concept seems likely.

Consider now the construct validity of the production functions. All production functions are approximations; they are simplified models of reality. Simplification is necessary in the welter of events which surround any manager, but clearly some models have more usefulness than others. Linear functions are the clearest. However, because of collinearity, it is difficult to include the interactions among production factors which are present in the real world.

Recall the discussion of the production surface in Chapter IV. The marginal rates of technical substitution for the factors in the linear production functions are constants. However, a graphic display of the isoquants of physican time and non-physician labor (Figure IV-2 on page 91) shows that the isoquants are anything but linear and smooth: they change constantly over the production surface and they have strong interaction effects.[157] Because the Cobb-Douglas-Reinhardt production function, equation {15}, includes all of the important factors from the linear equations, includes interactions between factors, and has the lowest coefficient of variation of production (see Table V-2), it is a better model than the linear functions.

Finally, whenever one uses linear regression for prediction, one must be cautious about extrapolation beyond the range of the data used to build the model. Use of these production functions probably will not be valid for estimating productivity of physicians in solo practice or physicians on a salary which includes no productivity incentive. Even within these constraints, the functions may not be applicable to clinics where the total number of professionals is greatly different from 58, which is the number of physicians on staff at the Western Montana Clinic. The reasons for emphasizing this range are decreasing returns to scale of production as clinic size increases, which was discussed earlier, and the consultations generated in a large organization which would not occur in a solo office or a small group.

EXTERNAL VALIDITY

Do the results of this study have applicability to other multispecialty group practices and to larger questions about production of medical services? There are reasons to be cautious about any extrapolations to other settings. Previous studies have shown that variations in the C factors, which are not germane to a single-clinic study, will affect the output in other clinic settings. C factors include clinic size, demand for services, demographics of the clinic service area, prepaid/fee-for-service mix, salary incentives for physicians, and "corporate culture," among others.

Evidence supporting the external validity of the production functions derived in this study includes comparisons with previous research. The production input factors of time, space, capital investment, and non-physician labor identified in prior investigations were found to be important in this research as well. Compare the factor elasticities from earlier reports in Table II-2 with the production factor elasticites from this study in Table IV-4 to see that for time and non-physician personnel they are very similar. The production elasticity of office space in this study is considerably higher than the single preceding report which used data obtained from clinics. The elasticity of capital investment is somewhat higher than the single report which used data from solo practice, but that data is nearly 20 years old and the divergence is not large.

The PRESS results listed in Table V-2 also indicate that the production functions have validity. For the Cobb-Douglas-Reinhardt model, PRESS is 4.1, which agrees well with the sums of squares error for the model of 3.0 and suggests that the means squares error of 0.06 for total production is a good indicator of the predictive capability of the model. The PRESS results for the linear models listed in Table V-2 also look auspicious. Nevertheless, even with the best form of the Cobb-Douglas-Reinhardt production function, there is an 8.7 percent coefficient of variation; using the linear production functions, the coefficients of variation are around 25 percent. How this degree of uncertainty in prediction compares with prior reports is unclear, since earlier authors have not disclosed them.

A third measure of external validity is to apply the predictive production function, equation {14}, to data from other clinics. This method of external validation must be interpreted cautiously for several reasons. The C factors may vary substantially, as just discussed. Some information comparing the other participating clinics with the Western Montana Clinic is shown in Table V-3. Geographic locations are quite diverse and there is a range of sizes. None of the clinics does much prepaid medical care. For the three clinics providing data, the degree of incentive pay for physicians seems equivalent: those on "100% incentive" do less cost accounting, which means that more expenses are born by the higher producers. However, one of the more important determinants of production, demand for services, cannot be quantified easily for different service areas, and this may substantially impact the results. Another major reason for caution about the external comparisons is the small sample size and lack of random sampling of either clinics or physicians. To be using PSP, to have management information sufficient to answer the questionnaire, and to be willing to put in the staff time to collect the data for this study requires a high level of management sophistication, which meant that clinics could not be chosen at random. The physicians were stratified by specialty, but to ask for a random sample within the specialty groups was beyond practicality. Furthermore, the PSP data could not be checked for accuracy as the Western Montana Clinic data were. As Reinhardt stated in 1975: "Detailed information on the rates of individual services produced by physicians is rarely

TABLE V-3. EXTERNAL VALIDATION SOURCES

CLINIC	WMC	LEX	MCC	WFC
SIZE	58	110	135	48
PREPAID	2.0%	-	4.5%	0.0%
INCENTIVE	85%	-	100%	100%
COST ACCT	50%	-	31%	15%

available to outside researchers."[158] This limitation on good data has not improved.

Data from the other multispecialty clinics was collected by questionnaire. Instructions and encouragement were done by phone. Sample questionnaires are included in the appendix.

The data from the Lexington Clinic were limited to PSP production in relative value units, specialty, and years worked at the clinic for 10 physicians in various specialties. This data was insufficient to carry out the mean square prediction error calculations. However, a multiple linear regression of specialty and experience on output was done. The model was not statistically significant, most probably due to the small sample size. From the regression, production elasticies of specialty (0.53) and experience (0.19) were determined. They compare favorably with values of 0.56 and 0.06 for WMC, as shown in Table IV-4.

Data from Pensacola on 15 physicians were more complete, but the output is in dollars rather than relative value units. After putting the predicitive production function, equation {14}, into a form so that the output is in dollars, the mean square error (MSE) for the regression was 45.3 and the predicted mean square error (PMSE) for the data from the Medical Center Clinic of Pensacola was 316.0. The major difference between the two results indicates that the predictive production function is of little use in Pensacola.

Data from the Wichita Falls Clinic on 15 physicians in various specialties were complete and the output was relative value units. However, the difference between the MSE (1.68) and the PMSE (24.5) was again very large, making the predictive function of little use in Wichita Falls either.

In addition to the possible reasons for the failure of the production function in other clinic settings listed above, collinearity of the independent variables deserves attention.[159] However, by all measures, collinearity is very low for the predictive production function.

The data which were provided from both clinics contained several physicians who produced more than double the revenues or relative value units of the leading Western Montana Clinic physician. Such high producers are certainly beyond the range of prediction of the production functions derived in this paper. Whether applying the methods used in this study to the other clinics will yield valid predictive or planning functions remains to be determined. The divergent evidence on external validity of the production functions from this study necessitates that any generalization be made with extreme caution.

CHAPTER VI. DISCUSSION OF RESULTS

The mathematical notation and the jargon of economics make it seem more dynamic than it is. "Elasticity" sounds vibrant. Yet a cross-sectional, descriptive analysis like this one is a still life. Nothing moves. Production functions derived from static observations presume that one physician can be transformed into another, which is preposterous.

Interpreting and applying the production functions demand a rich context of experience in health care. The amount of variation in production associated with various input factors has meaning only when seasoned by observations of the evolution of medical services over time. A background in management of physicians or a clinic is essential to make use of the output elasticities. The amount of variance not explained by the production functions and the failure of external validation simply show that important production factors in the health services setting remain poorly characterized.

Some prior explorers in physician productivity used aggregate data to show that the principal components of manufacturing productivity, namely capital, labor, time, and technology, could be used to analyze health services outputs. They showed that these traditional inputs account for between 40 and 60 percent of observed production. In contrast, this study focused on careful measurement of inputs and outputs. It demonstrates that the unexplained variance in physician production is not an artifact of crude measurement; rather, there are significant factors still lurking for future explorers to capture.

The present study fell short for two reasons. First, the population of physicians was too small. The expectation that precise scrutiny of an array of input factors, including a detailed set of "physician tastes," might better expain productivity was incompletely realized. Some of the problem was the low power of a number of the independent variables. Extending the regression analysis to physician time did allow consideration of more possible indirect inputs to production, but study of a larger clinic is a better solution to the problem of insufficient power. Factors with variances of the magnitude found in this study require a clinic of 100 to 200 physicians to do powerful regressions.

The second defect lies in the basic design of this investigation, which, like all of its predecessors, is descriptive. There is no conclusive evidence from this or any prior cross-sectional study about how manipulation of the inputs will change production. Re-examination of this clinic after several years in which time, capital, non-physician labor, space, and other inputs change could give a better indication of whether changing these inputs will really affect output of medical services.

But take a more optimistic view of this "dismal science." The methodology in this investigation is clearly better than that in prior studies. Expressing output in relative value units, looking carefully at details of physician behavior, assessing the power of the independent variables, and attempting external validation of the production functions have not been part of previous studies of physician output. A medical clinic can provide a laboratory to uncover some of the cause and effect relationships which larger econometric studies never will establish.

IMPLICATIONS FOR IMPROVING PRODUCTIVITY AT WMC

As in previous large scale reports of physician productivity, the important determinants of output of physician services at the Western Montana Clinic appear to be physician time, non-physician labor, capital equipment, and space. Minor determinants, some not previously associated with production, include continuing medical education, experience, and exercise. Income from non-medical sources has a minor negative association with production. Workaholism and more extensive office visits influence productivity indirectly through physician time input.

TIME. Time input is an important determinant of production. A 10% increase in time spent seeing patients may result in about a 6% increase in actual output of medical services. While interviewing professional staff for this study, it was evident that they do not have good information about how much time they were working compared to previous years or compared to their peers. All considered themselves "overworked and underpaid," yet this clearly was not the case when judged by national statistics or by the clinic physicians in the external validation sample.

Simple reports of time absent for vacation and meetings or of time spent in the office were made at WMC in the past, but such reports do not yield enough insight to be of much use. A useful feedback system for time spent with patients will have to be more complex, but computerization can simplify data collection. The information may well be worth the expense of setting up a monitoring system.

This study also indicates that there is considerable slack time in some physicians' workdays. Additional time in the office results in different increases in output for different physicians. There are at least two reasons for the variability: 1) some medical services are provided but billed, and 2) some physicians work more efficiently than others. Coaching about time management and billing may achieve notable production gains. Techniques for increasing efficiency are available and appear promising, even though they have not been shown to improve productivity in controlled studies (see pages 20 and 21).

NON-PHYSICIAN LABOR. Use of non-physician labor may be an important determinant of productivity. A 10% increase in office space/helpers could yield a 4% increase in output of medical services. Even though economists such as Uwe Reinhardt have for decades emphasized the importance of non-physician labor, physicians have been very slow to employ an optimum level of assistants. At the Mayo Clinic, which does a lot of clinical research as well as a lot of patient care, there are about nine non-physician staff for each doctor. Reinhardt estimated that the optimum level of assistants in a solo practice is about four. At WMC the current staff to doctor ratio is 3.1, which appears to keep "overhead" down, but may be limiting production even more. There are a number of barriers to optimum use of non-physician manpower, which include licensing restrictions, lack of training in delegation and management, jealousy, and fear of increased complexity. Overcoming these obstacles in a clinic should be relatively easy, as physicians observe the success of their colleagues' assistants. By assuming some of the risk of employing assistants, the clinic can encourage greater use of physician assistants in appropriate settings and discover if productivity gains do accrue.

CAPITAL EQUIPMENT. "Capital equipment" consists of two segments: capital investment and innovation. In the long term, both may be equally important for productivity gains.

Certainly it is hard to measure one without measuring the other. Although a 10% increase in capital equipment investment may yield just a 2% increase in medical services output, capital investment does not have the same limits as the 24 hour day. This study examined capital investment not simply within WMC, but considered all capital equipment available for the physician's use on a daily basis. This perspective emphasizes not only thoughtful investment of WMC funds, but careful planning together with other institutions, especially hospitals, about optimum acquisition of capital equipment.

Another important implication is that, to sustain a competitive position, some of the gains from any business must be reinvested. Despite high incomes, physicians traditionally live from hand to mouth. Some reach retirement with no savings. Others plan poorly and then try to make up for lost time with "get-rich-quick" schemes. Thus, instilling a philosophy of retaining earnings to invest in clinic operations faces resistance. The hurdles of retained earnings and cooperative investment can be jumped only by a farsighted executive committee.

The second part of capital equipment, technologic innovation, needs emphasis as well. Champions and critics alike focus on high technology scanners and open-heart surgeries. In a clinic, the more important technologic decisions may be more sublte. For example, consider the decline in cardiovascular mortality which can be gained by correction of risk factors, discussed on page 23. The issue is both quality and producitivity. Physicians, like all humans, deal poorly with masses routine details; computers handle the details well. McDonald showed 15 years ago that physicians accept computer reminders well.[160] Such reminders, made part of an electronic medical record, will satisfy patients' requests for preventive medicine and improve their health. At the same time, the reminders will increase utilization of examinations, laboratory tests, and X-rays. An electronic medical record is a technologic innovation and it requires substantial investment. The heart attack avoided will not make newspaper headlines, but the process does represent a notable step forward both for quality and productivity.

SPACE. Enough examination rooms to assure efficient flow of patients seems a prerequisite of productivity. How best to create the space is a different matter. Locations for patient convenience, such as at a shopping center, or for physician convenience, for example, near a hospital, must be balanced. Getting physicians to schedule office time in a manner to allow shared offices represents a low-overhead approach, but it is difficult to implement.

MINOR DETERMINANTS AND IRRELEVANT FACTORS. Continuing medical eduction is probably related to capital equipment and technology; it provides a small but definite boost to production. Retaining physicians and encouraging a healthy lifestyle surely make sense. However, limiting hiring by certain personal characteristics, such as male gender, high levels of medical school debt, or a non-working spouse can make at most a small impact on productivity. Although conventional wisdom points to the importance of such factors, practical and legal constraints plus the results of this study demonstrate that selective hiring based on these factors will not dramatically change output.

An important task of WMC management is to improve productivity. Increased output must be accomplished within constraints of professional culture, patient expectations, and government regulations. The production functions point out which pathways most deserve attention. The more sophisticated Reinhardt function also suggests that relatively small increases in the production factors will result in disproportionate increases in output, that is, positive returns to scale. The implication is that if management pushes ahead with well-considered initatives, WMC will prosper.

IMPLICATIONS FOR OVERALL HEALTH CARE PRODUCTIVITY

Despite assertions to the contrary, physician productivity certainly can be improved through technology and more efficient use of existing resources. Pope purports to show that physician productivity did not improve from 1976-

1986, but his measures of output, namely dollars spent on patient visits and the number of hospital days, do not encompass the work being done.[161] When the Harvard resourse-based relative value scale is in place in 1992, more sophisticated measurements of productivity for Medicare patients should be possible to clarify the productivity issue at a national level. The current study demonstrates the increased precision of the relative value measures of productivity over the crude measures of patient visits and financial measures. In a longitudinal study of production, changes in relative value units per dollar charged, if adjusted for inflation, could provide a way to take the content of the service into account. Nevertheless, even the relative value system will not measure some aspects of medical care: the quality the advice given in a brief office visit, for example.

Baumol's basic argument that physicians' productivity cannot be increased compares the physician's work to that of a musician playing a Scarlatti sonata on the harpsichord.[162][163] It took Colin Tilney 4 minutes and 31 seconds to play the cantabile andantino in D major, K 277, in 1988, which is the same time it must have taken Scarlatti himself to play it for King Joao the Fifth in Lisbon in 1738, 250 years earlier. Baumol would argue that there is no increase in Tilney's productivity over Scarlatti's. Yet I can hear Tilney play his performance over and over on a compact disk, and so can billions of other people, if they choose. Tilney's output for his performance must therefore be enormously greater than Scarlatti's performance for the King.

Returning to medicine, opinion polls have shown that citizens of this country value their medical care highly.[164] Demand for medical services will increase as the population ages and as long as patients do not have to pay the costs directly. With increasing numbers of physicians graduating from medical schools, service to patients will improve and medical costs as a share of the gross national product will continue to rise.

In this situation, if the government limits technology and encourages no-deductible insurance coverage, costs for medical care will skyrocket and productivity will languish. In contrast, if appropriate use of technology is encouraged and patients must pay part of the costs, competition in the medical sector will increase. Overall costs, adjusting for demographic changes, may stabilize. The value of this production study in the debate about increasing productivity of physicians is to show that methods determine results: careful measurements will show that productivity can improve and will point out the best ways to accomplish the improvement.

CHAPTER VII. SUMMARY AND CONCLUSIONS

This cross-sectional study probes factors associated with the productivity of 53 physicians, podiatrists, and psychologists at a multispecialty medical clinic. Output was measured in relative value units; 45 potential inputs were screened by multiple linear regression for importance. Particular care was used to look at details of physician behavior, measure the power of independent variables, and assess external validity of the results. The results confirm earlier research by demonstrating again the importance of physician time, capital equipment, non-physician helpers, office space, and physician experience in the production of medical services. Several minor determinants, such as continuing medical education, exercise habits, outside income, and workaholism, are weakly associated with output. Some factors, for example, gender, debts, complexity of work, and level of burnout, appear irrelevant to physician productivity. Some major factors remain to be discovered.

The principal result of this investigation is to emphasize that carefully designed research in a clinic setting has promise for dispelling some of the uncertainties which surround physician productivity.

APPENDICES

INDEPENDENT VARIABLE WORK SHEET FOR PROFESSIONAL STAFF

NAME: NUMBER:

I. TIME INPUT

*VAC [vacation and meeting time by WMC data and doctor estimate]:

CALL:

DENSITY:

II. NON-PHYSICIAN LABOR:

III. CAPITAL INPUTS

PROC:

*SPACE [number of offices and examination rooms used]:

*EQUIP [replacement value of equipment used at least once per day]:

IV. PHYSICIAN CHARACTERISTICS

*AGE:

*GENDER:

*EXP [how many years in practice?]:

*MARITAL:

*DEPEND [claimed on tax form]:

WAGE RATE [PSP intensity]:

*OUTSIDE [working spouse or comparable outside income]:

*DEBT [loans + mortgage > 75K]:

*SAVINGS [three months income = 1, less = 0]:

*PRIORITIES [list your three main activities other than medical practice. Rate the highest 1 --> lowest, 4. Take medial score.]:

*HOBBIES [hours per week]:

*OFFICE [previous solo?]:

*FELLOWSHIP IN SPECIALTY SOCIETY = 1, NONE = 0:

*CME [a national specialty meeting in 1990 = 1]:

REPUTATION [rating of desirability as own physician by WMC employees. 3 = I myself would see that doctor without hesitation for a medical problem --> 0 for I'd never see that doctor]:

*BURNOUT [0 for none -->10 for only ashes left scale]:

*SELF-ESTEEM IN MEDICINE [rate yourself on a 1 = 1st year resident --> 10 = world class for your specialty]:

*HEALTH [rate your own health on a 1 for nearly dead to 10 for perfect health scale]:

*EXERCISE [hours of aerobic exercise per week]:

*FATIGUE [by the end of most days in the office, I'm exhausted = 4, beat = 3, have some enthusiasm = 2, happy to come back in if they call me = 1]:

*WORKAHOLISM [funhog = 1, workaholic = 10]:

CHALLENGE [PSP complexity]:

*FLOW ["My interest and training are appropriate for nearly all the patients I see" = 4, "most" = 3, some = 2, very few = 1, none = 0]:

V. ANCILLARY INCOME [PROPORTIONATE]:

* = need to ask the staff member about this item directly.

INDEPENDENT VARIABLE WORKSHEET #2 FOR PROFESSIONAL STAFF

NAME: NUMBER:

I. TIME INPUT

ROUNDS [HOW MUCH TIME DID YOU SPEND MAKING HOSPITAL ROUNDS -INCLUDING READING EKGs, ETC., BUT NOT INCLUDING CATH, SURGERY, OR ENDOSCOPY - ON AN AVERAGE DAY 1N 1990?]:

WERE YOU ON VACATION BETWEEN SEPTEMBER 1 AND OCTOBER 15? HOW MANY DAYS:

HOW MUCH TIME DOES IT TAKE TO DO THE AVERAGE DELIVERY?

II. PHYSICIAN CHARACTERISTICS

ORGANIZED OUTSIDE INTERESTS [PLEASE ESTIMATE YOUR TIME SPENT PER WEEK IN ATHLETICS, COMMUNITY ACTIVITIES AND ORGANIZATIONS]:

WORKAHOLISM [FUNHOG = 1, WORKAHOLIC = 10]:

III. MISCELLANEOUS

VISSCHER CME [NATL SPEC MEETING?]:

KNAPP CH CATH
SPH ECHO

REPUTATION QUESTIONNAIRE

Reputation is a measure of public opinion about a doctor's skills and beside manner. Since you work at WMC, you are in a better position than most people to know about reputations of WMC doctors. Your opinion may come from personal experience, conversations with fellow workers, or any other source. Please rate each doctor listed below by writing a number from the following scale next to his or her name. All of the responses for each doctor will be totalled to give a reputation score. Your replies will be strictly anonymous.

0 = "I'd never see that doctor under any circumstances!"
1 = "I haven't heard anything good or bad to form a judgment."
2 = "Probably OK."
3 = "I myself would see that doctor without hesitation for a medical problem in his or her specialty."

PRODUCTION FUNCTION VALIDATION STUDY

1. LOCATION:

2. PHYSICIAN NUMBER:

3. SPECIALTY:

4. TOTAL GROSS PRODUCTION FROM 1990 (PSP reports a monthly average - please be sure that the average is based on at least 9 months of cumulative data. If you send figures from your accounting system, an annual total will be fine -- I'll convert it to monthly.)

a. PSP (MCGRAW HILL RELATIVE VALUE UNITS):
b. DOLLARS (FROM PSP):
c. DOLLARS (FROM YOUR CHART OF ACCOUNTS):

5. NUMBER OF EXAMINATION ROOMS AND OFFICES USED BY THE DOCTOR (PRORATE IF SHARED):

6. TOTAL YEARS THE DOCTOR HAS WORKED AT YOUR CLINIC:

7. PHYSICIAN TIME INPUT (Attention to detail is important here. Simple measures of time input do not correlate as well with production as a more complete determination of time spent. Use your management information when possible, but otherwise get the physician's best estimates. Not every blank will be relevant to every doctor, but fill in as many as possible.)

a. TOTAL HOURS THE CLINIC (OR DEPARTMENT OR SATELLITE) WAS OPEN IN 1990 DURING WHICH PATIENTS COULD BE SCHEDULED (Example: Monday to Friday, 8:30 to 5:30, all year is 9 hours * 52 weeks * 5 days, minus 10 holidays is 2340 - 90 = 2250 hours):

b. WORKDAYS ABSENT FROM CLINIC FOR 1990 (VACATION + REGULAR HALF DAYS OFF + MEDICAL MEETING + SICKNESS):

c. WEEKENDS AND HOLIDAYS ON CALL. TOTAL DAYS FOR 1990:

d. PORTION OF THE AVERAGE DAY DURING WHICH THE DOCTOR IS AVAILABLE FOR OUTPATIENT APPOINTMENTS (Example: 5/9 hours each day except 3/4 hours on Thurdays because afternoon off yields 23/40 = .575):

e. ESTIMATE OF DAILY TIME SPENT ON HOSPITAL ROUNDS IN HOURS:

f. BEST ESTIMATE OF DAILY TIME SPENT IN THE OPERATING ROOM (Do not include lounge time or transportation time):

g. BEST ESTIMATE OF DAILY TIME SPENT DOING PROCEDURES IN THE HOSPITAL:

h. NUMBER OF DELIVERIES (FROM PRODUCTION RECORDS):

i. ESTIMATED TIME FOR AVERAGE DELIVERY:

j. MONTHLY TIME SPENT ON NURSING HOME ROUNDS:

k. MISCELLANEOUS TIME SPENT IN PATIENT CARE EACH WEEK:

PRODUCTION FUNCTION VALIDATION STUDY - CLINIC DATA

Clinic size affects productivity through economies of scale. The amount of pay based on productivity influences productivity, as does the prepaid/fee-for-service mix. To predict productivity of physicians, these clinic factors as well as the individual factors on the other questionnaire need to be taken into account.

1. NAME OF CLINIC:

2. TOTAL NUMBER OF PHYSICIANS/PODIATRISTS/PSYCHOLOGISTS:

3. PERCENTAGE OF REVENUES FROM PREPAID SOURCES:

4. PERCENTAGE OF PHYSICIAN PAY BASED ON PRODUCTION:

5. PERCENTAGE OF EXPENSES WHICH ARE DIRECTLY ATTRIBUTED TO PHYSICIANS:

REFERENCES

1. Fuchs VR. The health sector's share of the gross national product. Science 1990;247:534-8.

2. Stout H. U.S. spending on health care keeps growing. Wall Street Journal 1991 Apr 23;Sect. B:4 (col. 5).

3. Fuchs VR, Hahn JS. How does Canada do it? A comparison of expenditures for physicians' services in the United States and Canada. New Engl J Med 1990;323:884-90.

4. Donabedian A, Wheeler JRC, Wyszewianski L. Quality, cost, and health: an integrative model. Medical Care 1982;20:975-92.

5. Barsky AJ. The paradox of health. N Engl J Med 1988;318:414-8.

6. Grossman M. The demand for health: a theoretical and empirical investigation. New York: National Bureau of Economic Research, 1972.

7. Fuchs VR. The health economy. Cambridge (MA): Harvard University Press, 1986:181-99,272-99.

8. Feldstein PJ. Health care economics. New York: John Wiley and Sons, 1988:15-32.

9. Newhouse JP, Friedlander LJ. The relationship between medical resources and measures of health: some additional evidence. J Hum Resour 1980;15:200-218.

10. Marder WD, Zuckerman S. Competition and medical groups. A survivor analysis. J Health Econ 1985;4:167-76.

11. Raelin JA. The clash of cultures. Managers and professionals. Cambridge (MA): Harvard Business School Press, 1985.

12. Baumol WJ. Price controls for medical services and the medical needs of the nation's elderly. Chicago: American Medical Association, 1988.

13. Hurdle S, Pope GC. Physician productivity: trends and determinants. Inquiry 1989;26:100-15.

14. Reinhardt UE. Physician productivity and the demand for health manpower. Cambridge (MA): Ballinger Publishing, 1975:63-9.

15. Hurdle and Pope, "Physician productivity."

16. Luft HS. Health maintenance organizations. Dimensions of performance. New Brunswick: Transaction Books, 1987:150-9.

17. Newhouse JP. The economics of group practice. J Hum Resour 1973;8:37-56.

18. Gaynor M, Pauly MV. Compensation and productive efficiency in partnerships: evidence from medical group practice. J Pol Econ 1990;98:544-73.

19. Gaynor M. Competition within the firm: theory plus some evidence from medical group practice. RAND J Econ 1989;20:59-76.

20. Held PJ, Reinhardt U, editors. Analysis of economic performance in medical group practices: final report. Princeton (NJ): Mathematica Policy Research, 1979.

21. Kimbell LJ, Lorant JH. Physician productivity and returns to scale. Health Serv Res 1977;12:367-79.

22. Manning WG, Liebowitz A, Goldberg GA, Rogers WH, Newhouse JP. A controlled trial of the effect of a prepaid group practice on use of services. New Engl J Med 1984;310:1505-10.

23. Ware JE Jr, Brook RH, Rogers WH, Keeler EB et al. Comparison of health outcomes at a health maintenance organization with those of fee-for-service care. Lancet 1986;1:1017-22.

24. Budd, GB. Productivity: the challenge for ambulatory services. J Ambulatory Care Manag 1988;11:1-12.

25. Deal TE, Kennedy AA. Corporate cultures. Reading (MA): Addison-Wesley Publishing Company, 1982.

26. Hurdle and Pope, "Physician productivity."

27. Gaynor and Pauly, "Compensation and productive efficiency."

28. Lorant JH, Kimbell LJ. Determinants of output in group and solo medical practice. Health Serv Res 1976;11:6-20.

29. Kletke PR, Marder WD, Silberger AB. The growing proportion of female physicians: implications for US physician supply. Am J Pub Health 1990;80:300-4.

30. Held and Reinhardt, "Analysis of economic performance."

31. Feldstein, "Health care economics," 189-92.

32. Simon JL. Effort, opportunity, and wealth. Oxford: Basil Blackwell, 1987.

33. Raelin, Clash of Cultures, 143.

34. Beer M, Spector B, Lawrence PR, Mills DQ, Walton RE. Human resources management. A general manager's perspective. Text and cases. New York: Free Press, 1985:416.

35. Frank RH. Passions within reason. New York: W.W. Norton and Company, 1988.

36. Maslow AH. A theory of human motivation. Psychological Rev 1943;50:371-396.

37. Perrow C. Complex organizations. A critical essay. 3rd ed. New York: Random House, 1986:79-118.

38. Csikszentmihalyi M. Beyond boredom and anxiety. San Francisco: Jossey-Bass, 1988.

39. Larranaga R. Calling It a Day. New York: Harper and Row, 1990.

40. Reinhardt UE. A production function for physicians' services. Rev Econ Stat 1972;54:55-65.

41. Reinhardt, "Physician productivity."

42. Held and Reinhardt, "Analysis of economic performance."

43. Lorant and Kimbell, "Determinants of output."

44. McCarthy TR. An economic study of the physician services market: the role of waiting time considered [dissertation]. Baltimore (MD): Univ. of Maryland, 1980.

45. Feldstein, "Health care economics," 180-6.

46. Gaynor, "Competition within the firm."

47. Gaynor and Pauly, "Compensation and productive efficiency."

48. Reinhardt, "Physician productivity," 66.

49. Baumol WJ, Blackman SAB, Wolff EN. Productivity and American leadership: the long view. Cambridge (MA): MIT Press, 1989.

50. Quinn JB, Baruch JJ, Paquette PC. Technology in services. Sci. Am. 1987;257:50-58.

51. Reinhardt UE, "Physician productivity," 105-7.

52. Fuchs VR, Garber AM. The new technology assessment. N Engl J Med 1990;323:673-7.

53. Linton AL, Naylor CD. Organized medicine and the assessment of technology -- lessons from Ontario. N Engl J Med 1990;323:1463-8.

54. Kane NM, Manoukian PD. The effect of the medicare prospective payment system on the adoption of new technology. N Engl J Med 1989;321:1378-83.

55. Reinhardt UE, "Physician productivity," 107-16.

56. Hurdle S, Pope GC. Improving physician productivity. J Ambul Care Manag 1989;12:11-26.

57. Ferraro KF, Southerland T. Domains of medical practice: physicians' assessment of the role of physician extenders. J Health Behav Sci 1989;30:192-205.

58. Cawley JF, Ott JE, DeAtley CA. The future for physician assistants. Ann Int Med 1983;98:993-7.

59. Luft HS, "Health maintenance organizations," 160-1.

60. Wennberg JE, Blowers L, Parker R, Gittelsohn A. Changes in tonsillectomy rates associated with feedback and review. Pediatrics 1977;59:821-6.

61. Fowkes FGR, Evans KT, Hartley G, Nolan DJ, Roberts CJ, Davies ER, Green G, Hugh AE, Power AL, Roylance J. Multicenter trial of four strategies to reduce use of a radiological test. Lancet 1986; 1:367-9.

62. Fowkes FGR, Hall R, Jones JH, Scanlon MF, Elder GH, Hobbs DR, Jacobs A, Cavill IAJ, Kay S. Trial of strategy for reducing the use of laboratory tests. Brit Med J 1986;292:883-5.

63. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectibility of man. New Engl J Med 1976;295:1351-1355.

64. Avorn J, Soumerai SB. Improving drug-therapy decisions through educational outreach. New Engl J Med 1983;308:1457-63.

65. Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? New Engl J Med 1989;321:1306-11.

66. Wennberg JE, Gittelsohn A. Dealing with medical practice variations. Health Affairs 1977;3:6-32.

67. Raelin, "Clash of cultures."

68. Fink DJ, Goldhar JD. Management for physicians: an annotated bibliography of recent literature. Ann Int Med 1980;92:269-75.

69. The MGMA is a trade association of single and multispecialty medical groups. Besides disseminating information about management of medical group practices, the MGMA does research in ambulatory health care, which includes collecting nationwide statistics on medical groups. The MGMA address is 104 Inverness Terrace East, Englewood, CO 80112-5306.

70. Medical Group Management Association Library Resource Center. Increasing productivity in group practice. Denver: The Association, 1987.

71. Sevel F. 8 ways to make medical offices more productive. Phys Manag 1984;24:204-223.

72. Reinhardt, "Physician productivity," 107-16.

73. Luft, "Health maintenance organizations," 160-1.

74. Hurdle and Pope, "Improving physician productivity."

75. Doherty H, Tekavec MM. Efficiency begins with delegation. Dental Manag 1983;23:56-7,60.

76. Greenfield AR. Physician productivity: a managerial challenge. J Ambul Care Med 1989;12:6-10.

77. Pierce A, Sweeney DR. Boost productivity through better scheduling. Physician Manag 1985;25:159-171.

78. McKenney PC. 12 ways to increase office productivity. Physician Manag 1987;9:65-71.

79. Womack GR. 10 ways to streamline your practice. Physician Manag 1983;23:264-268.

80. Mangan WF. Seeing more patients without shortchanging them. Med Econ 1985;62:74-81.

81. Coding tips for consultations, part I: initial. St. Anthony's CPT Coding for Physician Reimbursement 1989 Sept;1(10):1-3.

82. Coding tips for consultations, part II: follow-up. St. Anthony's CPT Coding for Physician Reimbursement 1989 Oct;1(11):1-2.

83. Sevel, "8 ways."

84. Owens A. Are you making every minute of your practice count? Med Econ 1981;58:66-71.

85. McCoy JL. Beating the clock -- time management for physicians. Med Group Manage 1986;33:46-51.

86. Chandler ET, Bloomfield RL. How to manage you time for greater productivity. Physician Manage 1983;23:212-223.

87. Hemenway D, Killen A, Cashman SB, Parks CL, Bicknell WJ. Physicians' responses to financial incentives. New Engl J Med 1990;322:1059-63.

88. Held and Reinhardt, "Analysis of economic performance," 139-210.

89. Greenfield, "Physician productivity."

90. Hillman AL, Pauly MV, Kerstein JJ. How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations? N Engl J Med 1989;321:86-92.

91. Gorin SNS. Expect the unexpected: consequences of the use of productivity indices. Health Care Strat Manag 1985 Apr;3(4):12-15.

92. Donabedian and others, "Quality, cost, and health."

93. Garvin DA. Managing quality. New York: The Free Press, 1988:84-5.

94. Sytkowski PA, Kannel WB, D'Agostino RB. Changes in risk factors and the decline in mortality from cardiovascular disease. N Engl J Med 1990;322:1635-41.

95. Becker MH, Janz NK. Practicing health promotion: the doctor's dilemma. Ann Int Med 1990;113:419-22.

96. Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med 1989;320:53-6.

97. Reinertsen JL. Transforming a group practice. A progress report on application of continuous quality improvement. Group Prac J 1991 Jan;40(1):9-14.

98. Anderson JG, Benson DS, Schweer HM, Gartner C, Jay SJ. AMBUQUAL: a computer-supported system for the measurement and evaluation of quality in ambulatory care settings. J Amb Care Manag 1989;12:27-37.

99. Reinhardt, "Physician productivity," 126.

100. Simon HA. Rational decision making in business organizations. Am Econ Rev 1979;69:493-513.

101. Smith KR, Miller M, Golladay FL. An analysis of the optimal use of inputs in production of medical services. J Hum Resour 1972; 7:208-224.

102. Linder SB. The harried leisure class. New York: Columbia University Press, 1970.

103. Gaynor, "Competition within the firm."

104. Baumol, Blackman, and Wolff, "Productivity and American leadership," 225-50.

105. Chew, W.B. No-nonsense guide to measuring productivity. Harvard Business Review 1988;66:110-119.

106. Seo KK. Managerial economics. 6th ed. Homewood (IL): Dow Jones-Irwin, 1984:303-30.

107. Reinhardt, "Physician productivity," 63-9.

108. Odiorne GS. The human side of management. Lexington (MA): Lexington Books, 1987:165-82.

109. Greenfield, "Physician productivity."

110. Lazarus SS. Physician services: productivity administrator's handbook. Denver: Center for Research in Ambulatory Health Care Administration, 1987.

111. Physician services: productivity user's manual. Denver: Center for Research in Ambulatory Health Care Administration, 1987.

112. Hadley J, Berenson RA. Seeking the just price: constructing relative value scales and fee schedules. Ann Int Med 1987; 106:461-6.

113. Beer and others, "Human resource management," 431-6, 476-97.

114. Relative Values Studies, Inc. Relative values for physicians. 3rd. ed. New York: McGraw-Hill Information Services Company, 1989.

115. American Medical Association. Physicians' current procedural terminology. 5th ed. Chicago: The Association, 1989.

116. Reinhardt, "Physician productivity," 104-7.

117. Lazarus, "PSP administrator's handbook," 6-15.

118. Department of Health and Human Services, Health Care Financing Administration. Medicare program: model fee schedule for physicians' services. Federal Register 1990; 55:36178-245.

119. Hsiao WC, Braun P, Dunn D, Becker ER. Resource-based relative values. JAMA 1988;260:2347-53.

120. Hsiao WC, Yntema DB, Braun P, Dunn D, Spencer C. Measurement and analysis of intraservice work. JAMA 1988;260:2361-70.

121. Dunn D, Hsiao WC, Ketcham TR, Braun P. A method for estimating the preservice and postservice work of physicians' services. JAMA 1988;260:2371-8.

122. Braun P, Yntema DB, Dunn D, DeNicola M, Ketcham T, Verrilli D, Hsiao WC. Cross-specialty linkage of resource-based relative value scales. JAMA 1988;260:2390-96.

123. Kelly NL, Hsiao WC, Braun P, Sobol A, DeNicola M. Extrapolation of measures of work for surveyed services to other services. JAMA 1988;260:2379-84.

124. Medical Group Management Association. Salary and production report. Denver: The Association, 1990.

125. Gonzalez ML, Emmons DW, editors. Socioeconomic characteristics of medical practice 1989. Chicago: American Medical Association, 1989.

126. Moore FD, Priebe C. Board-certified physicians in the United States. N Engl J Med 1991;324:536-43.

127. Reinhardt, "Physician productivity," 121-36.

128. Seo, "Managerial economics," 353-6.

129. Reinhardt, "Physician productivity," 127.

130. Seo, "Managerial economics," 323.

131. Chiang AC. Fundamental methods of mathematical economics. New York: McGraw-Hill Book Company, 1984:359-63.

132. Seo, "Managerial economics," 316-20.

133. Neter J, Wasserman W, Kutner MH. Applied linear regression models. Homewood (IL): Irwin, 1989:27.

134. Seo, "Managerial economics," 240-6.

135. Number cruncher statistical system [computer program]. Version 5.03. Kaysville (UT): NCSS, 1990.

136. Hintze JL. Number cruncher statistical system installation and reference manual. Kaysville (UT): NCSS, 1990:63.

137. Hintze JL. Software catalog. Kaysville (UT): NCSS, 1990:76-8.

138. Neter and others, "Applied linear regression models," 113-31.

139. Hintze, "Number cruncher reference manual," 174-5.

140. Neter and others, "Applied Linear Regression Models," 435.

141. McHenry C. Multivariate subset selection. J Royal Stat Soc 1978;27:291-6.

142. Goodall C. M-estimators of location: an outline of the theory. In: Hoaglin DC, Mosteller F, Tukey JW, eds. Understanding robust and exploratory data analysis. New York: John Wiley and Sons, 1983:339-400.

143. Neter and others, "Applied linear regression models," 113-31.

144. Neter and others, "Applied linear regression models," 407-10.

145. Neter and others, "Applied linear regression models," 235-6.

146. Neter and others, "Applied linear regression models," 120-1.

147. Neter and others, "Applied linear regression models," 403-4.

148. Cook TD, Campbell DT. Quasi-experimentation. Design and analysis issues for field settings. Boston: Houghton Mifflin Company, 1979:37-94.

149. Neter and others, "Applied linear regression models," 74.

150. Neter and others, "Applied linear regression models," 640.

151. Neter and others, "Applied linear regression models," 450-2, 467-8.

152. Neter and others, "Applied linear regression models," 466.

153. Harris RJ. A primer of multivariate statistics. New York: Academic Press, 1985:6.

154. Hintze JL. Number cruncher statistical system graphics reference manual Version 5.1. Kaysville, Utah, NCSS, 1990:128-34.

155. Simon, "Effort, opportunity, and wealth."

156. Odiorne, "The human side of management," 55-7.

157. Neter, "Applied linear regression models," 231-6.

158. Reinhardt, "Physician productivity," 66.

159. Neter and other, "Applied linear regression models," 470.

160. McDonald, "Protocol-based computer reminders."

161. Pope GC. Physician inputs, outputs, and productivity, 1976-1986. Inquiry 1990;27:151-160.

162. Baumol, "Productivity and American leadership," 124.

163. Baumol, "Price controls," 17.

164. Blendon, RJ. The public's view of the future of health care. JAMA 1988;259:3587-93.