Curt D. Mueller*
National Center for Health Services Research

Waiting for Physicians' Services: Model and Evidence

When services are distributed by appointment, consumers often incur costs of waiting beyond scheduled appointment times. A model of the physician's office is developed that recognizes that physicians jointly produce office services and accessibility. Testable implications of the model are generated, and estimates of reduced-form waiting time and money-price equations are presented. Results suggest that waiting times are responsive to demand and supply-side factors. Estimates are consistent with the theory of the allocation of time and suggest that lower waits will accompany growth in the supply of physicians.

Introduction and Background

Full prices of physicians' services include time as well as money price components. Time is spent in the acquisition of market information on alternative sources of care, prices, and the nature of services available at alternative sites. Physician visits often require substantial travel time expenditures. One source of time costs that has received considerable attention is the time patients spend waiting in physicians' offices. Findings from the National Medical Care Expenditure Survey indicate that in 1977 the average office wait for an appointment visit was approximately 18 minutes (Kasper and Berk 1981). If average hourly earnings for workers in private industry are used to impute a value for the opportunity cost of this time, the price of waiting was over 21% of the price of a routine office visit *before* deducting third-party reimbursements.[1] The waiting time cost component of full price can be expected to increase with future increases in

* I completed this work while at Syracuse University. I thank an anonymous referee for helpful comments. The views expressed in this paper are mine, and no official endorsement by either the National Center for Health Services Research or Department of Health and Human Services is intended or should be inferred.

1. In 1976, average hourly earnings for workers in private industry were $4.87 (U.S. Department of Commerce 1978); the average price of a routine office visit was $10.65 for general practitioners (Gaffney 1979).

(*Journal of Business*, 1985, vol. 58, no. 2)

0021-9398/85/5802-0003$01.50

the value of time and as insurance reduces direct out-of-pocket expenditures for care.

Models of demand for ambulatory physicians' services based on Gary Becker's (1965) seminal analysis of time allocation predict that time costs play a role analogous to money prices (see Grossman 1972; Acton 1973*a*; Miners 1981). These models predict that the degree of consumer responsiveness to changes in money prices is inversely related to the time price fraction of the full price of care and that increases in time costs will encourage consumers to substitute for time price-intensive services. Empirical evidence indicates that demand is inversely related to waiting time, that individuals who place relatively high values on their time are not as sensitive to money prices as those facing smaller time costs, and that patients substitute less time-intensive office-based services for care in longer-wait clinic settings as values of time increase (Phelps and Newhouse 1972; Acton 1973*a*, 1973*b*).

What distinguishes waiting from other sources of time costs is that waiting time is subject to short-run production decisions of suppliers. In the economics literature, waiting time has been viewed as a dimension of product differentness. Nichols, Smolensky, and Tideman (1971) characterize a competitive industry with a continuum of money and time price pairings. Money price is higher where waits are shorter (e.g., in specialty shops relative to discount stores), in part because it is costly to save time for full price–conscious consumers. Similar models are developed by DeVany, House, and Saving (1983) and House (1981). In these models (applied to dental care markets), firms are full price takers. A queuing model—incorporated into the firm's calculus in an earlier work (DeVany 1976)—generates waiting times that are dependent on firm capacity. In equilibrium, different money-nonmoney price mixes are observed, each valued at the market-determined full price.

Finally, waiting time is used as a short-run control variable of the price-setting firm in two analyses that have appeared in this *Journal*. In Olsen's (1976) model, certain inputs are used strictly in making goods available to buyers. These inputs—most appropriate in models depicting the production and sale of *goods* as opposed to *services* because the sales and production processes are separable—determine firm selling capacity and the time price incurred by consumers. Sloan and Lorant (1977) develop a model based on the notion that price-setting physicians determine waiting times by "balancing the efficiency of their operations against patient demand considerations" (p. 487). Waiting time is an argument in the demand and cost functions facing price setters. It is a potential source of efficiency because patient waits reduce input idleness and because waiting time can be used to obtain medical histories.

In this paper, an alternative model of waiting time determination is developed. This model—unlike existing models—explicitly recognizes the relationship between consumer wait and input idleness implicit in provision of services under scheduling schemes. Data permit an empirical examination of the effects of demand and physician supply-side factors on waiting time and money price. The paper concludes with a discussion of some implications.

A Model of the Physician Firm

Waiting and Production

Evidence from demand studies suggests that consumers value short waits. Indeed, studies appearing in medical journals indicate that physicians are cognizant of the opportunity costs of waiting (e.g., Villegas 1967; Landry 1976). Use of an appointment scheme is a means of eliminating excessive waits. Preappointment "outer office" queue time, which patients can use productively, substitutes for "inner office" waiting. Scheduling also benefits physicians. With appointment schemes, bottlenecks that might arise if patient arrivals were unscheduled can be avoided, and idleness of office personnel can be reduced as appointments extend patient contacts over periods of relatively slack demand.

Scheduling accommodates "preemptive priorities" (Schwartz 1975). Patients with "easily" treatable needs can be selectively scheduled around those requiring a full complement of professional personnel who might not be available at all times during office hours. Scheduling allows physicians to exercise priorities based on their assessment of needs. Certain services can be postponed while priority appointments are granted to patients with more urgent needs. Scheduling can be employed to accommodate patients with previously acquired property rights in care continuity. For example, "established" patients, referrals, and follow-up visits can be more readily scheduled than new patients without necessarily subjecting the latter to longer, unproductive office waits.

In appointment offices, inner office waiting time reflects a lack of synchronization between arrivals and the rate of production. If each patient's arrival coincides with completion of the previous patient's consultation, waiting need not arise. The synchronization of arrival and treatment times, however, is difficult to achieve because of "uncertainties" on the demand and supply sides. Patients may fail to arrive as scheduled, and treatment times may vary unexpectedly. If a patient's consultation time happens to fall short of the time allotted and the next patient has not arrived, the physician's time may be wasted. If appointments are scheduled to minimize input idleness, excessive waiting time costs may be incurred by patients.

The trade-off between office visits and patient waiting can be formalized by the relation

$$x = f(Q, A). \tag{1}$$

Inputs required to produce office visits Q are represented by x. "Access" A is an inverse (monotonic) function of expected per patient waiting time t which assumes positive values over the range in which t varies,[2]

$$A = A(t), \tag{2}$$

where $A > 0$ and $dA/dt < 0$. Given the physician's choice of office hours and other inputs, the physician faces a menu of efficient office visits–access combinations $f(Q, A)$ as illustrated in figure 1. Office visits can be increased with reductions in access (increase in waiting), or access can be improved with reductions in the expected level of office visit production.

The source of this trade-off is the substitutability of patient wait for input idleness. More intensive utilization of employed inputs in the production of services reduces idleness and increases patient waits. In figure 1, it is assumed that the substitution of office visits for access occurs at a decreasing rate. This assumption is supported by Monte Carlo studies of scheduling design. In Bailey's (1952) seminal analysis, for example, lengthening the appointment interval increases physician idleness at an increasing rate. Mean patient waiting time falls at a decreasing rate. Thus, successive unit reductions in patient waiting are associated with idle periods of increasing length. Increasingly greater reductions in output result as access improves.[3]

The cost of production is given by wx, the product of inputs employed by the physician—including the physician's time—and input prices w. In relation (1), input usage reflects the levels of office visits

2. In practice, access varies from patient to patient because of the stochastic nature of services provision. For simplicity, however, it is assumed that waiting times at each practice fluctuate randomly around a level t which reflects appointment scheme design.

3. Since Bailey's (1952) work, a number of simulation studies have been undertaken to examine the effects of patient and physician unpunctuality, variations in treatment times, and other practice "disruptions" on patient waiting and provider idleness (e.g., Fetter and Thompson 1966). An anonymous referee has observed that the quantity-access frontier of the model can also be used to characterize choices confronting hospitals. A hospital might receive "profit" or utility for a high occupancy rate (using bed-days as a measure of quantity of product) *and* improvements in accessibility to patients (measured, e.g., in terms of the number of days patients are delayed prior to admission for certain well-defined nonemergency procedures). The hospital faces a trade-off between bed-days and new patient access: higher occupancy rates result in admission delays, given capacity. An increase in hospital capacity—e.g., the number of beds—permits the hospital to increase its occupancy rate *and* decrease the number of patients turned away per production period.

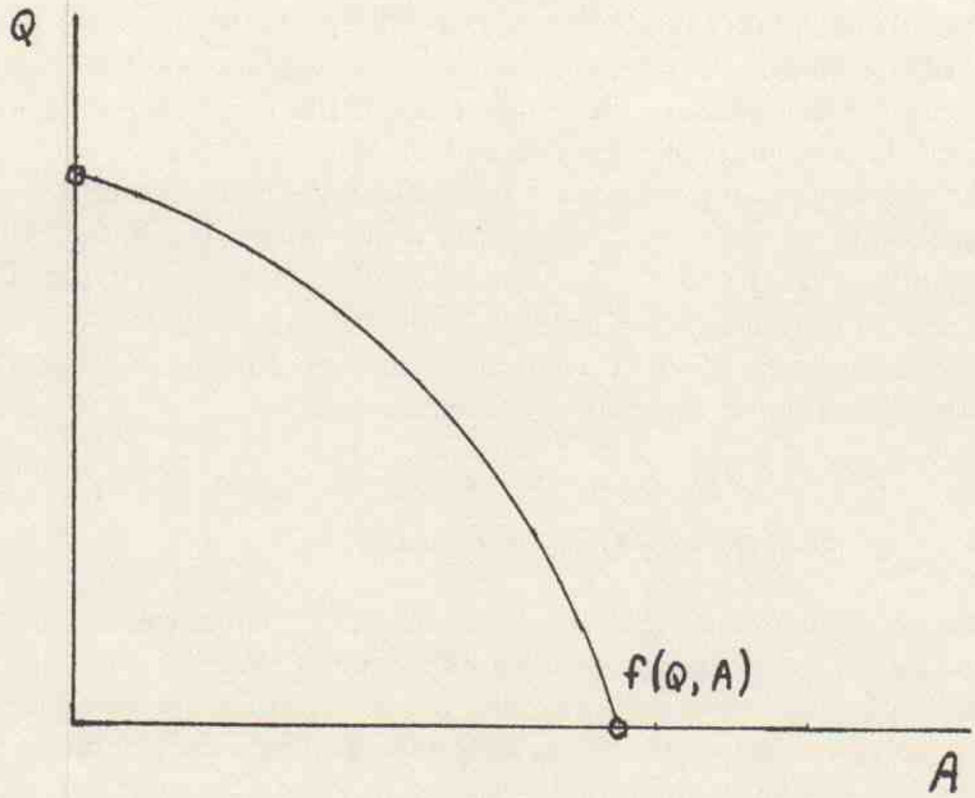

FIG. 1

and access. Thus, the cost of production may be formally represented by

$$C(Q, A; w) = wx = wf(Q, A). \qquad (3)$$

A ceteris paribus increase in the level of office visits requires the use of additional inputs, and cost increases ($\partial C/\partial A > 0$). In a similar fashion, ceteris paribus improvements in access (reductions in waiting time) entail increased input idleness. This requires the hiring of additional inputs if the level of office visit production is to be maintained ($\partial C/\partial A > 0$).

Demand and Waiting Time Determination

A general model. The specific office visit–access combination chosen by the physician depends on the demand for primary care services and the physician's behavior. In economic analyses of physician services markets, it is generally assumed that physicians are price setters (e.g., Sloan 1976; Sloan and Lorant 1977; Feldman 1979; Paringer 1980). Market prices are not competitively determined because potential patients frequently lack information on prices and because the nature or attributes of services vary from physician to physician. Market demand faced by the individual physician can thus be represented by the relation

$$Q^d = Q(p, t; M, v). \qquad (4)$$

Demand is negatively related to money price p and the cost of waiting. The cost of waiting vt is the product of per visit office waiting time and the market's (exogenous) per unit value of time v. Other exogenous demand determinants are represented by M.

The physician's output-access production decision can now be examined under a variety of behavioral assumptions. Following Sloan and Lorant (1977) and DeVany et al. (1983), assume that providers behave as if they maximize profits. As demonstrated in the Appendix, the optimum levels of office visit and access production are those that satisfy the following "necessary" conditions:

$$R_Q(Q; A, M, v) - C_Q(Q; A, w) = 0 \tag{5a}$$

$$R_A(A; Q, M, v) - C_A(A; Q, w) = 0. \tag{5b}$$

Condition (5a) implies that the level of office visit production is increased until its contribution to the physician's revenue (the marginal revenue of office visits, R_Q) equals its contribution to practice cost (marginal cost, C_Q), *given* the selected level of access. Condition (5b) insures equality of the marginal revenue (R_A) and marginal cost (C_A) of access given the level of office visits.

The primary purpose of this formalization is to examine the determination of office waiting time. The model as developed to this point is not particularly useful. The problem—recognized in the Sloan and Lorant (1977) analysis and in general by economists for some time (e.g., Archibald 1964)—is that joint maximization models of this type often fail to generate testable implications. Consider, for example, the effects of an exogenous increase in demand for short-wait office visits. Figure 2 contains a representation of condition (5b). The physician is initially maximizing profits at access level A_0 where $R_A = C_A$. An increase in demand increases the marginal revenue of access. R_A shifts outward along C_A, and the level of access begins to rise (waiting time falls). But an increase in demand also increases the profitability of office visit production (R_Q in condition [5a]). An increase in Q may generate additional shifts of the marginal revenue and cost of access schedules, depending on the specific properties of the revenue and cost specifications. In figure 2, an increase in the optimum level of office visits from Q_0 to Q_1 increases the marginal cost of access to C_A^1. This positive "cost effect" is dominated by a positive "revenue effect" (R_A^1) and access rises to A_1. It is conceivable that increases in Q will exert a relatively greater impact on the marginal cost of access than on R_A. If the cost effect dominates the revenue effect, waiting time rises. In short, effects of exogenous parameters on the levels of access, office visits, and money price cannot be predicted without additional assumptions.

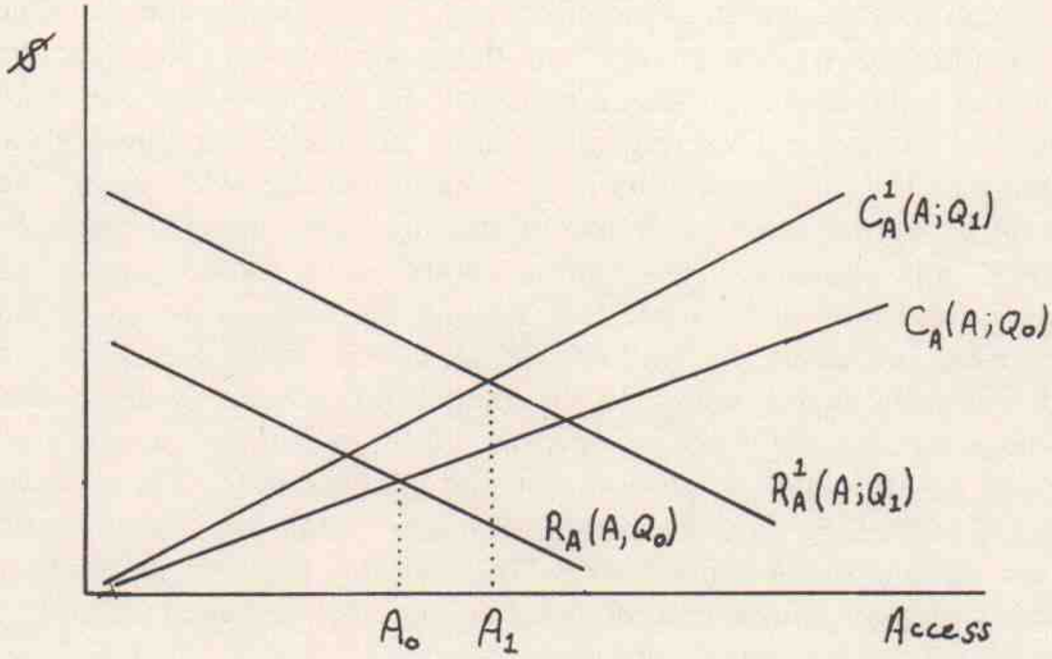

FIG. 2

Structural assumptions and implications. The set of equations (5) determines the optimal office visit–waiting time combination chosen by the physician. Thus, if specific revenue and cost expressions are specified, it may be possible to solve equations (5) explicitly for Q and A. Resulting reduced-form equations can then be used to examine implications of the structural model.

This procedure is detailed in the Appendix. An expression for revenue is derived using a market demand specification—of a type frequently used in economic analysis—which is extended to explicitly recognize time and money prices. Production cost is specified using a simple production possibilities frontier which embodies the desired properties illustrated in figure 1. Some implications of the model (based on the reduced-form equations in the Appendix) are summarized in table 1.

TABLE 1 Comparative Static Responses

	Exogenous Parameters		
Physician Variables	W	M	ν*
t	+	−	−
Q	−	+	+
x	−	+	+
p	+†	+	+

NOTE.—Results above hold under conditions stipulated in the text below and in Appendix.
* Results in this column hold when the earnings elasticity of demand $\epsilon_\nu > 0$ (see below and App.).
† This result holds if the time price elasticity of demand $-1 < \epsilon_{t\nu} < 0$ (see App.).

Recall that the waiting time effect of a change in demand could not be unambiguously determined from the general model. The structural demand and cost specifications imply that the revenue effect dominates the cost effect, and waiting time varies inversely with the level of demand. At first glance, this may seem counterintuitive. A practice might accommodate an increase in demand by reducing the time between appointments. The waiting room queue would increase in length, and one might argue that waiting time should increase. This argument presumes that patient waiting is substituted for idleness of currently employed inputs. In this model, an increase in demand increases inputs used in the production of both jointly produced outputs Q and A. Marginal increases in demand increase office visits and access. The waiting time component of full price falls, and money price rises. Among those consumers willing and able to incur higher direct money charges, those with high valuations of time benefit relative to those with lower opportunity time costs.

A frequently used, operational measure of one's value of time is earnings. In the structural demand specification (relation [vi] in the Appendix), earnings per unit time v inversely affect demand through the waiting time price vt (the price effect). But the earnings parameter also enters the demand specification directly. This is because earnings constitute the primary source of consumer income. Holding the time price constant, it is assumed that earnings directly affect demand (the shift effect). If the positive shift effect dominates the negative time price effect, the demand for office visits is directly related to market earnings.[4] Market earnings exert effects on physician behavior similar to those resulting from a general increase in demand. As summarized in table 1, physicians in higher earnings communities will produce more office contacts and lower-wait visits than physicians in lower-income communities, ceteris paribus.

Finally, consider the effects of a change in the cost of production. An increase in factor prices w increases the marginal cost of office visits. The marginal cost of access also increases because the cost of input idleness rises. Fewer inputs are demanded by the firm. Office visit production falls and waiting time rises. Empirical evidence suggests that the demand for office visits is time price inelastic; if so, money prices vary directly with factor prices.

4. The assumption that the positive shift effect dominates the negative time price effect is analogous to the claim that the income effect of a wage change dominates the negative time price substitution effect. There is some evidence that this is the case for physicians' services (e.g., Acton 1973*a*). It should be noted that the waiting time effect of unearned income (defined at the market level) might be investigated by treating wealth as an M-type variable in the demand specification. It is ignored in the text because suitable proxy data could not be obtained for all market areas as defined in the sample design. (See below.)

Empirical Tests of the Model

The regression model to be estimated, based on the reduced-form waiting time equation derived in the Appendix, is formally given by

$$\ln (t_i) = \theta_1 - \sum_{j=1}^{k} (\epsilon_{Mj}/\delta) \ln (M_{ij}) - (\epsilon_v/\delta) \ln (v_i) + (\epsilon_p/\delta) \ln (w_i). \quad (6)$$

The unit of analysis is the physician's practice. Mean waiting time in the *i*th physician's office is represented as t_i. M_{ij} represents k exogenous determinants of demand, v_i is the market valuation of time, and w_i is an index of factor prices. Coefficients to be estimated are functions of the following structural parameters: θ_1, a constant term; $\epsilon_p > 0$, the price elasticity of demand for office visits; ϵ_{Mj}, elasticities of demand with respect to the demand-shift parameters M_j; ϵ_v, the income elasticity of demand; and $\delta = 1 + (\epsilon_p - \epsilon_{tv})$, where $\epsilon_{tv} < 0$ is the time price elasticity of demand.

The effects of the exogenous variables M_j, v, and w in expression (6) correspond to those of table 1 so long as $\delta > 0$. This is a reasonable assumption. One implication of demand models that incorporate time costs is that $\epsilon_p \gtreqless |\epsilon_{tv}|$ as $p \gtreqless tv$; that is, the money price elasticity exceeds (is less than) the time price elasticity if the money price component of full price exceeds (is less than) the time price component (Acton 1973*a*). Because evidence from consumer surveys and demand studies suggests that money prices exceed waiting time prices, $\epsilon_p > \epsilon_{tv}$ and $\delta > 0$.

Data and Hypotheses

Data have been obtained from several sources. Mean waiting times in generalist and internist offices and other practice characteristics are from a public use file data tape compiled from the Physician Capacity Utilization Telephone Surveys. These surveys were administered in 1973 and 1974 by Mathematica, Inc., as part of HEW's Health Insurance Study. Over 8,000 generalists, internists, and pediatricians in 156 SMSAs and 63 nonmetropolitan multiple county sampling areas were surveyed.[5] The public use file's design to protect confidentiality of the survey respondents prohibits identification of precise practice locations within the sampling areas. Thus, each area constitutes a "market." Data on demand determinants and the price of time and factor price measures have been obtained from other sources by aggregat-

5. Discussions of sample selection, questionnaire design and administration, and a summary of findings with validity assessments are contained in Berry et al. (1976). The public use file is documented in Brown and Wooldridge (1976). Sampled pediatricians have been excluded from this analysis because their focus is on much more narrowly defined population groups relative to generalists and internists.

ing—data sources permitting—over those counties which compromise each sampling area.

Because suitable earnings data are unavailable at relevant levels of aggregation, the price of time is proxied by real 1973 disposable per capita income, IPC, in each physician's market area.[6] IPC contains unearned as well as earned income components. If demand for visits is directly related to unearned income, the unearned income component functions as an M-type variable in the model. Unearned income reinforces the predicted negative waiting time effect of a change in earnings, and a negative relationship between t and *IPC* is expected.

Two regressors are included in the reduce-form equation to assess the impact of health insurance coverage. REGM is the percentage of the population below age 65 with regular medical expense coverage in the physician's state.[7] Regular medical insurance—as defined by the Health Insurance Institute—provides benefits toward nonsurgical physicians' services rendered in the office, home, and hospital, and may provide benefits for laboratory tests performed outside of the hospital. MCARE, the percentage of the civilian population over age 65 in the physician's market area, is a proxy measure of the extent of Medicare coverage.[8] Both private and public insurance may directly affect demand because of moral hazard and because effective regular benefits coverage reduces the net money price of care. The model predicts that physicians will increase productive capacity as coverage (with cost-sharing) enrollments rises. Increased capacity facilitates improvements in access. On the other hand, effective coverage of office-based care is limited. As primary care services are not often well covered, REGM may fail to exert noticeable effects on waiting time.[9]

6. Income per capita data (aggregated from data at the county level) are from the American Medical Association (1974, 1975). Per capita incomes for markets of physicians surveyed in 1974 were deflated to 1973 dollars using the CPI. U.S. Bureau of Labor Statistics budget indices of 1973 living costs "at an intermediate level of living" for four-person families were used to adjust market IPC values for cross-sectional differences in the purchasing power of money income (U.S. Department of Labor 1975).

7. State-level data on numbers of people with regular medical coverage are from the Health Insurance Institute (1975, 1976). Estimates of the under-65 populations (by state) are from Sales Management, Inc. (1974, 1975). Weighted averages of REGM are computed for those markets which extend across state borders, where weights are fractions of the market area populations in each state (American Medical Association 1975).

8. MCARE values are computed using county data from U.S. Bureau of Census (1972).

9. Sloan and Lorant (1977) argue that major medical insurance coverage more completely subsidizes the purchase of "amenities" such as lower waits than "basic" insurance. Under major medical coverage, the patient pays a fraction of the total charge (defined by the plan's coinsurance provisions) once the deductible is satisfied. Benefits under basic-type "regular" plans may be limited by relatively stringent, service-specific fee schedules. Thus, if a patient seeks lower-wait, higher-money-priced care, the additional money charge incurred directly is apt to be greater under regular than under effective major medical coverage (pp. 495–96). On the other hand, it can be argued that for patients who have satisfied their deductibles because treatment regimens require

American Medical Association data on the distribution of physicians by county and specialty groupings are employed to construct two measures of physician availability. MDS measures physicians of the respondent's specialty grouping. If the respondent is a general or family practitioner, MDS is the number of generalists in the physician's market area per 10,000 population. If the respondent is an internist, MDS is the number of "medical specialists" per 10,000 population.[10] MDSA is a similarly constructed density measure designed to assess the impact of the availability of specialty substitutes. For generalist (internist) respondents, MDSA is the number of medical specialists (generalists) per 10,000 population. MDS and MDSA control for differences in the levels and elasticities of demand faced by physicians, as well as other nonmoney price barriers of accessibility. The structural model and empirical evidence suggest that t will be inversely related to MDS and MDSA.[11] It is expected that the effect of physician density in one's own specialty (MDS) will dominate.

The public use data file does not identify personal characteristics of providers that could serve as determinants of the value of the physician's time. County-level data have been used to compute estimates of the average quarterly wage (WAGE) paid to physician office employees in respondent market areas.[12]

Finally, data from the public use file are used to standardize for differences in practice technologies. The categorical variable *IM* equals zero if the physician is a general practitioner; *IM* equals one if the physician is an internist. The dummy variable FCFS is included to control for differences in patient priority schemes employed by physicians. If the physician treats patients on a first-come-first-served (fcfs) basis as well as by appointment, FCFS is set equal to one. Otherwise, FCFS assumes the value zero. Wait is hypothesized to be a positive

---

series of return visits, waiting time may be relatively less important on the margin. Sloan and Lorant find that the fraction of a physician's patients with major medical coverage (not necessarily effective) exerts a statistically significant, negative effect on office wait. Effects of major medical coverage cannot be examined in this study because data including major medical coverage under Blue Cross/Blue Shield plans are not available at the market level for the 1973–74 period.

10. MDS and MDSA are computed from physician and population data contained in American Medical Association (1974, 1975). "Medical specialists" include internists as well as physicians from a number of other major nonsurgical specialties. See either of the cited references for a complete listing.

11. Structural assumptions imposed in this analysis imply that output and access levels are lower in "monopoly" than in "competitive" (socially optimal) market structures (see Levhari and Peles 1973). Negative relationships between waiting times and providers per capita are observed by Sloan and Lorant (1977), House (1981), and DeVany et al. (1983).

12. WAGE values are estimated using county data on the numbers of office employees and corresponding payrolls in 1973 and 1974, from U.S. Bureau of the Census (1974, 1975). Wages for 1974 were converted to 1973 dollars; adjustments for intermarket differences in purchasing power are described in n. 8 above.

function of FCFS: fcfs arrivals may disrupt production continuity. It may be argued that in offices where both appointment and fcfs patients are seen, demand pressures of appointment patients lead to adjustments of the appointments schedule, taking into account the incidence of fcfs arrivals. But there is no reason to expect waiting time in such offices to be less than in appointment-only offices.

The number of examining rooms per full-time equivalent physician associated with the practice, including the physician's private office, ROOMS, controls for differences in practice size which may affect scale economies and differences in relative efficiencies of office visit–access production. Data on practice size (in terms of numbers of physicians) are not available from the public use file.

Estimation Results

Table 2 contains summary statistics; ordinary least squares regression estimates of the waiting time equation appear in table 3. Only 3% of the variation in the logarithm of waiting time is explained by the exogenous and standardizing variables. Low R^2 values—common in analyses of survey data where the individual is the unit of analysis—are not atypical in waiting time studies. House (1981), for example, reports R^2 values in the .01–.02 range. One explanation is that stochastic factors beyond direct control by physicians—for example, random variations in arrival and length of treatment times—may account for much of the

TABLE 2 **Variable Definitions and Descriptive Statistics of the Physician Sample (N = 5,637)**

Variable	Definition	Mean	Standard Deviation
t	Mean office wait, minutes	20.72	17.65
IPC	Real 1973 income per capita, 00s	42.37	5.51
REGM	% civilian population under age 65 with regular medical coverage	72.91	12.17
MCARE	% civilian population age 65 and over	9.78	2.71
MDS	Physicians per 10,000 population, respondent's specialty group	2.50	.88
MDSA	Physicians per 10,000 population, substitute specialty group	2.05	.92
WAGE	Real 1973 quarterly wage of office personel, 000s	2.57	.54
IM	= 0 for generalists = 1 for internists	.46	. . .
FCFS	= 0 for appointment office = 1 if also fcfs	.08	. . .
ROOMS	No. of examining rooms, including physician's office	3.62	1.33

Source.—Author's tabulations of Physician Capacity Utilization Telephone Survey and supplementary data.

TABLE 3 Reduced-Form Estimates: Waiting Time and Money Price

	Dependent Variable = Ln (Waiting Time)	Dependent Variable = Ln (Money Price)
Variable[a]	Estimate (S.E.)	Estimate (S.E.)
IPC	−.351**	.190**
	(.115)	(.031)
REGM	−.022	−.188**
	(.072)	(.019)
MCARE	.049	−.035*
	(.052)	(.014)
MDS	−.135**	.066**
	(.038)	(.010)
MDSA	−.071*	.050**
	(.028)	(.007)
WAGE	.135*	.051**
	(.058)	(.016)
IM	−.124**	.229**
	(.027)	(.007)
FCFS	.074	−.014
	(.047)	(.013)
ROOMS	.250**	−.002
	(.038)	(.010)
Constant	3.757**	2.093**
	(.518)	(.139)
	$R^2 = .034$	$R^2 = .250$
	$F(9,5627) = 19.67$	$F(9,5131) = 189.89$
	$N = 5,637$	$N = 5,141$

SOURCE.—Author's estimates, using Physician Capacity Utilization Telephone Survey and supplementary data.

[a] All independent variables except IM and FCFS are entered as natural logarithms.

* Significant at the 5% level, two-tailed *t*-test.

** Significant at the 1% level, two-tailed *t*-test.

total variation. Nevertheless, the magnitude of the *F*-statistic implies that the null hypothesis of no relationship between the dependent and independent variables is rejected at the 1% level.

Waiting time is more responsive to real per capita income than to any of the other demand and supply-side determinants. The IPC elasticity is negative and statistically significant at conventional levels. Negativity is consistent with the model's prediction and the notion that physicians are cognizant of consumer time costs. Sloan and Lorant (1977) report inverse relationships between mean waits and incomes of individual physicians' patients; results here suggest that waits are less in higher-income market areas, ceteris paribus. There is no evidence that the extent of insurance coverage affects waiting time costs. The effect of REGM, while negative as expected, is statistically indiscernible and small in magnitude. MCARE also fails to affect waiting time.

Waiting time is inversely related to physician density. The elasticity with respect to physicians per capita of similar specialties (MDS) exceeds (in absolute value) the elasticity of substitutable physicians per capita. Parameter estimates suggest that waiting time costs will rise with reductions in physician density. However, an extremely large decrease in the number of physicians per capita is necessary to substantively increase office waits. A one minute increase in t (relative to the sample mean) "requires" a ceteris paribus 36% decrease in MDS.

Office waiting time is directly related to WAGE, the price of physician office personnel. The elasticity of +0.14 is statistically significant, in contrast to elasticities computed from the Sloan-Lorant estimates. The latter, based on indices of wages of hospital employees in the physician's county, are smaller by a factor of seven.

The focus of this paper has been on the determination of waiting time as a component of the full price of ambulatory care. As a final "test" of the model, the implied reduced-form money price equation (derived in the Appendix and summarized in table 1) is estimated. The dependent variable is the natural logarithm of the money price of a routine office visit (adjusted for intermarket differences in the cost of living). Estimates—obtained using data for those physicians of the waiting time sample who reported money prices—appear in the third column of table 3.[13]

Routine office visit prices for internists exceed prices for generalists. Together, the equations of table 3 suggest that while characteristics of demand or production confronting specialists account for higher money prices of office-based care, monetary costs may be somewhat offset by shorter office waits. The directional impact of WAGE is also consistent with estimates from other pricing studies (e.g., Sloan 1976) and the model developed above. Higher production costs can be expected to directly affect both money and time cost components of full price.

On the demand side, income directly affects the money price of care. The notion that consumers will attempt to substitute away from time intensive commodities may underlie the direct relationships between money price, access, and market income. The REGM and MCARE coefficients are not consistent with the expectation that money prices directly reflect the extent of the market with insurance coverage. One possible explanation is that the state-level data are poor proxies for the extent of market-area insurance coverage.

The estimates in table 3 reveal a positive association between money

13. Of the 5,637 physicians included in the waiting time analysis, 5,141 provided usable price information. Descriptive statistics and estimates of the waiting time equation for the subsample reporting both waiting time and money prices have not been reported for purposes of brevity but are available from the author. Prices for physicians surveyed in 1974 were adjusted for inflation using the physicians' services component of the Consumer Price Index. Adjustments for intermarket differences in purchasing power were made as indicted in n. 8 above. After adjustments, the mean and standard deviation of price are \$8.93 and \$2.64, respectively.

price and physician density. This finding certainly conflicts with the notion of the standard monopoly model that increases in physician density reduce the profit-maximizing physician's market share, resulting in lower prices. The source of this apparent inconsistency, however, is not known. A direct relationship between price and physician supply might be observed if physicians do not behave as profit maximizers, but instead behave as satisficers. The desire to maintain a target income is one such behavioral objective that is often discussed. But rejection of the standard profit maximization assumption is by no means required. Pauly and Satterthwaite (1981), for example, argue that consumer search is apt to be less efficient in markets with relatively many monopolistically competitive physicians. Less efficient search reduces price sensitivity and permits physicians to receive higher prices. Their price equation estimates yield a negative relationship between market price and physicians per capita only *after* attempting to control for intermarket differences in consumer information.

Summary and Implications

The model developed in this paper recognizes that waiting time, as a component of full price, is a cost incurred by effective demanders. The structural model explicitly recognizes the role of waiting time in production. Services and accessibility are jointly determined. Service providers face a trade-off: high rates of output can be maintained with relatively lengthy consumer waits, or output rates can be reduced as input idleness substitutes for consumer waiting. Specific output-wait-money price combinations are determined in response to factors affection the demand for and supply of services.

Empirical results using data on physicians can be viewed from two perspectives. First, estimation is a means of testing the model. Directional effects of demand- and supply-side factors are generally consistent with the model's predictions. Second, waiting time results supplement existing studies on the effects of factors believed to underlie variations in waiting times independent of any formal behavior model. Results, in conjunction with the money price estimates, are suggestive of the impacts of economic factors on full prices of primary care.

Waiting time is negatively related to real income. In higher-income markets where money prices are greater, waiting times are less ceteris paribus. This is consistent with economic models of time allocation which predict that as the value of time rises, consumers substitute money price for time price–intensive commodities. This has implications on access to primary care services. Less waiting does not appear to facilitate access in lower income markets where relatively greater members of people are apt to face more stringent financial barriers.

A surplus of primary care physicians is expected by 1990. Evidence to date questions whether enhanced supplies will help—through market forces—to contain health care costs. Estimates suggest that although money prices may be higher in relatively physician-dense markets, lower waits may help somewhat in holding down full prices. The extent to which increases in physician density reduce waiting time costs depends on patient opportunity costs of time. Waiting cost reductions are greater for those with higher values of time. Thus, if income proxies time costs, enhanced physician availability facilitates access for those most able to pay. Evidence also indicates that production costs directly affect waiting time as well as money prices. Within a joint production framework, wage increases commanded by office employees increase the cost of input idleness, raising full prices of primary care.

Appendix

The physician is assumed to make production decisions as if he or she is maximizing profits, defined as the difference between revenue and cost. Revenue—the product of money price and the number of office visits produced—can be written as a function of Q and A by solving $A = A(t)$ (relation [2] in the text) for t, substituting the result into the demand expression (4), solving the latter for p, and multiplying the result by Q:

$$R(Q, A; M, v) = p(Q, A; M, v)Q. \tag{i}$$

Using the cost expression (3) in the text, profit is given by

$$\pi = R(Q, A; M, v) - C(Q, A; w).$$

The first-order necessary conditions for an internal optimum combination of Q and A (relations [5]) are the following:

$$\partial\pi/\partial Q = R_Q(\cdot) - C_Q(\cdot) = 0 \tag{iia}$$

$$\partial\pi/\partial A = R_A(\cdot) - C_A(\cdot) = 0 \tag{iib}$$

(subscripts denoting partial derivatives).

Structure is imposed on the model by first assuming the production possibility frontier (relation [1], illustrated in fig. 1) can be represented by

$$x = K_1Q^2 + K_2A^2, \tag{iii}$$

where $K_1, K_2 > 0$, and

$$A = 1/t, \tag{iv}$$

$t > 0$. It follows that production cost (relation [3]) becomes

$$C(Q, A; w) = w(K_1Q^2 + K_2A^2). \tag{v}$$

Let the market demand for office visits assume the conventional multiplicative form

$$Q^d = \alpha_0 p^{-\alpha_1}(tv)^{-\alpha_2}M^{\alpha_3}v^{\alpha_4} \tag{vi}$$

where $\alpha_i > 0$, $i = 0, \ldots, 4$.

Relations (iv)–(vi) can be used to solve equations (ii) simultaneously for Q and t in terms of the exogenous parameters. The natural logarithms of these "reduced-form" equations are given by the following:

$$\ln t = \theta_1 + \epsilon_p/\delta(\ln w) - \epsilon_m/\delta(\ln M) - \epsilon_v/\delta(\ln v), \qquad \text{(vii)}$$

$$\ln Q = \theta_2 - \epsilon_p/\delta(\ln w) + \epsilon_m/\delta(\ln M) + \epsilon_v/\delta(\ln v), \qquad \text{(viii)}$$

where θ_1 and θ_2 are constant terms, $\epsilon_p = \alpha_1 > 0$, $\epsilon_m = \alpha_3 > 0$, $\epsilon_v = \alpha_4 - \alpha_2 \lesseqgtr 0$, $\delta = 1 + \epsilon_p + \epsilon_{tv} > 0$ (as discussed in the text), and $\epsilon_{tv} = -\alpha_2 < 0$. The money price reduced form—obtained by substituting the antilogs of (vii) and (viii) into (vi), solving for p, and taking its natural logarithm—is given by

$$\ln p = \theta_3 + (1 + \epsilon_{tv})/\delta\ (\ln w) + \epsilon_m/\delta\ (\ln M) + \epsilon_v/\delta\ (\ln v), \qquad \text{(ix)}$$

θ_3 a constant term.

Comparative static effects appearing in table 1 are obtained from equations (vii)–(ix) and (iii).

References

Acton, Jan Paul. 1973*a*. *The Demand for Health Care among the Urban Poor with Special Emphasis on the Role of Time*. Santa Monica, Calif.: Rand Corp.

Acton, Jan Paul. 1973*b*. *Demand for Health Care When Time Prices Vary More than Money Prices*. Santa Monica, Calif.: Rand Corp.

American Medical Association, Center for Health Services Research and Development. 1974. *Distribution of Physicians in the U.S., 1973*. Chicago: AMA.

American Medical Association, Center for Health Services Research and Development. 1975. *Physician Distribution and Medical Licensure in the U.S., 1974*. Pt. 1. Chicago: AMA.

Archibald, G. C. 1964. Profit-maximizing and non-price competition. *Economica* 31 (February): 13–22.

Bailey, Norman T. J. 1952. A study of queues and appointment systems in hospital outpatient departments, with special reference to waiting times. *Journal of the Royal Statistical Society*, Ser. B, 14:185–99.

Becker, Gary S. 1965. A theory of the allocation of time. *Economic Journal* 75 (September): 493–517.

Berry, Charles, et al. 1976. *Report on the Physician Capacity Utilization Telephone Surveys*. Vol. 1. Princeton, N.J.: Mathematica Policy Research.

Brown, Jay, and Wooldridge, Judith. 1976. *Documentation of Physician Capacity Utilization Telephone Surveys: Public Use Files 1 and 2*. Princeton, N.J.: Mathematica Policy Research.

DeVany, Arthur S. 1976. Uncertainty, waiting time, and capacity utilization: A stochastic theory of product quality. *Journal of Political Economy* 84 (June): 523–541.

DeVany, Arthur S.; House, Donald R.; and Saving, Thomas R. 1983. The role of patient time in the pricing of dental services: The fee-provider density relation explained. *Southern Economic Journal* 49 (January): 669–680.

Feldman, Roger. Price and quality differences in the physicians' services market. *Southern Economic Journal* 45 (January): 885–91.

Fetter, R. B., and Thompson, J. D. 1966. Patients' waiting time and doctors' idle time in the outpatient setting. *Health Services Research* 1 (Summer): 66–90.

Gaffney, John C., ed. 1979. *Profile of Medical Practice, 1978*. Chicago: American Medical Association.

Grossman, Michael. 1972. *The Demand for Health: A Theoretical and Empirical Investigation*. New York: National Bureau of Economic Research.

Health Insurance Institute. 1975. *Source Book of Health Insurance Data, 1974–75*. New York: Health Insurance Institute.

Health Insurance Institute. 1976. *Source Book of Health Insurance Data, 1975–76*. New York: Health Insurance Institute.

House, Donald R. 1981. A full-price approach to the dental market: Implications for price determination. *Journal of Health Politics, Policy and Law* 5 (Winter): 593–609.
Kasper, Judith A., and Berk, Marc L. 1981. *Waiting Times in Different Medical Settings: Appointment Waits and Office Waits,* National Health Care Expenditures Study Data Preview, no. 6. Washington, D.C.: Department of Health and Human Services.
Landry, M. 1976. There are good and bad ways to set up an appointment system. *Canadian Medical Association Journal* 115 (July): 160–61.
Levhari, David, and Peles, Yoram, 1973. Market structure, quality, and durability. *Bell Journal of Economics and Manageable Science* 4 (Spring): 235–48.
Miners, Laurence A. 1981. The family's demand for health: Evidence from a rural community. In *Advances in Health Economics and Health Services Research,* Vol. 2, ed. Richard M. Scheffler. Greenwich, Conn.: JAI.
Nichols D.; Smolensky, E.; and Tideman, T. N. 1971. Discrimination by waiting time in merit goods. *American Economic Review* 61 (June): 312–23.
Olsen, E. Odgers. 1976. The role of time in the neoclassical theory of the firm. *Journal of Business* 4 (October): 541–54.
Paringer, Lynn. 1980. Medicare assignment rates of physicians: Their response to changes in reimbursement policy. *Health Care Financing Review* 1 (Winter): 75–89.
Pauly, Mark V., and Satterthwaite, Mark A. 1981. The pricing of primary care physicians' services: A test of the role of consumer information. *Bell Journal of Economics* 12 (Autumn): 488–506.
Phelps, Charles E., and Newhouse, Joseph P. 1972. Effects of coinsurance: A multivariate analysis. *Social Security Bulletin* 35 (June): 20–28.
Sales Management, Inc. 1974. *1974 Survey of Buying Power*. July 8.
Sales Management, Inc. 1975. *1975 Survey of Buying Power*. July 21.
Schwartz, Barry. 1975. *Queuing and Waiting*. Chicago: University of Chicago Press.
Sloan, Frank A. 1976. Physician fee inflation: Evidence from the 1960s. In *The Role of Health Insurance in the Health Services Sector,* ed. Richard N. Rosett. New York: National Bureau of Economic Research.
Sloan, Frank A., and Lorant, John H. 1977. The role of waiting time: Evidence from physicians' practices. *Journal of Business* 50 (October): 485–507.
U.S. Bureau of the Census. 1972. *City and County Data Book.*
U.S. Bureau of the Census. 1974. *County Business Patterns, 1973* (Various issues by state).
U.S. Bureau of the Census. 1977. *County Business Patterns, 1974* (Various issues by state).
U.S. Department of Commerce. 1978. *Statistical Abstract of the U.S. 1978.*
U.S. Department of Labor, Bureau of Labor Statistics. 1975. *Handbook of Labor Statistics, 1975.*
Villegas, Eduardo L. 1967. Outpatient appointment system saves time for patients and doctors. *Hospitals* 41 (April): 52–57.