Other Articles

The Impact of Financial Incentives on Physician Productivity in Medical Groups

*Douglas A. Conrad, Anne Sales, Su-Ying Liang,
Anoshua Chaudhuri, Charles Maynard, Lisa Pieper,
Laurel Weinstein, David Gans, and Neill Piland*

Objective. To estimate the effect of financial incentives in medical groups—both at the level of individual physician and collectively—on individual physician productivity.

Data Sources/Study Setting. Secondary data from 1997 on individual physician and group characteristics from two surveys: Medical Group Management Association (MGMA) Physician Compensation and Production Survey and the Cost Survey; Area Resource File data on market characteristics, and various sources of state regulatory data.

Study Design. Cross-sectional estimation of individual physician production function models, using ordinary least squares and two-stage least squares regression.

Data Collection. Data from respondents completing all items required for the two stages of production function estimation on both MGMA surveys (with RBRVS units as production measure: 102 groups, 2,237 physicians; and with charges as the production measure: 383 groups, 6,129 physicians). The 102 groups with complete data represent 1.8 percent of the 5,725 MGMA member groups.

Principal Findings. Individual production-based physician compensation leads to increased productivity, as expected (elasticity = .07, $p < .05$). The productivity effects of compensation methods based on equal shares of group net income and incentive bonuses are significantly positive ($p < .05$) and smaller in magnitude. The group-level financial incentive does not appear to be significantly related to physician productivity.

Conclusions. Individual physician incentives based on own production do increase physician productivity.

Key Words. Compensation, managed care, financial incentives, physician productivity

This article addresses an increasingly important issue in medical practice and health services research—the effect of financial incentives on physician behavior. Specifically, this study estimates the impact of financial incentives

on physician productivity in medical groups. The medical group represents a natural organized setting for observing variation in financial incentives and the impact of those incentives on physician behavior. For example, salary compensation and income-sharing methods are only observable within groups, not solo practices. Moreover, the majority of active patient care physicians practice in medical groups of three or more members (Gaynor 1994), and that proportion has been growing over time. The American Medical Association (Havlicek 1996) reports that the number of physicians in groups rose by 18.5 percent from 1988 to 1991 and 14.3 percent from 1991 to 1995. Simultaneously, group practice size has been increasing, and the share of physicians in solo practice has declined from more than one-third in 1991 to just more than one-quarter in 1995 (Gaynor and Haas-Wilson 1999).

In the medical group, two “tiers” of financial incentives bear on physician productivity: (1) the method of payment by health plans and organized purchasers (e.g., employer groups contracting directly with provider organizations) to the medical group; (2) the method used by the medical group to compensate individual physicians (Hillman, Welch, and Pauly 1992). Both the plan-to-group and group-to-individual physician incentive tiers are expected to influence individual physician behavior: the first via a “group” effect, the latter through the direct effect of individual physician compensation.

The current study builds on a strong tradition of empirical research regarding the effects of financial incentives on physician productivity. The major previous econometric studies by Reinhardt, Pauly, and Held (1979), Gaynor and Pauly (1990), and Gaynor and Gertler (1995) all drew on a

This research was supported by a grant from the Health Care Financing and Organization (HCFO) Initiative of the Robert Wood Johnson Foundation.

Address correspondence to Douglas A. Conrad, Ph.D., Professor, University of Washington, Department of Health Services, 1959 NE Pacific Street, Box 357660, Seattle, WA 98195-7660. Anne Sales, Ph.D., is an Assistant Professor, University of Washington, Department of Health Services, and is also with the Veterans Administration Puget Sound Health System, Health Services Research and Development Program. Su-Ying Liang, Ph.D., is a Research Assistant Professor, School of Pharmacy, University of California at San Francisco. Anoshua Chaudhuri is a Ph.D. Candidate, Department of Economics, University of Washington. Charles Maynard, Ph.D., is a Research Associate Professor, University of Washington, Department of Health Services, and is also with the Veterans Administration Puget Sound Health System, Health Services Research and Development Program. Lisa Pieper, M.B.A., and Laurel Weinstein, B.A., are Research Associates with the Center for Research, Medical Group Management Association (Englewood, CO); Neill Piland, Dr.P.H., is Director of the Center. David Gans, M.A., is Director, Survey Operations, Survey Operations Department, Medical Group Management Association.

nationwide survey of medical group practices conducted by Mathematica Policy Research during the period March to June of 1978.

One important contribution of the current paper is to update the prior econometric work to the current managed care and policy environment, using a nationwide sample of medical groups responding to two surveys (1997 data) of the Medical Group Management Association: The Compensation and Production Survey and the Cost Survey. Second, the rich data set provided by the MGMA surveys allows us to account for the role of a variety of potential productivity “drivers” within the medical group: ownership form, presence of monitoring mechanisms, size of the group, physician specialty mix, and individual physician characteristics. Third, this research examines a wider range of ownership forms and specialty types of medical group practice—non-primary care single-specialty groups, primary care groups, and multispecialty groups—than previous empirical studies of physician productivity. In contrast, the analyses of Gaynor and Pauly (1990) and Gaynor and Gertler (1995) were restricted to primary care groups and the partnership form of practice. Fourth, by virtue of the broader array of specialty groups in the MGMA sample, the current study will be able to distinguish differential responses to financial incentives and organization design features among primary care physicians (PCPs), medical specialists, and surgical specialists.

OBJECTIVE

The primary purpose of this study is to estimate the effects of physician compensation method on the productivity of individual physicians in medical groups. Secondarily, we also examine the impact of other key organization design features on physician production: (a) physician choice of direct patient care hours worked, (b) number of physicians in the group (“group size”), (c) ownership form, (d) monitoring of physician direct patient care hours, and (e) “at-risk” managed care revenues as a percent of group total revenues.

We posit that individual physicians might either select into or stay in medical groups based on unobserved personal attributes correlated with compensation method that cause them to favor particular organization design features. Accordingly, to obtain unbiased and consistent estimates of the incentive effects of compensation, we include these design features in the behavioral production functions estimated in this article and also test for their potential endogeneity in the production equations.

LITERATURE REVIEW

The theoretical literature in economics clearly establishes that “high-powered” compensation methods, which directly tie the agent’s compensation to individual production, are expected to lead to increased productivity (Lazear 1999; Baker, Jensen, and Murphy 1988). The medical group (firm) seeks multiple objectives: joint economies in production (Alchain and Demsetz 1972), an optimal blend of competition and cooperation among team members (Gaynor 1989; Holmstrom 1982), and maximum valuation of the group’s “brand name” or reputation (Lee 1990; Getzen 1984). The potential tradeoff between financial incentives as “insurance” against the economic risk borne by agents and the beneficial incentive effects of high-powered, piece rate-type compensation further complicates the design of efficient incentive mechanisms (Gaynor and Gertler 1995).

Gaynor and Pauly’s (1990) empirical analysis of medical group partnerships found that, as individual physician compensation changed from zero to 100 percent sensitivity to individual production, physician productivity increased by 28 percent. The same study also found significant diseconomies of group size, and the authors posited that (productivity) incentives decrease with group size.

Prior studies of physician productivity (Gaynor and Pauly 1990; Gaynor and Gertler 1995) have found that the level of physician production in medical groups is positively and significantly related to individual production-based compensation. This finding holds irrespective of whether production is measured per physician hour worked (Gaynor and Pauly 1990) or per physician (Gaynor and Gertler 1995). Another important result in the latter article is that the null hypothesis of *nonjoint* production, or insignificant “team effects” on productive efficiency, is not rejected. This suggests that estimation of individual physician-level production functions within medical groups is appropriate.

Prior empirical work, as reflected in the articles by Lee (1990) and Gaynor and Gertler (1995), explicitly analyzes the *organizational design* of medical groups. If aspects of organization design and productivity are codetermined, one must treat such group-level characteristics as endogenous in estimating physician production functions. Gaynor and Gertler’s results suggested that only compensation method and average price per office visit were endogenous in their production function estimates at the individual physician level.

THEORETICAL MODEL

Structure of Production

The theoretical model underpinning this study is neoclassical production theory, modified to incorporate the behavioral effects on (unobservable) physician effort of personal physician characteristics and of different incentive and organization design mechanisms of the medical group. The production model closely follows the earlier work of Gaynor and Pauly (1990), and takes the form of a “behavioral” production function:

$$\begin{aligned} &\text{Quantity of output produced per physician per unit of time} \\ &= f\{\text{physician hours worked, capital input, non-physician labor} \\ &\quad \text{input, administrative input, personal physician characteristics,} \\ &\quad \text{organizational incentives, and design mechanisms}\} \quad (1) \end{aligned}$$

Organization-level incentives and design features are posited to influence the physician’s (unobserved) level of effort, and these factors are potentially endogenous with respect to the individual physician. This endogeneity might arise through self-selection and “retention” of individual physicians into particular medical groups and their accompanying incentive and design features based on personal characteristics, or by the individual physician’s direct influence on the choice of particular organizational incentives and designs.

The method of individual physician compensation is an important incentive affecting individual effort, as is the size of the group. We hypothesize that increased group size will dampen individual effort through “free-rider” effects, while compensation based on one’s own production will spur increased effort. Increased *organization-level* risk-bearing by medical groups—in the form of capitation payment, withholds, and risk funds—also is posited to influence physician productivity through an indirect, “group-level” effect on the norms and practice styles of physicians, by inducing a shift toward a more conservative, less production-intensive mode of practice.

Potentially Endogenous Choices in Production

We assume that price and marginal cost are *exogenously* determined at the individual physician level, but the variables for own production-based compensation—physician direct patient care hours worked, group size, monitoring, and organization-level financial risk—are treated as potentially endogenous. We draw from economic theory and previous empirical work in

choosing exogenous instruments to explain variation in the above potentially endogenous variables. The details are presented as part of a web-based appendix accompanying this article.

METHODS

Sources of Data

The main sources of the data for this study are the 1998 Cost Survey and 1998 Physician Compensation and Production Survey (both data for 1997) of the Medical Group Management Association (MGMA). The Cost Survey provides annual information on the input factors of production for the medical group: full time equivalent (FTE), nurses, administrative personnel, midlevel providers, physicians (by specialty type), square footage of office space, and occupancy (building) cost for the medical group's facilities.

The Physician Compensation and Production (PCAP) Survey is the source of data on typical physician compensation methods for three strata of physicians. The strata are "new" physicians (less than two years in the group), established specialists, and established primary care physicians, or PCPs (defined as family practice, general internal medicine, or pediatrics by MGMA). The PCAP Survey provides data on the nature of the medical group (single-specialty versus multispecialty), ownership form of the practice (partnership, for-profit business corporation, limited liability company, professional corporation, or not-for-profit foundation), and annual average weekly professional hours in the practice (clinical and nonclinical) for three physician strata: medical specialists, surgical specialists, and PCPs. The measures of monitoring (tracking direct patient care hours) and percentage of the group's revenues in "at-risk" managed care contracts (capitation or fee-for-service subject to withhold) also derive from the PCAP Survey. Individual physician-specific detail (for full-time physicians only) is provided on annual production (including gross charges, ambulatory encounters, RBRVS units), specialty type, gender, years in specialty, and percent of total professional hours spent in direct patient care.

Sample of Medical Groups

The 1997 universe of medical group members of MGMA was 5,725 practices. In comparison, the American Medical Association (AMA) Census of Medical Groups (1996) reported 19,478 medical groups in the United States, comprised of 210,811 physicians. The MGMA membership sample is not a random draw from the universe of medical groups, but tends to overrepresent

larger medical groups. The MGMA sample is the only one of its type with broad national representation, a very large sample of medical groups of all specialty types, and comprehensive data on physician compensation methods, input factors, ownership form, organization design, incentive variables, extent of “at-risk” managed care contracting, and various metrics for production at the individual physician level. The final pooled sample size for this study varies depending on the dependent measure of “production.” In the analyses with charges as the dependent variable, 383 medical groups and 6,129 physicians practicing in those practices reported complete data for the multivariate analyses reported in this article. When RBRVS units are used as the production measure, the sample size declines to 102 groups and 2,237 physicians with complete data.

This sample shrinkage poses a conundrum for the study. The conceptually superior measure of physician production—real resource output, with heterogeneous outputs weighted by their relative resource intensity (i.e., the RBRVS measure)—is available with otherwise complete data only for a comparatively small sample of MGMA practices. As a measure of real output, gross charges (the sum across services of price multiplied by quantity) are potentially distorted by service-specific price variation across groups, but are available for an appreciably larger sample of MGMA groups. Thus, we present both sets of results in this article and discuss differences between them.

Statistical Methods

Individual-level production functions are estimated in two forms: ordinary least squares (OLS) regression and two-stage least squares (2SLS) regression. The functional form used for the production function is the transcendental logarithmic (translog) model adopted by Reinhardt (1972) in estimating physician services production function equations.

The OLS estimation of technical production functions is appropriate if “technology”—that is, the vector of input–output coefficients relating the factors of production to the level of production—and incentives and organization design are exogenously determined with respect to the individual physician. This assumption is often applied in the estimation of “technical” production functions, but is examined within the current study by comparing OLS estimates of production function parameters to 2SLS estimates. Details of the specification tests of OLS and 2SLS models are presented in the web-based appendix.

Robust standard errors of all regression coefficients in the production equations are generated by applying the Huber (1967) and White (1980)

“sandwich” estimator to correct for errors that are not likely to be independently and identically distributed. Second, a correction for clustering of physician observations within medical groups (another source of potential nonindependence in the error term) is applied to produce consistent estimates of standard errors.

Measures in the Empirical Model

Two dependent measures are used alternately in the production function equations: (a) the natural log of gross (billed) charges generated by the individual physician (*lncharge*); (b) the natural log of the resource-based relative value scale (RBRVS) units produced (*lnRBRVS*).

The independent variables in the model are: (1) input factors of production, (2) individual physician characteristics, (3) state-level regulatory factors, (4) local market characteristics, and (5) organizational design and financial incentive features. Only the “supply-side” variables—inputs, individual physician characteristics, and organizational features—are included in the production functions. The role of the other factors is as exogenous variables (“instruments”) to explain organizational choices. Because the focus of this article is on individual physician production, those other factors are explained in the web-based appendix to this article.

Input factors. Factor input measures are the natural log of direct patient care hours of the physician (*lclinhrs*), full time equivalent (FTE) medical administrative personnel per FTE provider (*medadl*), medical patient care personnel per FTE provider (registered nurses, licensed practical nurses, and medical assistants) per FTE provider (*medper*), and available physical capital per provider (using the proxy of the natural log of building cost per FTE provider, *lbcost*).

Individual physician characteristics. Physician characteristics included in the production function model are gender (*female* = 1), self-reported specialty type, and years in practice (experience, or *exp*).

Organization design and incentive features. The variables representing organization-level (medical group) design and incentives are: (a) percent of the individual physician’s compensation from various methods based on individual production; fixed salary; bonus; capitation; equal share of group net income; or other; (b) natural log of total number of FTE physicians in the group (*lnftep*); (c) a dummy variable for whether the group tracks physician direct patient care hours (*monitor*); (d) whether the group is multispecialty (with primary care and other specialties: *multi_ps* = 1; with only primary care

specialties: *multi_p* = 1; with only nonprimary care specialties: *multi_s* = 1; reference category is single-specialty); (e) whether the group is not-for-profit (*o_nfp* = 1 if not-for-profit, 0 otherwise); (f) whether the group is owned by a hospital or hospital system (*ishosp* = 1); its physicians are employed or under contract with a hospital or system (*hospee* = 1); or the group owns a hospital (*hosown* = 1); and (g) the percentage of the group's revenues from "at-risk" managed care contracts (*mgdcare*).

RESULTS

Descriptive Statistics

The means, standard deviations, minima, and maxima for each of the variables in the production function analyses are presented in Table 1. The typical physician's compensation is predominantly production-based (62.7 percent); equal shares compensation accounts for roughly 10.4 percent, and the average "at-risk" group revenue share is 36.9 percent. Roughly 17 percent of the physician sample is female. Not-for-profit ownership arrangements are relatively common in the sample (23.3 percent of physicians), and multi-specialty groups with primary care and other specialties predominate (60 percent). A small minority of the physicians practice in medical groups owned by a hospital or system (<3 percent), and a slightly larger share (~ 4.5 percent) are in medical groups that own a hospital.

Specification Tests

On balance, we conclude that OLS estimation of the behavioral production function is consistent and efficient relative to 2SLS. The specification tests, summarized in the web-based appendix, do not reject that the null hypotheses that compensation method, managed care, ownership form, monitoring mechanisms, and physician hours worked are individually and jointly exogenous. Thus, the presentation of results weights most highly the OLS results. For purposes of comparing this study's 1997 findings with those of the most closely related study—Gaynor and Pauly (1990), based on 1978 data—we also present estimates of the 2SLS model, instrumented for the physician compensation variable only.

Estimates of the Behavioral Production Function

Incentive Variables. In the pooled physician sample (Table 2), the effect of individual production-based compensation on physician productivity is

Table 1: Summary Descriptive Statistics

<i>Summary Statistics: Final pooled sample (for the analysis)</i>					
<i>Variable</i>	<i>#Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
charge	6129	693582.9	546423.1	806	8055000
lncharge	6129	13.18644	.7586887	6.692084	15.9018
prod	6129	62.74399	41.3457	0	100
equal sh	6129	10.40969	24.35687	0	100
cap	6129	.7898515	4.794595	0	50
bonus	6129	2.732273	8.159418	0	100
other	6129	.971565	6.439723	0	100
clinhrs	6129	42.666882	12.2842	0	80
lclinhrs	6129	3.702501	.3523545	.4706284	4.382039
bcost	6129	33442.63	17045.71	310.6111	133320.3
lbcost	6129	10.27163	.6102627	5.738542	11.80051
medper	6129	.9583997	.6997104	0	7.011495
medpersq	6129	1.408045	2.487391	0	49.16106
medad	6129	1.274639	.6594321	0	5.333333
medadsq	6129	2.059486	2.025067	0	28.44445
FTEP	6129	78.79801	96.40075	3	370.8
lnFTEP	6129	3.590156	1.326757	1.098612	5.915663
exp	6129	13.68424	8.88616	0.3	56
expsq	6129	266.2094	316.389	.09	3136
monitor	6129	.3834231	.4862596	0	1
multi_ps	6129	.6000979	.4899179	0	1
multi_p	6129	.033937	.1810819	0	1
multi_s	6129	.0398107	.1955303	0	1
ishosp	6129	.0265949	.1609093	0	1
hospee	6129	.2125959	.4091774	0	1
hospown	6129	.0445423	.2063135	0	1
female	6129	.167564	.3735092	0	1
o_nfp	6129	.2329907	.4227709	0	1
mgdcare	6129	.3688116	.3869997	0	1
expsp	6129	26.41477	23.84926	.2788325	165.4673

positive as expected, but statistically significant only in the regression comparing each of the different “component” methods of compensation (column 2) and in the run with RBRVS as the real output measure (column 5). The production-based compensation coefficient is not statistically significant in the 2SLS regression. Converting the OLS compensation coefficients in column 2 to elasticities, a 10 percent increase in the share of compensation tied to individual physician production is associated with a 0.7 percent increase in productivity. The RBRVS runs in column 5 imply a 2.9 percent increase in real output for a 10 percent increase in the share of production-based

Table 2: Pooled Sample Regression Analyses (robust t-statistics in parentheses)

<i>Analysis Sample:</i>	(1)	(2)	(3)	(4)	(5)
<i>All Physicians</i>					
<i>Dependent Measure:</i>					
<i>Model Specification</i>	<i>LnCharges</i> <i>OLS</i>	<i>LnCharges</i> <i>OLS/components</i>	<i>Elasticities</i> <i>(OLS/components)</i>	<i>Ln Charges</i> <i>2SLS Model</i>	<i>LnRBRVS</i> <i>OLS</i>
Regressors					
Incentive Measures					
production-based	.0007 (1.29)	.0012 (2.33)**	0.0753	.0021 (0.65)	.0047 (1.76)*
bonus		.0060 (3.24)***	0.0164	<i>production-based</i>	
equal shares		.0015 (2.09)**	0.0156	<i>compensation</i>	
capitation		−0.0054 (2.92)***	−0.0043	<i>is instrumented</i>	
other		−.0008 (0.44)			
mgdcare	−.0431 (0.48)	−0.0175 (0.20)		−.0670 (0.28)	−.0612 (0.14)
Factor Inputs					
lbcost	.0716 (1.85)*	.0733 (2.19)**	0.0716**	.0733 (1.93)*	−.0591 (0.56)
medper	−.0460 (1.05)	−.0501 (1.18)	−0.0016*	−.0482 (1.03)	−.8574 (1.41)
medpersq	.0157 (2.07)**	.0152 (2.10)**		.0157 (1.91)*	.1679 (1.64)
medad	−.0938 (1.08)	−.0818 (0.95)	0.0490*	−.0901 (1.05)	−.5462 (0.89)
medadsq	.0573 (2.16)**	.0559 (2.15)**		.0538 (1.98)**	.1665 (0.80)
lclnhrs	.4670 (4.64)***	.4626 (4.55)***	0.4626***	.4558 (4.32)***	.6802 (4.17)***
lnftcp	.0092 (0.37)	−.0021 (0.09)		.0058 (0.21)	−.1745 (1.95)**
monitor	−.0314 (0.79)	−.0445 (1.16)		−.0283 (0.71)	.4602 (1.08)
multi_ps	−.1182 (1.73)*	−.0743 (1.05)		−.1418 (1.59)	.9521 (1.79)*
multi_p	−.2760 (2.69)***	−.2491 (2.44)***		−.3077 (2.46)***	.5513 (1.87)*
multi_s	−.1970 (1.91)*	−.1796 (1.80)*		−.2086 (1.93)*	.3351 (0.88)
o_nfp	−.1861 (3.57)***	−.1782 (3.56)***		−.1725 (2.77)***	.0946 (0.37)

Table 2: Continued

Analysis Sample: <i>Dependent Measure:</i> Model Specification	(1) All Physicians		(2) (Pooled Sample)		(3) Elasticities		(4) (Pooled Sample)		(5) (Pooled Sample)	
	LnCharges		LnCharges		LnCharges		LnCharges		LnCharges	
	OLS		OLS/components		(OLS/components)		2SLS Model		LnRRVS	
ishosp	.1013	(0.62)	.1117	(0.74)			.0977	(0.64)	-.9.0876	(58.00)***
hospec	.0332	(0.86)	.0303	(0.78)			.0422	(0.92)	-.5523	(1.04)
hospow	-.0542	(0.53)	-.0654	(0.69)			-.0583	(0.57)	9.4911	(14.16)***
exp	.0377	(7.57)***	.0369	(7.47)***	0.0025		.0361	(5.91)***	.0102	(1.38)
expsq	-.0009	(7.66)***	-.0009	(7.49)***			-.0008	(6.39)***	-.0005	(1.72)*
exp	-.0063	(5.13)***	-.0064	(5.25)***			-.0063	(5.05)***	-.0002	(0.04)
female	-.1962	(8.17)***	-.1943	(8.23)***			-.1988	(8.14)***	-.1242	(3.21)***
Output Statistics:	R ² = .5444		R ² = .5495				R ² = .5418		R ² = .7060	
	N = 6,129 (383 grps)		N = 6,129 (383 grps)				N = 6,129 (383 grps)		N = 2,237 (102 grps)	

Note: *** denotes $p < .01$; ** denotes $p < .05$; * denotes $p < .10$

compensation. The group-level incentive variable (*mgdcare*) is not significant in any of the pooled sample analyses.

Table 3 stratifies the production function estimation by physician specialty category. Production-based compensation is significantly related to productivity ($p < .10$) for primary care physicians and medical specialists. The group-level incentive, *mgdcare*, is significantly negative only for surgical specialist productivity, measured in terms of gross charges.

Input Factors. The output elasticity of capital input (*lbcost*) is approximately .07. The stratified estimates in Table 3 using charges as the output measure suggest that the elasticity of output to capital input is quite similar across the three specialty categories—ranging from .07 to .11. The output elasticity of physician clinical hours (*lclinhrs*) is the largest among the inputs—approximately .46 in the pooled sample with charges as the output measure and .68 with output measured in RBRVS units. The stratified estimates in Table 3 reveal that this elasticity is statistically significant and relatively large for all three physician-specialty categories—for both charges and RBRVS output measures. The sum of the gross charges-measured output elasticities for all inputs equals 0.58, which implies decreasing returns to scale among this sample of group practices.

Physician Characteristics. Experience (*exp*) is positively associated with productivity, with production increasing at a diminishing rate as experience increases (*expsq*). The statistically significant negative coefficient on the interaction term, *expsp*, indicates that productivity increments associated with increased experience are attenuated among specialties that are more production-intensive (measured by the average level of gross billed charges for all sample physicians in a given specialty). The average elasticity of output with respect to experience (at the sample mean) is relatively small: .0025. The specialty category analyses of Table 3 and the 2SLS results are qualitatively the same in sign and significance.

Female physicians' productivity, measuring output as gross charges, is approximately 20 percent ($p < .05$) lower than that of male physicians in the pooled sample estimates. The difference in RBRVS-measured productivity is smaller—roughly 12 percent, and the stratified analyses of Table 3 show the same pattern.

Medical Group Characteristics. Medical group size (*lnflep*) is not significantly related to productivity in the pooled physician sample or in any of the specialty categories when output is measured as gross charges. However, measuring real output as RBRVS, estimated productivity for the pooled physician sample is reduced by approximately 1.7 percent as group size

Table 3: Specialty Category Sub-sample Analyses (robust t-statistics in parentheses)

Analysis Sample:	(1)		(2)		(3)		(4)		(5)		(6)	
	Primary Care		Medical Specialists		Surgical Specialists							
Dependent Measure:	LnRRVS		LnRRVS		LnRRVS		LnRRVS		LnRRVS		LnRRVS	
Model Specification	LnCharges	OLS	LnCharges	OLS	LnCharges	OLS	LnCharges	OLS	LnCharges	OLS	LnCharges	OLS
Regressors												
Incentive Measures												
production-based	.0007 (0.96)		.0091 (1.83)*		.0014 (1.79)*		.0073 (1.21)		.00003 (0.04)		-.0007 (0.45)	
bonus												
equal shares												
capitation												
other												
mgdcare	.0225 (0.20)		-.7872 (1.04)		-.0421 (0.28)		1.6141 (1.38)		-.3247 (2.60)***		-1.0466 (1.31)***	
Factor Inputs												
lbcost	.1059 (1.64)		.2236 (0.61)		.1161 (2.50)***		-.2068 (0.96)		.0730 (1.13)		-.0779 (0.70)	
medper	-.1740 (1.63)		-1.7775 (1.76)*		-.3250 (1.83)*		-.2837 (0.38)		.1576 (2.51)**		.0405 (0.22)	
merpersq	.0459 (1.18)		.1741 (0.71)		.1501 (1.95)*		-.1898 (0.78)		-.0125 (1.31)		.0255 (0.55)	
medad	.2077 (1.75)*		-3.5551 (2.31)**		-.2752 (2.18)**		.4844 (0.67)		-.0511 (0.48)		.5182 (1.16)	
medadsq	-.0516 (1.33)		1.4028 (2.12)**		.0872 (2.15)**		-.4850 (1.40)		.0467 (1.69)*		-.1541 (1.21)	
lclnhrs	.4305 (2.81)***		.5274 (1.73)*		.5647 (5.8572)***		.5205 (3.84)***		.5034 (3.81)***		.6724 (2.01)**	
monitor	-.0004 (0.01)		1.8091 (2.52)**		-.1485 (1.97)**		.0796 (0.15)		.0889 (1.47)		.1774 (1.24)	
lnftep	-.0117 (0.37)		-.1197 (0.67)		.0233 (0.50)		-.3964 (2.88)**		.0022 (0.05)		-.0940 (1.19)	
multi-ps	-.1569 (1.58)		2.2983 (2.62)***		-.0773 (0.86)		1.1058 (2.48)***		-.1411 (1.21)		-.0297 (0.10)	
multi-p	-.3560 (3.20)***		1.4571 (1.26)		N/A		N/A		N/A		N/A	
multi-s	N/A		N/A		-.3421 (2.89)***		1.3394 (2.67)***		.0381 (0.43)		-.1042 (0.56)	

o-nfp	-.2190 (4.14)***	.6323 (1.51)	-.2113 (2.52)***	-.1052 (0.36)	.0474 (0.57)	-.0300 (0.13)
ishosp	.1418 (0.86)	-9.3753 (19.64)***	0.0473 (0.26)	-9.1648 (35.62)***	.2448 (3.97)***	-9.1902 (51.86)***
hospee	-.0085 (0.17)	-2.1160 (2.61)***	.0693 (1.00)	-.5537 (0.94)	.0853 (1.33)	.1158 (0.78)
hospown	.0299 (0.27)	10.3061 (8.16)***	-.5848 (2.93)***	12.0170 (6.26)***	.0418 (0.34)	N/A
exp	.0488 (5.87)***	.0339 (2.14)**	.0178 (2.33)**	.0214 (1.45)	.0482 (3.73)***	-.0080 (0.30)
expsq	-.0007 (4.91)***	-.0006 (1.32)	-.0006 (3.05)***	-.0010 (1.30)	-.0014 (8.17)***	-.0006 (2.10)**
exsp	-.0229 (3.63)***	-.0133 (1.24)	-.0025 (1.14)	.0019 (0.27)	-.0034 (0.92)	.0069 (0.89)
female	-.2235 (7.08)***	-.1444 (1.90)*	-.1813 (4.66)***	-.1145 (2.63)***	-.2356 (3.87)***	-.1921 (2.36)***
Output Statistics:	N = 2,515 (145 grps) R ² = .3303	N = 921 (34 grps) R ² = .8114	N = 2,082 (174 grps) R ² = .4483	N = 815 (50 grps) R ² = .6189	N = 1,532 (197 grps) R ² = .3399	N = 501 (54 grps) R ² = .8845

Note: *** denotes $p < .01$; ** denotes $p < .05$; * denotes $p < .10$

increases by 10 percent. The stratified specialty category analyses indicate that the effects are largest and statistically significant only for medical specialists, but the direction of effect is consistently negative.

The effect of monitoring physician direct patient care hours (*monitor*) is not statistically significant in either the OLS or 2SLS estimates for the pooled physician sample. Four of six comparisons in the stratified analyses are also statistically insignificant and the two significant comparisons are of opposite sign.

Not-for-profit organizational form (*o_nfp*) is associated with roughly a 17–18 percent reduction in productivity in the pooled sample estimates of Table 2—measuring output as charges. Except for surgical specialists, the same qualitative results emerge in the stratified estimates of Table 3. However, none of the not-for-profit ownership effects is statistically significant using the RBRVS output measure.

Hospital ownership of the medical group (*ishosp*) is not significantly related to physician productivity in the pooled sample using charges as the output measure. However, hospital ownership of the group is significantly negatively related to RBRVS-measured productivity in the pooled sample. This negative relationship between hospital ownership and RBRVS productivity is consistent across specialty categories. Ownership of a hospital by the medical group (*hosppwn*) is associated with significantly higher physician RBRVS-measured productivity for primary care and medical specialists and in the pooled sample estimates.

The estimated effect of multispecialty group structure on productivity differs by output measure and by specialty category. Measuring output by gross charges in the pooled sample, physicians in groups with both primary care and other specialists (*multi_ps* = 1) are somewhat less productive than physicians in single-specialty groups, but the differences are modest and statistically significant ($p < .10$) in only one comparison. Conversely, RBRVS-measured productivity is substantially higher overall for physicians in such groups. This higher estimated productivity applies to primary care and medical specialists, but differences are insignificant for surgical specialists.

Physicians in primary care-only multispecialty groups (*multi_p* = 1) are significantly less productive than those in single-specialty groups when output is measured by gross charges, but somewhat more productive ($p < .10$) in terms of RBRVS units. Overall, measuring output by gross charges, physicians in groups with only non-primary care specialties (*multi_s* = 1) are marginally less productive ($p < .10$). However, RBRVS-measured productivity is not significantly different overall and significantly *higher* for medical specialists.

DISCUSSION

Financial Incentive Effects

The primary purpose of this study was to address the effects of individual physician compensation method and extent of medical group risk-bearing arrangements with managed care plans on physician practice productivity. Our empirical results suggest that physicians in medical groups that base a higher share of the typical physician's compensation on his or her own production are more productive, other factors held constant. Our estimates of the individual components of compensation imply a similar magnitude of effect of compensation method to that reported by Gaynor and Pauly (1990). The current study also sheds some light on the productivity effects of other "nonneutral" compensation methods: equal shares, bonus, and capitation. As one would expect, given the "free-rider" aspects of distributing of equal shares, equal shares compensation has a substantially smaller effect on productivity than individual production-based methods; but the effect is *positive* and statistically significant. Moreover, "bonus" compensation promotes higher levels of productivity, even though the bonuses are not generally tied directly to the physician's own production. For example, 24 percent of the reporting practices base the bonus on administrative or governance responsibility, 19 percent on patient satisfaction (Medical Group Management Association 1998b, p. 22). Individual compensation tied to capitation has the smallest proportionate effect on productivity, but is statistically significant ($p < .01$) and negative, as expected a priori.

The percentage of managed care revenues in at-risk arrangements is not generally significantly related to productivity, with the single exception of surgical specialists and only for the gross charges measure. This is consistent with the prior results of Gaynor and Pauly (1990), who found no significant differences in productivity between partnerships according to whether or not the majority of the medical group's patients were "prepaid."

Factor Inputs

Output elasticities for physician direct patient care hours from the specialty category-specific regressions are arguably more informative than the pooled production function results. The estimated output elasticities for physician time input—which vary from .43 to .67 across specialty categories—are in the neighborhood of the value of .53 estimated by Gaynor and Pauly (1990) for PCPs in their 1978 sample.

The mixed results for the nonphysician labor inputs are disappointing. We suspect that random measurement error is the source of an *errors-in-variables* bias in this article's estimates of the effects of nonphysician medical personnel and administrative personnel on productivity, particularly since the measures of nonphysician inputs are group averages, not physician-specific.

Individual Physician Characteristics

Physician experience and gender are consistently significantly related to productivity, after adjusting for physician specialty. For the pooled sample of all physician specialties, experience is related to comparatively modest effects on productivity: a 10 percent increase in experience relative to the average for all physicians (roughly 14 years) translates into approximately a 0.025 percent increase in productivity.

Differences related to gender are considerably larger—female physician productivity is estimated to be some 12–23 percent lower than that for male physicians, depending on output measure and specialty category. One important explanation for the seemingly lower productivity of female physicians is the measurement of hours worked in this study. The PCAP Survey queried total professional hours worked per week by specialty category, but not actual hours worked by individual physician. The survey did query the individual physician's percent of professional time in direct patient care, however, and that figure was multiplied by the specialty category average for each physician to compute estimated direct patient care hours per week for the individual physician. Thus, if female physicians work fewer total hours than the average, estimated direct patient care hours would be overstated and productivity (output per hour) understated for female physicians. Other factors held constant (specialty type, for example), female physicians did report significantly lower patient care time percentages, by approximately 11 percent.

Organization Design

Our results suggest diseconomies of group size, when output is measured as RBRVS units. Also, our estimated size-elasticity of -0.17 in the pooled physician sample is very close in magnitude to the comparable estimate of Gaynor and Pauly (1990) for primary care partnerships.

We are puzzled by our mixed findings regarding the effect of not-for-profit ownership on physician productivity. Whereas we find the expected significant negative relationship when output is measured as gross charges, the estimated effects are not significant (and even of the wrong sign) for the RBRVS

output measure. If not-for-profits charged systematically lower prices, this could account for the difference, but we tested this conjecture with two different proxies for price (see the web-based appendix), and it was not supported.

Our finding that physicians in hospital-owned medical groups are less productive, when output is measured by RBRVS, is supported by previous empirical work (cf., Cleverley 1997; Morrison 1999). The consistent RBRVS productivity differences across physician specialty categories lend some credibility to a real effect, but the lack of significant effects on gross charges (admittedly a conceptually inferior output measure) encourage further scrutiny in a representative sample of the universe of medical groups. Future research also is merited to explore whether—as implied by our RBRVS results—primary care physicians and medical specialists really are more productive in groups which own their hospital, and, if so, how those productivity gains are achieved.

LIMITATIONS OF THE STUDY

While the authors believe that this study has yielded several important findings of potential interest to practicing clinicians and decision makers in physician organizations, there are certain limitations that must be acknowledged (discussed in detail in the web-based appendix accompanying this article). Certain limitations stem from potential response biases in the MGMA survey data and from specific features of the variables in the MGMA data.

Arguably most important, the inferences from the current study are not necessarily generalizable to the universe of all medical groups in the United States because the study sample of MGMA members was not drawn randomly from a census of medical groups. Comparisons between the MGMA member sample and the AMA Census of Medical Groups indicate that the geographic distribution and specialty type of group (single versus multispecialty) of MGMA members in the aggregate are similar to those of the medical group universe. However, MGMA member groups are, on average larger: their mean size (21.0 FTE physicians in MGMA respondent groups versus 10.7 in the AMA Census) and percentage of groups with 25 or more physicians (15.2 percent versus 5.0 percent in the AMA Census). Thus, the range of observable data for the current study's empirical models is truncated in comparison to what would be available from a random sample of all groups. The relatively small sample of MGMA groups with complete data on RBRVS units also limits the robustness of our productivity estimates. We have dealt with this problem by distinguishing our

charges-based estimates from the RBRVS-based estimates and attempting to logically and empirically reconcile the differences.

The absence of physician-level case mix factors in the MGMA data—a shortcoming shared by all other physician productivity studies of which we are aware—implies that productivity estimates of central interest in this study might be biased if case mix and physician financial incentives are correlated. Also, certain observed productivity differences—for example, by ownership form, physician gender, multispecialty vs. single-specialty group structure—might reflect unobserved case mix differences.

IMPLICATIONS FOR POLICY AND MANAGEMENT PRACTICE

This study does not pretend to offer any direct implications for *public* policy, in the sense of legislation or regulation. However, the authors do believe that several of our findings merit close consideration by managers and organizational decision makers in medical groups.

Physician compensation on the basis of individual production does appear to stimulate physician productivity, even in an environment in which medical groups are assuming increased risk for unanticipated patient care costs. Our estimates of the size of this positive effect of “high-powered” physician incentives vary substantially—depending on the empirical model used, specialty category, and measure of physician output. Thus, our findings provide robust *qualitative* support for the use of individual production-based incentives in spurring physician productivity, and also indicate that bonus and equal share arrangements are productivity catalysts. Our results also confirm that capitation at the individual physician level (as applied to a minor degree by the groups in this study) discourages productivity. Thus, the negative effect of capitation on productivity would have to be counterbalanced by other incentives or design features.

This article’s findings of *diseconomies* associated with group size and decreasing returns to scale are consistent with a recent review article (Pauly 1996), which concluded that the growth of large multispecialty groups is probably not the result of economies of scale. This empirical evidence reinforces the subtlety required to “right-size” medical group practices and indicates that further consolidation and “scaling up” of medical group practices may be counterproductive—at least in terms of productive efficiency. This study’s implications for the efficiency of alternative multispecialty group

structures are not clear-cut, but—emphasizing the results using RBRVS output measures—we conclude that PCPs and medical specialists appear to be more productive in groups with PCPs and nonprimary care specialists, compared to single-specialty groups. Thus, there may be economies of *scope* in combining PCPS and medical specialists within group practice.

The lower measured productivity of female physicians—whether measured in visits per hour or total revenues (adjusted for fee-level differences)—must be viewed with some caution. First, the measures of average hours worked may overstate female physicians' hours worked relative to male physicians'. Second, the case mix of female physicians might differ systematically from that of male physicians. Finally, female physicians might differ from men in practice style and professional preferences in ways that lower *measured* productivity, but that yield equal or superior outcomes. This possibility highlights the importance of developing and counting more comprehensive and clinically relevant measures of the outcomes of physician services.

REFERENCES

- Alchain, A. A., and H. Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review* 62 (5): 777–95.
- American Medical Association. 1996. *Medical Groups in the U.S.: A Survey of Practice Characteristics*. Chicago: American Medical Association.
- Baker, G. P., M. C. Jensen, and K. J. Murphy. 1988. "Compensation and Incentives: Practice vs. Theory." *Journal of Finance: Papers and Proceedings* 43 (July): 593–616.
- Cleverley, W. O. 1997. "Factors Affecting the Valuation of Physician Practices." *Healthcare Financial Management* 51 (12): 71–3.
- Gaynor, M., and P. Gertler. 1995. "Moral Hazard and Risk Spreading in Partnerships." *Rand Journal of Economics* 26 (4): 591–613.
- Gaynor, M., and D. Haas-Wilson. 1999. "Change, Consolidation, and Competition in Health Care Markets." *Journal of Economic Perspectives* 13 (1): 141–64.
- Gaynor, M., and M. V. Pauly. 1990. "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practice." *Journal of Political Economy* 98 (3): 544–73.
- Getzen, T. E. 1984. "A 'Brand Name' Theory of Medical Practice." *Journal of Industrial Economics* 33: 199–215.
- Hillman, A. L., W. Welch, and M. Pauly. 1992. "Contractual Arrangements between HMOs and Primary Care Physicians: Three-Tiered HMOs and Risk Pools." *Medical Care* 30 (2): 136–48.
- Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. Berkeley: University of California Press.

- Holmstrom, B. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13 (Autumn): 324-40.
- Lazear, E. P. 1996. "Performance Pay and Productivity." Working paper no. 5672, July. Cambridge, MA: National Bureau of Economic Research, p. 33.
- . 1999. "Personnel Economics: Past Lessons and Future Directions." *Journal of Labor Economics* 17 (2): 199-236.
- Lee, R. H. 1990. "Monitoring Physicians: A Bargaining Model of Medical Group Practice." *Journal of Health Economics* 9: 463-81.
- Medical Group Management Association. 1998a. *Cost Survey*. Englewood, CO: Medical Group Management Association.
- Medical Group Management Association. 1998b. *Physician Compensation and Production Survey*. Englewood, CO: Medical Group Management Association.
- Morrison, W. H. 1999. "What Are IDSs Doing about Poorly Performing Physician Networks?" *Healthcare Financial Management* 53 (10): 33-6.
- Pauly, M. V. 1996. "Economics of Multispecialty Group Practice." *Journal of Ambulatory Care Management* 19 (3): 26-33.
- Reinhardt, Uwe E. 1972. "A Production Function for Physician Services." *Review of Economics and Statistics* 54 (1): 55-66.
- White, H. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50 (1): 1-25.