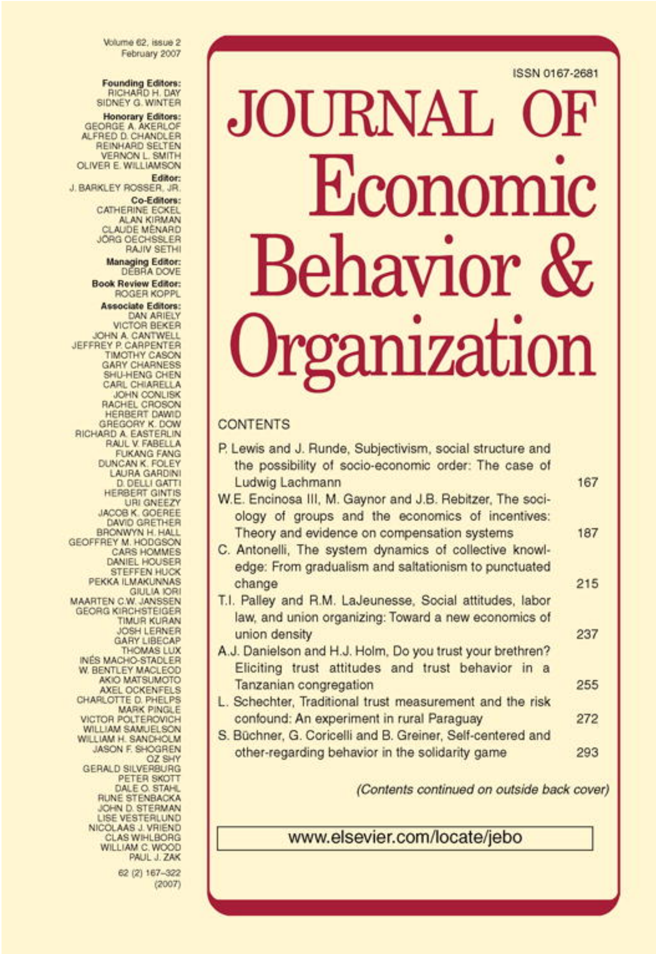
Volume 62, issue 2
February 2007
Founding Editors:
RICHARD H. DAY
SIDNEY G. WINTER
Honorary Editors:
GEORGE A. AKERLOF
ALFRED D. CHANDLER
REINHARD SELTEN
VERNON L. SMITH
OLIVER E. WILLIAMSON
Editor:
J. BARKLEY ROSSER, JR.
Co-Editors:
CATHERINE ECKEL
ALAN KIRMAN
CLAUDE MÉNARD
JÖRG OECHSSLER
RAJIV SETHI
Managing Editor:
DEBRA DOVE
Book Review Editor:
ROGER KOPPL
Associate Editors:
DAN ARIELY
VICTOR BEKER
JOHN A. CANTWELL
JEFFREY P. CARPENTER
TIMOTHY CASON
GARY CHARNESS
SHU-HENG CHEN
CARL CHIARELLA
JOHN CONLISK
RACHEL CROSON
HERBERT DAWID
GREGORY K. DOW
RICHARD A. EASTERLIN
RAUL V. FABELLA
FUKANG FANG
DUNCAN K. FOLEY
LAURA GARDINI
D. DELLI GATTI
HERBERT GINTIS
URI GNEEZY
JACOB K. GOEREE
DAVID GRETHER
BRONWYN H. HALL
GEOFFREY M. HODGSON
CARS HOMMES
DANIEL HOUSER
STEFFEN HUCK
PEKKA ILMAKUNNAS
GIULIA IORI
MAARTEN C.W. JANSSEN
GEORG KIRCHSTEIGER
TIMUR KURAN
JOSH LERNER
GARY LIBECAP
THOMAS LUX
INÉS MACHO-STADLER
W. BENTLEY MACLEOD
AKIO MATSUMOTO
AXEL OCKENFELS
CHARLOTTE D. PHELPS
MARK PINGLE
VICTOR POLTEROVICH
WILLIAM SAMUELSON
WILLIAM H. SANDHOLM
JASON F. SHOGREN
OZ SHY
GERALD SILVERBURG
PETER SKOTT
DALE O. STAHL
RUNE STENBACKA
JOHN D. STERMAN
LISE VESTERLUND
NICOLAAS J. VRIEND
CLAS WIHLBORG
WILLIAM C. WOOD
PAUL J. ZAK
62 (2) 167–322
(2007)
ISSN 0167-2681
JOURNAL OF
Economic
Behavior &
Organization
CONTENTS
P. Lewis and J. Runde, Subjectivism, social structure and the possibility of socio-economic order: The case of Ludwig Lachmann 167
W.E. Encinosa III, M. Gaynor and J.B. Rebitzer, The sociology of groups and the economics of incentives: Theory and evidence on compensation systems 187
C. Antonelli, The system dynamics of collective knowledge: From gradualism and saltationism to punctuated change 215
T.I. Palley and R.M. LaJeunesse, Social attitudes, labor law, and union organizing: Toward a new economics of union density 237
A.J. Danielson and H.J. Holm, Do you trust your brethren? Eliciting trust attitudes and trust behavior in a Tanzanian congregation 255
L. Schechter, Traditional trust measurement and the risk confound: An experiment in rural Paraguay 272
S. Büchner, G. Coricelli and B. Greiner, Self-centered and other-regarding behavior in the solidarity game 293
(Contents continued on outside back cover)
www.elsevier.com/locate/jebo

Journal of Economic Behavior & Organization
Vol. 62 (2007) 187–214

JOURNAL OF
Economic Behavior
& Organization

www.elsevier.com/locate/econbase

The sociology of groups and the economics of incentives: Theory and evidence on compensation systems

William E. Encinosa III [a], Martin Gaynor [b], James B. Rebitzer [c,*]

[a] *Agency for Health Care Research and Quality, United States*
[b] *Carnegie Mellon University, NBER, and Leverhulme Centre for Market and Public Organisation, United States*
[c] *Department of Economics, Weatherhead School of Management, Case Western Reserve University, 11119 Bellflower Road, Cleveland, OH 44106, United States*

Received 9 June 2003; received in revised form 21 February 2005; accepted 12 January 2006
Available online 28 February 2006

Abstract

When working together, people engage in non-contractual and informal interactions that constitute the sociology of the group. We use behavioral models and a unique survey of medical groups to analyze how group sociology influences physician incentive pay and behavior. We conclude that informal interactions among group members influence pay practices and behaviors, but the relationship is complex. No single aspect of group sociology is entirely consistent with all the patterns in the data. Factors emphasized in the economic theory of agency, notably risk aversion, also shape pay policies, but these factors cannot account for all the observed empirical relationships.

JEL classification: D21; D22; J40; J41

Keywords: Incentives; Social norms; Physicians

1. Introduction

Economic models of compensation treat pay practices as a solution to an incentive problem. High levels of performance require high levels of effort and, beyond some minimal point, providing this effort is costly. Firms desiring high levels of performance from employees should therefore

* Corresponding author. Tel.: +1 216 368 5535; fax: +1 216 368 5039.
E-mail address: james.rebitzer@case.edu (J.B. Rebitzer).

0167-2681/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2006.01.001

link economic rewards closely to an individual's productive contribution. Yet this tight linkage is hardly a universal characteristic of pay systems. Firms exhibit enormous variation in the degree to which compensation responds to individual performance. Explaining this diversity is one of the fundamental tasks for the economics of organizations (Baker et al., 1988; Prendergast, 1999).

This paper introduces the concept of "the sociology of groups" into a theoretical and empirical analysis of variation in pay systems. The term "group sociology" refers to the non-contractual and informal interactions that occur between members of work groups. These interactions can take the form of activities (mutual help, mutual monitoring) and psychological experiences (guilt, envy, shame, greed, peer pressure). It is widely believed among economists, sociologists, and human resource professionals that group sociology influences the design of incentive contracts, but theory and evidence regarding these influences has been scarce. We argue below that group sociology influences both the benefits and costs of incentives and therefore the kind of incentive arrangements that firms will use.

Although the economic issues we discuss are quite general, our analysis focuses on pay practices in a narrowly defined setting: medical groups. We concentrate on these organizations for three reasons. First, the medical groups we study tend to be small and to have a flat organizational structure. This simplicity, combined with the fact that the key revenue generating activities (e.g., patient office visits) are regularly recorded for billing purposes, makes it feasible to link compensation to the performance of individual physicians.

The second reason for our focus on medical groups is that our data allow us to observe (rather than infer) the incentive formula that prevails in the group. Our sample of medical groups, like professional groups in general, relies upon administrative rules that specify how partners share in the income generated by other partners in the firm. Information concerning these sharing rules allows us to examine a group's incentive system directly. Third, our data on medical groups contains information about the attitudes and behaviors of individual physicians in the group. The combination of a simple organization, direct information about incentive structures and separately collected information about the attitudes and behaviors of the physicians who make up the group make possible new and novel tests of economic and sociological incentive theories.

We analyze the sociology of medical groups in terms of three types of informal interactions among physicians: (1) the intra-group income comparisons that lead to income norms, (2) the intra group effort comparisons and mutual monitoring that result in effort norms, and (3) mutual help activities.

Income and effort norms play a prominent role in the study of group sociology, but the meaning of these terms is often unclear. In this paper, we use "group norms" to refer to the consequences that interpersonal comparisons of income and effort have on the actions and psychological experiences of group members. The phenomena we have in mind (envy, shame, guilt and peer pressure) are of self-evident importance in social and economic life, but have so far played only a peripheral role in modern microeconomic studies of compensation systems. It is not hard, however, to see how interpersonal comparisons could influence incentive design. In the case of medical groups, if physicians resent having a low rank in the firm's income hierarchy, groups may prefer to avoid workplace tensions by adopting incentive pay schemes that do not create too much of a gap between high performers and others. Interpersonal comparisons of effort will also influence optimal pay practices. Those groups able to marshal sufficient peer pressure and mutual monitoring to support stringent effort norms will also require less in the way of pay-for-performance.

The third type of group interaction we analyze the mutual help activities. Physicians in group settings must decide how much effort they will devote to seeing their own patients versus helping other members of the group. The "helping" activities we have in mind are informal and therefore hard to reward via incentive pay. Some of the helping activities can improve the general reputation of the group and in this way benefit all its members (e.g., making sure that receptionists are polite when they answer the phone), while other helping activities make it easier for other doctors in the group to generate their own revenues (e.g., offering advice to another physician struggling with a tricky case). Both types of informal helping activities will be discouraged by high powered incentives that reward physicians for seeing more patients. Put differently, incentive pay arrangements are costly because they may discourage valuable, mutual help activities. The behavioral logic of mutual help activities is not new to economics; indeed it is formally the same as the familiar multi-task models of optimal incentive design (Holmström and Milgrom, 1991). Classifying these mutual help activities as "sociological" rather than economic is convenient in our empirical setting, however, because this label highlights the role of informal interactions among group members.

Although the analysis in this paper is specific to the institution we study, its implications are far more general. The issue of the use of low-powered versus high-powered incentives occurs in all organizational settings. The use of profit-sharing to provide incentives is also common. Finally, partnerships with readily metered output are quite common (e.g., law, accounting, consulting). Confining our analysis to medical groups is convenient for the reasons stated above, but our results have quite general relevance.

In what follows, we present a series of models designed to capture various dimensions of group sociology. We use these models to generate hypotheses that can be explored using our data. Our findings indicate that the sociology of groups matters for compensation systems, but that no single type of informal interaction is sufficient to explain the patterns in the data. We also find evidence that risk aversion matters, but risk aversion alone cannot explain observed variations in incentive pay. Indeed the empirical results are most consistent with a model in which risk aversion and a number of different informal interactions among group members simultaneously influence the incentive pay policies of organizations.

The paper proceeds as follows. The next part, Section 2, presents a theoretical analysis of incentive pay in groups. Empirical findings are reported in Section 3. The paper concludes by considering the implications of our results for the economic analysis of organizational design.

2. Theory

In this section, we develop a series of models of optimal incentive design in medical groups. In particular, we focus on deriving observable predictions that distinguish the models. Since we have a number of different models with varying details, we focus on an intuitive presentation. Details and derivations are available to interested readers in our working paper (Encinosa et al., 2000).[1] We begin by presenting the set-up of a basic model in Section 2.1. Our general approach for analyzing the incentive design problem is to incorporate a term into agents' preferences that we call the "expected cost of incentives." This term captures factors that optimally lead firms to choose to deviate from high-powered incentives. In each subsequent subsection, we consider several stylized models that lead to such deviations: risk aversion (Section 2.2), income norms (Section 2.3), effort norms (and mutual monitoring) (Section 2.4), and, finally, mutual help activities (Section 2.5). Throughout we contrast the results that derive from the standard model of risk aversion in Section

[1] Available at http://www.heinz.cmu.edu/∼mgaynor/papers/norms_abstract.htm.

Table 1
Model hypotheses

Model	Testable prediction	Table with econometric evidence
Risk aversion		
Large groups choose lower-powered incentives since they give up more insurance by using high-powered incentives than do small groups	$\partial\alpha/\partial n < 0$	Table 3B
	$\partial e/\partial\alpha < 0$	Table 4
	More risk averse doctors choose groups with lower α	Table 3A and B
Group income norms		
Large groups are less likely to adopt low-powered incentives since they are less effective at equalizing incomes than small groups	$\partial\text{Prob}(\alpha = 1/n)/\partial n < 0$	Table 3A
	$\partial e/\partial\alpha > 0$	Table 4
	$\partial\alpha/\partial n > 0$ is possible	Table 3A
Group effort norms		
Small groups more likely to have low-powered incentives because they have stronger effort norms than large groups	$d\text{Prob}(\alpha = 1/n)/dn < 0$	Table 3A
	$e = e^*$	Table 4
	$de/d\alpha = 0$	Table 4
Mutual help activities		
Large groups will use lower-powered incentives since small groups are more effective at maintaining a given level of individual return to mutual help activities	$d\alpha/dn < 0$	Table 3B
	$e < e^*$ and $de/d\alpha > 0$	Table 4
	$d\nu/d\alpha < 0$	Table 5

2.2 with those that derive from the models that capture some aspect of group sociology in Sections 2.3–2.5. These contrasting results can be followed in Table 1.

2.1. Model set-up

Doctors form groups in order to share fixed costs. Some of these costs may be specific to a given specialty (e.g., special purpose equipment or nurses with particular skills) while others are generic to any medical practice (space, equipment, etc.).[2] In every case, however, the group shares a common administrative structure that collects revenues from patients and insurance companies. This common accounting system makes it easy and convenient for physicians to link revenues to the activities of individual physicians. In what follows, we consider the design of optimal compensation arrangements within groups.

Consider a partnership of n doctors. Each individual doctor generates revenue, R, according to

$$R(e_i) = e_i + \varepsilon_i, \tag{1}$$

[2] Sharing space and costs are key parts of the American Medical Association's definition of a medical group practice (Havlicek, 1996).

where e_i is the effort exerted by partner i. We capture the random aspect of revenue by ε_i, a mean zero random variable having variance σ_ε^2. For simplicity, we assume that all individuals in the partnership are identical and the error term is independently distributed across individuals.[3]

The group allows each partner to keep a fraction, α, of her revenues and puts $(1-\alpha)$ into a common pool that is divided equally among the remaining partners. Taking the number of partners in the group to be exogenously determined at n, we can write the expected income of individual i as

$$E(Y_i) = \alpha e_i + (1-\alpha)\frac{\sum_{j\neq i} e_j}{n-1}, \quad 1/n \le \alpha \le 1. \tag{2}$$

We assume that individual doctors derive utility from income and that there is a private cost of effort, $ce_i^2/2$, where c is a positive parameter. Preferences for physician i are represented by[4]

$$U_i = E(Y_i) - \frac{ce_i^2}{2} - E[\text{Cost of Incentives}]. \tag{3}$$

The expected cost of incentives term, E[Costs of Incentives], is central to our analysis. We will use this term to capture costs of high-powered incentives due to risk aversion or group sociology. If the marginal cost of incentives were negligible, the group's incentive problems would be trivially solved by setting $\alpha = 1$.

We assume that each individual physician supplies his or her privately optimal effort given monetary incentives (α), the cost of effort, and the cost of incentives systems. The first-order condition determining effort supply is

$$\frac{\partial U_i}{\partial e_i} = \alpha - ce_i - \frac{\partial E[\text{Cost of Incentives}]}{\partial e_i} = 0 \tag{4}$$

or, rearranging,

$$e_i = \frac{1}{c}\left(\alpha - \frac{\partial E[\text{Cost of Incentives}]}{\partial e_i}\right).$$

Solving this first-order condition for e_i yields an effort supply function, $e(\alpha)$, in which physician effort is an increasing function of α. If the cost of incentives were negligible, then $\alpha = 1$ and the *first best effort* would be $e^* = 1/c$. When the cost of incentives are not negligible, the group's incentive design problem is to choose the value of α that maximizes the utility of a representative member, subject to the individual effort supply function, $e(\alpha)$:

$$\max_{\alpha} U = Y(e) - \frac{ce_i^2}{2} - E[\text{Cost of Incentives}] \quad \text{s.t. } e = e(\alpha). \tag{5}$$

Most economic explanations for variation in incentive pay revolve around the expected costs created by high-powered incentives. In the next sections, we analyze the incentive costs due to risk aversion and to informal group interactions. In some instances the marginal cost of incentives

[3] Homogeneity is reasonable if work propensities and abilities are observable by others in the group. As long as partners share some fraction of income, high output partners will end up subsidizing low output partners. Thus the best any physician can do would be to join a group comprised of other, equally productive individuals (Farrell and Scotchmer, 1988).

[4] All the theoretical results that rely on Eqs. (2) and (3) generalize to settings where revenues are a concave function of effort and the cost of effort function is convex.

increases with group size, while in others the marginal cost of incentives falls with respect to group size. These features will play an important role in our empirical investigation of the determinants of incentive pay practices.

2.2. Incentive costs due to risk aversion

Much of the economic literature on incentive design emphasizes the importance of risk aversion. As a consequence we use this model as our point of comparison with models of group sociology. Since this is such a familiar model,[5] we present only an intuitive summary (again, for a full exposition see Encinosa et al., 2000). Risk aversion on the part of individual physicians makes high powered incentives costly. Risk aversion leads to the marginal costs of incentives being greater in large groups than small groups. The reason for this is simple: large groups offer better insurance (against income fluctuations) due to the law of large numbers. As a consequence, high-powered incentives are costlier for large groups since this means sacrificing more insurance. This is the result we use in our empirical work: the risk aversion model predicts a negative relationship between group size and the strength of incentives. We also use the implication that more risk averse physicians will prefer groups with lower-powered incentives in order to achieve more insurance.

2.3. Group income norms and incentive pay

In this section, we model the costs of incentives due to income norms. Income norms here mean that doctors compare their income with the incomes of others in their group and are made worse off if others have higher income. Our analysis parallels that for risk aversion, but we reach the opposite conclusion. If individuals care about relative income, then the income differentials resulting from high powered incentives can cause tensions within the group.[6] We treat the expected level of these social tensions as the expected cost of incentives due to income norms. Where risk aversion causes the marginal cost of incentives to *increase* with group size, income norms cause the marginal cost of incentives to *decrease* with group size.

In what follows, we briefly present the social psychology that lies behind the income norms model. We then develop a formal model of income norms. Because the structure and intuition associated with this model are not familiar, we describe it in some detail.

2.3.1. The social psychology of income comparisons

Models of group norms highlight the informal, interpersonal comparisons that take place in groups. In order to bring these comparisons into a microeconomic model, we need to specify how individuals assess (and react to) differences between themselves and others. Conventional economic theory does not offer much insight into the ways in which these comparisons are made. We rely, therefore, on three behavioral regularities that have emerged from experimental studies of economic behavior:[7]

[5] See Gaynor and Gertler (1995) or Lang and Gordon (1995) for models of optimal linear incentives with risk averse partners.

[6] Fehr and Schmidt (1999) refer to this as "inequity aversion". This is also sometimes referred to as "status concerns". See also Fehr et al. (2001) and Huck et al. (2003) for models which take up similar issues.

[7] For empirical and theoretical analyses of reference dependence, see Babcock et al. (1996), Frank (1985), Kahneman et al. (1986), and Levine (1993). For an extensive discussion of loss aversion in income see Tversky and Kahneman (1991).

- *Reference dependence*: Utility is determined by absolute and relative income. For any given earnings level, an increase in the earnings of the reference group reduces an individual's utility.
- *Loss aversion*: The marginal utility gain from doing "better" than the reference group (*x* dollars more income) is less than the marginal utility lost by doing "worse" than the reference group (*x* dollars less income).
- *Saliency*: The effect on utility of interpersonal income comparisons increases with an individual's similarity, proximity and exposure to the reference group. Similarly, the effect of interpersonal comparisons increases the more directly that individuals compete for important resources.

2.3.2. *Income comparisons and the marginal cost of incentives*

Our model of interpersonal income comparisons follows directly from the behavioral assumptions presented above. We can capture these relationships with the following utility function for partner i in a group with n partners:

$$U_i = E(Y_i) - \frac{ce_i^2}{2} - E(\text{Inequity}_i), \quad \text{with } E(\text{Inequity}_i) \equiv E\left(\frac{\left\{\beta_1 \sum_{j\neq i}\max(0,\ Y_j - Y_i)\right\} + \left\{\beta_0 \sum_{j\neq i}\min(0, Y_j - Y_i)\right\}}{n-1}\right). \quad (6)$$

Parameters β_1 and β_0 reflect the utility consequences of unequal earnings within the firm. The first expression in braces is the total utility lost to individual i due to others in the reference group having greater earnings. The second term in braces is the utility gain to i from having earnings greater than others in the reference group. We introduce the loss aversion assumption by specifying that $\beta_1 > \beta_0 \geq 0$. We simplify our exposition (with no loss of generality) by setting $\beta_0 = 0$.[8]

Conditional on partner i's income, expected inequity is

$$E[\text{Inequity}_i | Y_i] = \frac{\beta_1}{n-1} \sum_{j\neq i} \int_{Y_i}^{\infty} (Y_j - Y_i) f(Y_j)\, \mathrm{d}Y_j \quad (7)$$

where $Y_j - Y_i = \left[\alpha - \frac{1-\alpha}{n-1}\right][(e_j + \varepsilon_j) - (e_i + \varepsilon_i)]$.

Re-arranging and integrating over ε_i (assuming the ε_i are i.i.d.), we get (for $\alpha \geq 1/n$)

$$E(\text{Inequity}_i) = \beta_1 A\theta + \frac{\beta_1 A}{2}\left(\frac{\sum_{j\neq i} e_j}{n-1} - e_i\right) \quad (8)$$

where $A \equiv \left[\alpha - \frac{1-\alpha}{n-1}\right]$ and $\theta \equiv \int_{-\infty}^{\infty}\int_{\varepsilon_i}^{\infty}(\varepsilon_j - \varepsilon_i) f(\varepsilon_j)\,\mathrm{d}\varepsilon_j\, f(\varepsilon_i)\,\mathrm{d}\varepsilon_i > 0$.Thus, in equilibrium (where $e_j = e_i$), $E(\text{Inequity}_i) = \beta_1 A\theta$. Note that $A \geq 0$ with the equality holding only when

For discussions of saliency see Frank (1988) and Baron and Pfeffer (1994).

[8] If we maintained our assumption that $\beta_1 > \beta_0$ and also allowed $\beta_0 > 0$, then $E(\text{Inequity}_i) = (\beta_1 - \beta_0)A\theta$ in Eq. (8) in equilibrium. None of our results would be changed by this. Another feature of this set-up is that preferences regarding relative income are assumed to be invariant with respect to group size. Alternatively, if we stipulated that $(\beta_1 - \beta_0)$ decreases with group size due to the closer social interactions of small groups, this would only strengthen our reported findings. In unpublished work, we also demonstrate that our results hold when $(\beta_1 - \beta_0)$ increases with group size, provided the increase is not "too fast".

firms adopt the equal sharing rule $\alpha = 1/n$. The expression $A\theta$ represents the expected earnings differential between partners i and j due to differences in *random* shocks. The parameter β_1 captures the degree of social tension resulting from interpersonal income comparisons.

Eq. (8) depicts the relationship between expected inequity, incentive pay and group size. Two aspects of this relationship are important for our purposes. First, if group members do not share income ($\alpha = 1$), the expected income inequality is independent of group size. Second, whenever groups share income ($\alpha < 1$) big groups have higher levels of expected inequity than small groups.[9] It follows immediately from these results that increases in α have a bigger marginal effect on intra-organizational income inequality in large groups. Put differently, because small groups are better than large groups at equalizing incomes (for any $\alpha < 1$), the marginal costs of incentives due to income norms are bigger in small groups.

2.3.3. *Income comparisons and the marginal benefit of incentive pay*

If income norms only influenced the cost of incentives, then the preceding discussion would be sufficient to demonstrate that $\partial\alpha/\partial n < 0$. The story is more complicated, however, because income norms also alter the marginal benefits of incentives by increasing an individual's responsiveness to incentive pay.

The intuition follows directly from the logic of income norms. A doctor can reduce the probability that another doctor earns more than she does by working harder than others in the group. The externality resulting from this return to *relative* effort creates a rat-race type dynamic within partnerships such that doctors over-respond to work incentives (Landers et al., 1996). The effect of this "rat race" on effort is magnified in larger groups because, as we have already discussed, income inequalities are greater in larger groups. For any given level of α, less of the income produced by working harder than others is shared with any other individual in the partnership. Increasing relative effort, therefore, translates into bigger gains in relative income for doctors in larger groups.[10]

We have shown that increases in group size increase both the marginal costs and the marginal benefits of incentives. For this reason, the income norms model does *not* generate a determinate prediction of the effect of group size on incentives. The model does, however, produce a testable prediction: that large groups will be less likely to adopt equal sharing rules.

To see this, consider the result from Eq. (8) that the marginal costs of social tensions resulting from income comparisons is $\beta_1 A_\alpha \theta$. Groups will adopt equal sharing rules ($\alpha = 1/n$) when these marginal costs exceed the marginal benefits of more high powered incentives:

$$\frac{\partial E(U)}{\partial e}\frac{\partial e}{\partial \alpha} \le \beta_1 A_\alpha \theta. \tag{9}$$

The values of β_1 that ensure that (9) holds with equality represent the minimum amount of social tension sufficient to prevent groups from setting α above $1/n$. We can solve for these values of β_1 by evaluating (9) as an equality with $\alpha = 1/n$. The result is

$$\frac{1}{\left(\theta c - \frac{n-1}{2n}\right)(A_\alpha)^2} = \beta_1. \tag{10}$$

9 To see why, consider a doctor who compares herself to another, more fortunate doctor in the group. The less fortunate doctor gets $(1-\alpha)/(n-1)$ of the fortunate doctor's income while the fortunate doctor gets to keep α. Thus, as the group size increases, the income differential between more and less fortunate doctors grow.

10 See our working paper (Encinosa et al.) for the derivation.

Notice that $\partial\beta_1/\partial n = (\beta_1/n(n-1))(1+\theta\beta_1 c A_\alpha^2) > 0$. It takes a greater degree of social tension to sustain equal sharing rules in large groups than small ones. This prediction contrasts with the risk aversion model and forms the basis for an empirical test.[11]

2.4. Group effort norms and incentive pay

If group members care about relative income, they are also likely to care about relative effort. We follow our analysis of income norms, therefore, with an analysis of the interpersonal effort comparisons that support effort norms.[12]

We use the term effort norm to refer to the informal interactions that make it costly for individuals to perform below the level of others in their work group. Effort norms can be sustained by feelings of guilt or shame when not carrying one's "fair share" of the group's work (Kandel and Lazear, 1992). Alternatively effort norms can be the result of informal processes of monitoring and sanctions within the work group. Finally, effort norms can result from the praise a group member receives from working harder than others in the group. Our analysis can incorporate all of these processes and all are likely to operate in real world settings.

We have in mind a setting where it is more efficient for the group to resolve its incentive problems through a combination of incentive pay and peer pressure rather than through legally binding contracts or the threat to dismiss group members who work below the group norm. This presumption is not unreasonable. In many instances work effort is non-contractible because it is assessed in subjective ways that are difficult to record and difficult for a third party to verify. Even if effort were contractible, dismissing individuals who work below the group norm may not be desirable if dismissal or subsequent hiring entails substantial costs.

We write the expected utility of the ith partner as[13]

$$E(U_i) = E(Y_i) - \frac{ce_i^2}{2} - \gamma\left(\sum_{j\neq i}\frac{e_j}{n-1} - e_i\right), \quad \text{with } \gamma > 0. \tag{11}$$

The parameter γ is a positive constant indicating the size of the penalty that sub-norm performers receive. The larger is γ, the greater is the penalty for working below norm.

Changes in γ and in c influence equilibrium work effort by altering the marginal cost of effort. The two parameters, however, are different in one important respect. Reductions in c cause an increase in first-best effort levels because at the margin, the individual is getting more pleasure (or less disutility) out of her work. In contrast, increases in γ bring forth more work effort without altering first-best effort levels. Put differently, under norms the individual works harder not because the work is more palatable, but because the social environment provides sanctions against those who work less than others and rewards for those who work harder than others.

[11] We get this clean result because the "rat-race" effect of income norms on work incentives is negligible under equal sharing rules.

[12] Income and effort norms are conceptually distinct and have distinctive empirical predictions. In real world settings, however, income norms are likely to be most salient where effort comparisons are most salient. In unpublished work, we incorporated this assumption into our analysis of income norms. The resulting income norms model is considerably more complex than the one in this paper, but the core predictions remain unchanged.

[13] This representation of effort norms is adapted from Kandel and Lazear. Our model differs from theirs in that we derive optimal incentive pay rather than assuming equal sharing ($\alpha = 1/n$) prevails. See also Huck et al. (2003).

Effort norms do not create any incentive costs in equilibrium. For this reason, groups will pick the level of α that generates first-best effort levels. When effort norms matter, however, groups can achieve first-best effort with $\alpha < 1$. To see this, recall that first-best effort occurs when $e = 1/c$. Differentiating (11) leads to the following first-order condition for labor supply in the presence of norms:

$$\frac{\partial EU}{\partial e} = \alpha - ce + \gamma = 0. \tag{12}$$

Eq. (12) implies that first-best effort occurs when $\alpha + \gamma = 1$. If, for example, $\gamma = 0.3$, then first-best effort is achieved by setting $\alpha = 0.7$. Thus, effort norms, like income norms, cause groups to operate with $\alpha < 1$. Similarly, the logic of effort norms makes equal sharing rules more likely in small groups. The reason for this is that the incentive implicit in equal sharing rules ($\alpha = 1/n$) falls as group size increases.[14] It is therefore less likely that large groups will have norms (γ) stringent enough to sustain first-best effort with equal sharing.

Empirically, one can distinguish effort norms from income norms by examining the relationship between α and work effort. In settings where only income norms matter, incentives are costly and groups with $\alpha < 1$ operate with less than first-best work effort. Indeed, under income norms one should observe a positive relationship between incentive pay and work intensity. In settings where only effort norms matter, however, all groups will be operating at first-best effort, and there will no relationship between effort and incentive pay.[15]

2.5. *Mutual help and incentive pay*

We have so far considered two different informal interactions among group members that influence incentive pay, income and effort norms. The final aspect of group sociology we consider are mutual help activities. Mutual help activities increase group productivity but do not directly generate revenue for the individual. The presence of valuable mutual help activities thus creates incentive costs that lead groups to set $\alpha < 1$. The intuition is clear.[16] High-powered individual incentives will cause individuals to shirk on (uncompensated) help to others. Doctors will only get a return on their mutual help through the part of their income that comes from the pooled sharing of income across the group. Thus, lower incentives (i.e., greater use of sharing rules) are adopted to encourage more mutual help. Note that sharing rules allow each physician to keep $(1 - \alpha)/(n - 1)$ of the revenues that their helping activities generate. Since these individual returns to mutual help activities fall as group size increases, larger groups will have to operate with lower values of α just to maintain a given level of mutual help.[17]

[14] That is, $\alpha = \frac{1}{n}$ if $\gamma \geq \gamma^* = 1 - \frac{1}{n}$, and $\frac{\partial \gamma^*}{\partial n} = \frac{1}{n^2} > 0$. A second implication of Eq. (12) is that α varies within group size categories only if γ varies across groups.

[15] These *ceteris paribus* assumptions describe special cases chosen to highlight features of the theory. In taking the theory to the data, other more complicating factors must also be considered.

[16] For an explicit derivation see Encinosa et al. (2000).

[17] It is worth noting here that our approach to modeling the "sociology" of mutual help is decidedly economic: lowering α raises the share received of other's productivity, $(1 - \alpha)$. A more radical departure from conventional economic logic would be to posit that explicit incentives degrade intrinsic, altruistic motives (Kreps, 1997).

2.6. Summing up

We have analyzed four explanations for groups choosing low-powered incentives ($\alpha < 1$): risk aversion, group income norms, group effort norms, and mutual help activities. Each of these models yields hypotheses that we can test with our data on medical groups. These are summarized for convenience in Table 1, along with a reference to the tables that contain the econometric evidence relevant to the specific hypothesis.

3. Empirical analysis

In this section, we investigate empirically the relationships summarized in Table 1. We begin by describing our data (Section 3.1). We then analyze the determinants of incentive intensity (Section 3.2) and the effect of incentive pay on work effort (Section 3.3) and mutual help activities (Section 3.4).

3.1. Data

The data we use in this study are from a national random sample of medical group practices conducted by Mathematica Policy Research during the period March–June of 1978. These data are uniquely suited for our purposes because they contain information about key group level characteristics (a measure of incentive pay, group size, characteristics of the group's practice and clientele) as well as survey data from individuals who are members of the group.[18]

Information on incentive pay comes from a question asking each group about the compensation of physicians who had an ownership stake in the practice[19,20]:

> "Excluding fringe benefits, what percentage of the total amount the group distributes to owner physicians is distributed on the basis of productivity?"

The answers to this question are coded in the variable *Incentive Pay I* with responses ranging from 0 to 100. *Incentive Pay I* differs from the empirical measures of incentives used in other studies in that it describes the incentive *policy* of the group without reference to the ex-post realizations of that policy, individual earnings.

Incentive Pay I does not correspond exactly to the theoretically appropriate incentive parameter, α, because it does not include the incentive effect of revenues that individuals receive after the money is pooled and divided among partners. For this reason we construct a second incentive pay variable, Incentive Pay II = Incentive Pay I + (100 − Incentive Pay I)/Group Size). The key

[18] A group was defined as a medical practice having three or more full-time equivalent physicians. For details see Gaynor and Pauly (1990) as well as a data available from the authors upon request in Appendix A.

[19] The question forces the respondent to allocate *compensation* across four categories: productivity, straight salary, equal shares, and other. The emphasis on the allocation of *compensation* rather than gross revenues to the partnership is important. If the question asked about the allocation of gross revenues, then fixed employment costs (if they rose less rapidly than group size) would reduce the fraction of revenues devoted to incentive pay or any other sort of pay.

[20] In the survey individual productivity is defined as "billings, patient visits or some other individual productivity measure". Thus, the survey instrument states if "none of the physician's compensation from the group was directly related to individual physician productivity; it was all based on equal shares of group net income or some similar criteria." Similarly, if the physician's compensation was based entirely on individual productivity, then the survey states "no part of the physician's income from the group was from equal shares, seniority, board certification and the like; it was all based on billings, patient visits or some other individual productivity measure".

variable for our purposes, *Equal Sharing*, is a dummy variable equal to 1 when Incentive Pay II = 100/group size and 0 otherwise.

Table 2 presents descriptive statistics for the distribution of the variables *Equal Sharing*, *Incentive Pay I* and *Incentive Pay II* by group size. Information on the size of the group was collected in six categories, each measuring the number of full-time equivalent physicians: 3–5, 6–7, 8–15, 16–24, 24–49 and 50+.[21] At the time of this survey, physician groups tended to be small: 46 percent of the 794 groups in our sample were in groups with three to five physicians and only 2.4 percent were in groups with 50+ physicians. This last figure is inflated because very large physician groups were over sampled in the original survey.

Column (4) in Table 2 presents the proportion of groups in each size category having equal sharing rules. Increased group size is associated with a reduced propensity to adopt equal sharing rules. Similarly both the mean and median values of *Incentive Pay I* and *Incentive Pay II* increase with size for all except the 2.4 percent of groups in the largest size category. In our view, this break in pattern suggests that many of the largest groups are quite different organizations than smaller groups. Specifically, we suspect that the largest groups in our sample are more likely than smaller groups to be associated with research and teaching entities. In academic medicine, high powered incentives linked to such revenue generating activities as seeing patients are likely to be counterproductive.

3.2. *The determinants of incentive intensity*

The income and effort norms models predict that incentive pay may increase with group size while the risk aversion and mutual help models always predict the opposite relationship. The risk aversion model also predicts that more risk averse physicians will be found in groups having relatively low powered incentives.

In this section, we analyze the relationship between group size, individual risk preferences, and incentive pay. We will show that, consistent with the income and effort norms models, incentive intensity increases with group size. However, consistent with the risk aversion model, groups with low powered incentives tend to have more risk averse physicians.

Table 3 presents estimates of the relationship between group size and incentive pay. The dependent variable in panel A is *Equal Sharing*, a variable taking values of 1 or 0, and all the estimates reported are from probits. The dependent variable in panel B is *Incentive Pay II*, a variable that ranges from $100/n$ to 100, and all the estimates in panel B are from censored normal regressions.[22] Otherwise the two panels are identical and, as we shall see, they contain similar results.

The estimates in column (1) of panel A regress *Equal Sharing* against dummy variables indicating group size categories. The omitted size category is the smallest size category. The regressions also contain a vector of variables that condition on the characteristics of the practice and its clientele. These coefficients are not presented in the paper, but they are available from the authors upon

[21] In the regressions that follow, each group was assigned the mid-point of its size category. The top category was assigned a value of 113.5, a figure derived by assuming that the empirical distribution of the two largest group sizes follows a Pareto distribution. We also experimented with different values (ranging from 60 to 160) and found our basic results were not sensitive to these different assumptions.

[22] The term, censored normal regression, refers to a generalization of tobits in which each observation can be left and right censored at a different point. We treat Incentive Pay II as being left censored at Incentive Pay II = $100/n$ and right censored at Incentive Pay II = 100.

Table 2
Descriptive statistics for group size and incentive pay measures

Distribution of group size[a]			Groups with equal sharing rules[b]		Distribution of Incentive Pay I[c]			Distribution of Incentive Pay II[d]			Groups with no sharing	
(1) Group size category	(2) Number of groups in size category	(3) Percent groups in size category (%)	(4) $\alpha = 1/n$ (%)	(5) Mean	(6) 25th percentile	(7) Median	(8) 75th percentile	(9) Mean	(10) 25th percentile	(11) Median	(12) 75th percentile	(13) $\alpha = 1$ (%)
3–5	365	46	54.2	31.2	0	0	70	48.4	25	25	77.5	19.7
6–7	100	12.6	42.0	38.9	0	22.5	75	48.3	15.4	34.4	78.8	20.0
8–15	153	19.3	21.6	55.4	15	60	100	59.3	22.4	63.5	100	28.1
16–24	85	10.7	23.5	55.5	10	60	95	57	14.4	62	92.3	22.3
25–49	72	9.1	6.9	65.2	50	70	95	66	51.3	70.8	95.1	20.8
50+	19	2.4	31.4	29.7	0	20	46	30.4	0.88	20.7	46.5	10.5
All	794	100.0	38.3	42.49	0	37.5	90	52.67	25	40.8	90	21.5

[a] Group size refers to the number of physicians in the group.
[b] Equal sharing occurs when Incentive Pay II = 100/Group Size.
[c] Incentive Pay I is the percent of compensation (excluding fringe benefits) distributed to owner physicians on the basis of individual productivity.
[d] Incentive Pay II = Incentive Pay I + ((100 − Incentive Pay I)/Group Size).

Table 3
Determinants of Incentive Pay

Independent variables	Dependent variable: Equal Sharing						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: determinants of Equal Sharing[a]							
Group size: 6–7	−0.245 (−1.409)						
Group size: 8–15	−0.717 (−3.842)						
Group size: 16–24	−1.020 (−3.909)						
Group size: 25–49	−1.056 (−3.264)						
Group size: 50+	−0.367 (−0.768)						
Inverse group size		4.369 (4.611)	4.269 (2.803)	3.646 (3.317)	3.675 (2.298)	5.728 (5.417)	5.903 (4.453)
Importance of regularity of income 1 = no importance; 4 = very important							0.311 (2.431)
Importance of pay for performance 1 = no importance; 4 = very important							−0.985 (−8.042)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of groups	583	583	299	454	231	500	413
Likelihood ratio test	109.500	104.570	61.280	69.040	10.890	104.520	149.390
Log-likelihood	−326.831	−329.297	−171.142	−245.144	−154.023	−269.565	−192.978

Independent variables	Dependent variable: Incentive Pay II						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: determinants of incentive pay parameter[b]							
Group size: 6–7	6.105 (0.615)						
Group size: 8–15	26.072 (2.586)						
Group size: 16–24	34.634 (2.703)						
Group size: 25–49	26.494 (1.726)						
Group size: 50+	−21.377 (−0.863)						
Inverse group size		−147.980 (−2.775)	−182.538 (−2.153)	−84.184 (−1.440)	−207.468 (−1.732)	−201.687 (−3.655)	−199.274 (−3.417)
Importance of regularity of income 1 = no importance; 4 = very important							−12.507 (−2.289)
Importance of pay for performance 1 = no importance; 4 = very important							49.083 (8.895)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	−1658.687	−1663.007	−795.486	−1377.330	−548.929	−1449.729	−1167.866
Number of groups	583	583	299	454	231	500	413

In panels A and B, column (3) refers to groups comprised solely of owner physicians. Estimates in column (4) are for groups having full-time managers. In panels A and B, estimates in column (5) are for groups comprised only of internists, pediatricians and/or general practioners. In panels A and B, estimates in column (6) are for groups having no productivity guidelines. The models in panels A and B also include variables describing characteristics of the group practice and its clientele. Group practice variables included in the regressions measure the percent of revenues from HMOs, the percent of revenues from outside specialty referrals, the percent of the group that is board certified, whether the group is multi- or single specialty, the specialty composition of the group (percent physicians in internal medicine, general practice, specialty surgeon, OB/GYN, and pediatrics). Patient variables include the percent of clientele who are white, the income distribution of clients, and the percent who are covered by Medicare and Medicaid. The coefficients (and descriptive statistics) for these variables are available from the authors upon request.

[a] All estimates in panel A are probits. Numbers in parentheses are *z*-scores.

[b] All estimates in panel B are censored normal regressions and values in parentheses are *t*-statistics.

request.[23] The negative sign on the size dummy variable coefficients indicate that larger groups are *less* likely to adopt equal sharing rules than smaller groups. The magnitude of these effects, however, is not directly interpretable from the probit coefficients. Converting the coefficients to derivatives, we find that increases in group size substantially reduce the probability of equal sharing. Moving from the smallest group size (three to five physicians) to the next larger (six to seven) reduces the probability of equal sharing rules by 8.5 percentage points. Moving from the smallest group size to the fourth largest (16–24 physicians) reduces the probability of equal sharing by 28.5 percentage points. This is a substantial change given that the probability of equal sharing is 38 percent for the sample as a whole. The effect of being in the largest group size (50+) is also negative, but the magnitude is small. As noted above, this may be due to the presence in this category of groups involved in academic medicine. The small number of groups in this largest size category (19) makes it difficult to estimate size effects precisely. The 95 percent confidence intervals for this coefficient range from −1.30 to 0.57. Thus, while we cannot reject the statement that the true size coefficient for groups greater than 50 is zero, we also cannot reject the thesis that the true coefficient is the same as that for groups having 24–49 physicians.

Estimates in column (1) of panels A and B tell similar stories. Increases in group size are associated with an increase in incentive intensity as measured by *Incentive Pay II*, although the marginal effect of size falls as the size of the group increases. The largest groups appear to break from this pattern because they have very low incentive parameters. Here, as in panel A, the standard error of this estimate is very high, and the 95 percent confidence interval ranges from −70.06 to 27.30.

Columns (2) in panels A and B repeat the analysis of incentive pay presented in column (1), but in these estimates group size dummy variables are replaced with the continuous variable *Inverse Group Size*. Log-likelihood tests do not reject the restrictions implicit in the use of the inverse of group size variable.[24] The point estimates are also close to those derived from the dummy variable specification.[25] For example, moving from a group with four physicians to one with 20 physicians reduces the estimated probability of equal sharing rules by 32 percentage points. This change in group size is associated with an increase in incentive pay roughly comparable to that reported in column (1) of panel B.[26]

One possible concern with the analysis presented thus far is that we may not be using the appropriate size variable. The theoretical discussion focused on the number of partners while our group size variables measure the total number of physicians in the group. For groups in the bottom

[23] Group practice variables included in the regressions measure the percent of revenues from HMOs, the percent of revenues from outside specialty referrals, the percent of the group that is board certified, whether the group is multi- or single specialty, the specialty composition of the group (percent physicians in internal medicine, general practice, specialty surgeon, OB/GYN, and pediatrics). Patient variables include the percent of clientele who are white, the income distribution of clients, and the percent who are covered by Medicare and Medicaid.

[24] Twice the difference in the log-likelihood between columns (2) and (1) in panel A of Table 3 is 4.93 and 8.64 in panel B. For a 95 percent confidence level, the critical value of the chi-square distribution with four degrees of freedom is 9.48. We therefore cannot reject the restrictions.

[25] These results hold even when we control for the possible endogeneity of group size. Our instrument for group size is the average response of doctors in a group to the question “Taking everything into consideration, what group size do you prefer?” The IV model produced very similar results. These IV estimates are available from the authors upon request.

[26] It is possible that the estimated relationship between group size and measures of incentive pay in column (2) (panels A and B) of Table 3 are artifacts due to measurement error in the top size category. We think this possibility is remote, however, because we get the same pattern when we use dummy variables to indicate group size (see column (1) in Table 3). We also believe that the results are not likely due to mechanical relationships between group size and incentives. Substituting Incentive Pay I for Incentive Pay II yields results quite similar to those in column (2) of Table 3.

five size categories this discrepancy is not likely to pose a substantial problem because a large fraction of the physicians working in a group are owners (in 63.54 percent of the groups with three to five doctors, all physicians had an ownership stake in the group). Things are different in the very large group practices. Here only 25.93 percent of groups were comprised solely of physicians who were owners/shareholders. As a check on the importance of this measurement error, we re-estimated the results in column (2) for groups composed only of physician owners. This reduces the number of groups in the sample from 583 to 299. These estimates are reported in column (3) of panels A and B. Restricting the sample in this way does not alter the relationship between group size and the probability of having equal sharing rules.

An alternative explanation for the results in Table 3 might be that introducing high powered performance incentives requires a fixed expenditure on systems that monitor individual performance. If these expenditures are substantial, small groups may be unwilling to make the investment.

This fixed cost explanation relies on two assumptions: (i) there are substantial fixed costs associated with implementing incentive pay systems, and (ii) these costs increase as incentive intensity increases. We contend that neither of these assumptions are likely to hold in the medical groups we study.

The fixed costs of incentive systems are likely to be negligible because much of the information needed to calculate incentives is information about physician billings and/or patient visits that are already collected by groups in order to obtain reimbursement from insurers. In addition, low powered incentives entail neither more nor less cost than high powered incentives. Equal sharing rules still require that revenues are appropriately measured, placed in a common pool and accurately divided. There is no reason to believe that $\alpha = 1/n$ is any less costly, in this regard, than $\alpha = 1$ or any value in between. Indeed, it is easy to imagine settings where $\alpha = 1$ is the least cost system because each physician simply keeps the revenues he or she generates. Since medical groups *must* select a value for α, and since the fixed costs of implementing α do not fall as α falls, there is no reason why fixed costs should lead small groups to favor low powered incentives.

We do not have any data on the fixed cost of incentive systems in medical groups. We can make some indirect inferences; however, if we assume that whatever the added costs of setting up high powered incentives, they are smaller for groups having full-time managers. If this assumption is correct, then under the fixed cost hypothesis, the relationship between size and incentive intensity should be attenuated for groups with full-time managers.

The results for groups with full-time managers are presented in column (4) of Table 3. Comparing columns (2) and (4) of panel A, we do find a small reduction in the magnitude of coefficient on *Inverse Group Size*, but it is well within the 95 percent confidence interval for column (2).[27] Similarly, the absolute value of the coefficient on *Inverse Group Size* falls between columns (2) and (4) of panel B, but the value in column (4) is also well within the 95 percent confidence interval of the coefficient in column (2).[28] The bottom line is that we do not reject the hypothesis that the relationship between size and incentive intensity are unchanged between columns (2) and (4) of panels A and B.

The results presented in Table 3 might also reflect the importance of joint production between physicians in a practice. Consider a hypothetical practice composed of two hand surgeons and an anesthesiologist. If these three doctors perform their surgeries together, then an equal sharing rule might only reflect the fact that it is impossible to attribute revenues to any single individual.

[27] This confidence interval ranges from 2.5 to 6.22.

[28] This confidence interval ranges from −252.717 to −43.242.

We can investigate the importance of joint production by restricting attention to specialties where revenues are generated by individual physicians seeing patients individually in their offices. For this reason we re-estimated column (2) for groups composed entirely of general practitioners, internists and/or pediatricians, specialties for which joint production is unlikely to be important. These estimates, presented in column (5) of Table 3, reveal the same negative relationship between size and incentive intensity observed in column (2).[29]

Our discussion thus far assumes that groups choose α and that individuals then choose their optimal effort levels. An alternative incentive instrument would be to require that individuals, on average, achieve a certain level of performance as a condition of employment. We can investigate the importance of such productivity guidelines for our key results using data collected from our survey of group practices. Groups were asked to respond to the following yes/no question:

> "Does the group have a formal policy or explicit guidelines on expected productivity for physicians?"

Roughly 17 percent of groups reported having productivity guidelines.

Column (6) in Table 3 presents estimates of incentive equations for groups without productivity guidelines. For these groups there is the same strong negative relationship between size and incentive intensity that we observed in earlier estimates. In unpublished estimates for groups having productivity guidelines, however, *Inverse Group Size* does not have a significant effect on the probability of adopting an equal sharing rule. This pattern is what we would expect if productivity guidelines are substituting for incentive pay. It is worth noting, however, that only 123 groups had productivity guidelines and the coefficients estimated for this subset are therefore not measured with much precision.[30]

In our theoretical analysis, we assumed away cross-group differences in risk aversion and assumed that individuals had no other preferences regarding the form of compensation. In the real world, however, cross-group differences in risk aversion and other preferences regarding the nature of pay are likely to matter in the design of optimal incentive systems. Ceteris paribus, a group composed of more risk averse members will have lower powered incentives. Similarly, groups whose members believe productivity based pay is "fair" will tend to choose high powered incentives.

We examine these possibilities in columns (7) of Table 3. We assess risk aversion by taking the group's average response to a question asking about the importance of regular income (*Importance of Regularity of Income*).[31] Group average responses to a similar question asking about the

[29] Adding specialty dummies to the estimates in column (5) of panel A in Table 3 produces the following coefficient (t-statistic) on inverse group size: 3.108 (1.895). In panel B, the analogous coefficient (t-statistic) on inverse group size is −138.93 (−1.158). Both these results are within the 95 percent confidence interval for the coefficients presented in the text, but in both cases the t-statistics fall. We conclude from this that adding the specialty dummies to the regression forces us to rely on cell sizes too small to yield precise estimates.

[30] Column (6) in Table 3 look only at groups without productivity guidelines. The corresponding coefficient (std. error) of the variable inverse group size for groups with productivity guidelines is −1.862022 (2.894975) for panel A and 256.9904 (186.2838) for panel B.

[31] The wording of the question was "listed below are some factors that physicians might consider when choosing a new practice. Please check in the columns below how important each of the factors . . . is to you." The factor measuring risk aversion is "regularity of income (lack of fluctuation)". Responses are coded in a four point scale with 1 = of little or no importance and 4 = very important. Interestingly, regressing the group average importance of regular income on the group average importance of linking pay and productivity yields an R^2 of nearly zero (0.0015).

importance of linking pay to productivity (*Importance of Pay for Performance*) is used to capture preferences regarding the form of pay.

In panel A of Table 3, the coefficient on *Importance of Regularity of Income* was positive and statistically significant while the coefficient on *Importance of Pay for Performance* was negative and significant. In panel B, the analogous coefficients were also statistically significant and their signs reversed. These results indicate that preferences regarding the form of compensation matter: groups with more risk averse physicians are more likely to operate with low powered incentives and groups whose members prefer linking pay to productivity operate with high powered incentives.

Before leaving the discussion of risk preferences, it is important to reconsider the predictions of the risk aversion model regarding equal sharing rules. Under risk aversion, rational groups should never choose equal sharing rules because when $\alpha = 1/n$, the marginal cost of additional incentives is zero. The only way to reconcile equal sharing with the risk aversion model is to assume that groups selecting equal sharing rules are doing so because they have mistakenly chosen incentives that are too low-powered. If we denote optimal incentives by α^*, it is easy to show that $(\alpha^* - 1/n)$ increases with group size. Put differently, the cost of mistakenly choosing equal sharing rules is greater in larger groups. If the likelihood of mistakes decreases with the costs of a mistaken action, the risk aversion model *can* be reconciled with the results of panel A of Table 3.[32] Such a reconciliation is not possible with the results of panel B. We conclude, therefore, that the patterns in Table 3 *taken together* cannot be explained solely on the basis of risk aversion. We can also conclude, however, that the patterns in the data are consistent with a model in which risk aversion is one of a number of factors determining incentive pay.

3.3. Incentive pay and work intensity

Models of income norms and effort norms both predict that smaller groups will be more likely to adopt equal sharing rules than larger groups. The two models can be distinguished, however, by the relationship between equal sharing rules and work intensity. In income norms models, the social costs created by income comparisons cause groups to adopt incentives that generate less than first best effort levels. In effort norms models, groups set $\alpha < 1$ because they can achieve first-best effort levels with low powered incentives. Thus, if effort norms alone were accounting for the patterns described in Tables 2 and 3, we should observe no relationship between incentive pay and work intensity.

As discussed above, the Mathematica survey asked individual physicians to report the number of office visits and office hours they worked in the week prior to the survey. In Table 4, we use the number of office visits (columns (1), (2), (5), and (6)) and the number of office hours (columns (3), (4), (7), and (8)), as measures of work intensity.

Columns (1) and (2) of Table 4 examine the relationship between *Equal Sharing* and *Log of Office Visits*. In interpreting the coefficient on *Equal Sharing*, it is important to note that it is the result of both incentive effects and selection effects. Groups with more incentive pay or more stringent work norms will, by virtue of these features, elicit higher levels of work intensity from their members than other groups. These same groups will, however, tend to attract physicians better able to tolerate high levels of work intensity. Both incentive and selection effects operate in our model, but in ways subtly different than in other settings. Lazear (1989) argues that individuals with a low marginal cost of effort will be drawn to firms offering piece rates. In our set-up,

[32] The mutual-help model, in contrast, does not predict that $(\alpha^* - 1/n)$ increases in n.

Table 4
Work intensity and incentive pay[a]

Independent variables	(1) Log office visits[b]	(2) Log office visits[c]	(3) Log office hours[b]	(4) Log office hours[c]	(5) Log office visits[b]	(6) Log office visits[c]	(7) Log office hours[b]	(8) Log office hours[c]
Equal Sharing	−0.003 (−0.039)	−0.164 (−2.765)	0.020 (0.424)	−0.016 (−0.315)				
Incentive Pay II					0.001 (0.445)	0.001 (2.057)	-0.001 (−0.875)	0.000 (0.277)
Inverse group size		0.248 (0.295)		0.375 (0.576)		0.150 (0.176)		0.366 (0.567)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of groups	204	285	204	289	204	285	204	289
R^2	0.3189	0.35826	0.31834	0.32936	0.3197	0.2789	0.2419	0.1547

These models also include variables describing characteristics of the group practice and its clientele as described in Table 3. The coefficients (and descriptive statistics) for these variables are available from the authors upon request.

[a] Numbers in parentheses are *t*-statistics calculated using White's formula for heteroskedasticity consistent standard errors.

[b] Estimated for groups in the smallest size category.

[c] Estimated for groups not in the smallest size category.

individuals for whom high effort is not very costly will be drawn to groups with lots of incentive pay *or* to groups with little incentive pay but stringent work norms.

The equation in column (1) is estimated for groups in the smallest size category, three to five physicians. The coefficient on *Equal Sharing* in this equation is small and imprecisely measured. The estimates in column (2) are for groups with more than five physicians. In contrast to column (1) the coefficient on *Equal Sharing* is −0.1637 with a *t*-statistic above 2.7. Thus, ceteris paribus, equal sharing in the smallest groups has virtually no effect on the number of office visits while in larger groups equal sharing reduces the number of office visits in a week by roughly 16 percent. Similar results are found if we substitute *Incentive Pay II* for the *Equal Sharing* variable (see columns (5) and (6)).[33]

Columns (3), (4), (7), and (8) in Table 4 estimate the effect of incentive pay on *Log Office Hours*. In contrast to the results for *Log Office Visits*, none of the coefficients on our measures of incentive pay, *Equal Sharing* or *Incentive Pay II*, are large or statistically significant. We do not know why it is that incentive intensity influences the number of office visits rather than the hours actually worked, but the result does not appear unreasonable in the context of physician practices.[34] One can imagine physicians making rigid commitments to each other about when they will be in the office, if for no other reason than to ensure appropriate coverage and to make the problem of scheduling patient visits easier. Given these commitments, the hard to observe dimension of work effort is the number of visits physicians choose to fit into a given number of office hours. Taking more time with patients who might need or desire more attention may be pleasurable to a physician, even though the slower churn of patients reduces revenues. It makes sense to address this sort of agency problem with incentive pay (rather than rigid time allocations) because each physician can use his or her judgment in allocating time and attention across the pool of patients.

Taken together, the results in Table 4 indicate that work intensity, as measured by office visits per week, increases with incentive pay for large groups but not for small groups. We interpret these results as indicating that our effort norms model may account for variation in incentive pay for groups in the smallest size category, but not for the variation observed in larger groups. Put differently, our results suggest that in small groups, effort norms reduce the costs due to free-riding that accompany low powered incentives. Groups in larger size categories, however, pay a productivity price for choosing equal sharing rules.[35]

We find additional support for our interpretation of Table 4 from data on the distribution of productivity guidelines across groups. Large groups with equal sharing rules are more than twice as likely to adopt productivity guidelines as other groups. This pattern suggests that larger groups with equal sharing rules are trying to find alternatives to costly incentive pay and ineffective group norms.[36]

It is interesting to speculate why small groups with $\alpha = 1/n$ might sustain first best effort levels while larger groups might not. The explanation we emphasized in our theoretical analysis is

[33] Very similar results are found if we re-estimate columns (1), (2), (5) and (6) conditioning on office hours worked in the previous week. Similar results are also found if we pool across group size categories and include a size/incentive interaction term.

[34] Less than 35 percent of the physicians in our sample indicated that they set their own hours.

[35] The finding of "shirking" in larger groups with low powered incentives is consistent with Nalbantian and Schotter's (1997) experimental study of productivity under group incentives with groups having 12 members.

[36] For groups in the smallest size category, 12.6 percent of those with equal sharing rules and 13.7 percent with more high powered incentives adopted productivity guidelines. For bigger groups, the analogous figures are 31 percent and 13.3 percent.

Table 5
Incentives and mutual help

Independent variable	Dependent variable: intra-group consults	
	(1)	(2)
Incentive Pay II	−0.0025 (−2.0158)	
Equal Sharing		0.2244 (2.3013)
Inverse group size	−1.9002 (−2.2802)	−2.1081 (−2.5441)
Constant	Yes	Yes
Number of groups	413	413
R^2	0.076	0.0744

All estimates are OLS. Numbers in parenthesis are *t*-statistics calculated with heteroskedasticity consistent standard errors. In columns (1) and (2) we can reject the hypothesis that all coefficients are jointly significant at the 2% and 1% confidence levels, respectively. Physicians were asked how frequently they consulted with other doctors in the group about their patients: 3*x*/day; 2*x*/day; 1*x*/day, 2–3*x*/week; 1*x*/week; 1*x*/month <1*x*/month. The dependent variable is the average response for the group expressed as consults per day. The 25th, 50th and 75th percentile of this variable are 0.840, 1.428, and 2.166, respectively. These models also include the variables describing characteristics of the group practice, and its clientele as described in the notes to Table 3. The coefficents and descriptive statistics for these variables are available from the authors upon request.

perhaps the simplest. When $\alpha = 1/n$, incentive intensity falls as group size increases. Thus less stringent effort norms are required to sustain first-best work effort in smaller groups. In terms of our earlier notation, this would mean that $\partial\gamma^*/\partial n > 0$. An alternative explanation may be that the psychological sanctions for violating group effort norms (guilt, shame, etc.) are more keen in smaller groups.[37] Untangling these competing explanations would require data that is not currently available.

3.4. *Incentive pay and mutual help activities*

In this section, we consider whether incentive pay discourages mutual help activities. The focus of our investigation is the frequency with which physicians consult one another about cases. Making oneself available for consultation is just the sort of activity highlighted by our model of mutual help. Agreeing to discuss another partner's case is likely to help the other doctor deliver medical services to his or her patients. Increasing incentive pay increases the opportunity cost of providing this help. Evidence that physicians in groups with high powered incentives engage in less consultation would thus be indirect evidence that concern over mutual help activities may also shape incentive pay decisions.

The Mathematica survey asked individual physicians how frequently they consulted with other doctors in their group about their patients. We use the group average response to this question to indicate the amount of time and energy doctors devote to mutual consultation. Table 5 presents estimates of the relationship between incentive pay and the frequency of consultation within the group.

The consultation equation is presented in column (1). The negative and significant coefficient on *Incentive Pay II* suggests that increases in incentive pay are associated with reductions in the frequency with which doctors in the group consult one another. The estimated coefficient appears

[37] In terms of our theoretical framework, this would suggest that the saliency of social sanctions falls as group size increases, or equivalently $\partial\gamma/\partial n < 0$.

to be behaviorally as well as statistically significant. In a group of four physicians, increasing incentives from equal sharing to full incentive pay reduces the frequency of consultations by 0.19 per day. This represents a greater than 13 percent reduction from the mean number of consultations per day (1.5). Column (2) replaces *Incentive Pay II* with *Equal Sharing*. The positive coefficient on *Equal Sharing* suggests that physicians in groups with equal sharing rules have an average of 0.22 more consultations per day than physicians in other groups. Unfortunately our data do not allow us to tell if this fall-off in consultations would produce a significant change in the quality, cost or quantity of medical care provided by the group. Similarly, we do not know if the increase in incentive pay causes a fall off in consultation or if groups for whom consultation was more important (because of some unobserved feature of the practice) would choose a more "group oriented" pay system.

3.5. An unresolved question: why do groups choose $\alpha = 1$?

The descriptive statistics in column 13 of Table 2 indicate that a sizeable proportion of medical groups operate without physicians sharing any income. According to the models in this paper, groups should choose $\alpha = 1$ only when high powered incentives are costless. It is easy to show that in the presence of any incentive costs (from risk aversion, income norms, effort norms or mutual help activities) groups should choose $\alpha < 1$. Are groups with $\alpha = 1$ groups for whom incentives are costless? Our data on risk aversion suggests not. We find that 47 percent of groups with $\alpha = 1$ have members who are on average more risk averse than the mean for groups with $\alpha < 1$.

If costless incentives do not explain the observed mass points at $\alpha = 1$, what might? We suggest three plausible explanations. The first is that $\alpha = 1$ is simply an approximation to choosing an optimal α close to 1. Groups, for example, might optimally set $\alpha = 0.9$, but find it easier just to round up to 1.

A second hypothesis relies on ideas concerning procedural justice. Perhaps group members value high powered incentives not simply because of the rewards they expect, but because linking pay to individual productivity appeals to their sense of justice. These procedural concerns could be modeled by entering α directly as a term in the utility function. If fairness created strong enough preferences for high α, groups might choose $\alpha = 1$ even when there are non-trivial costs to high powered incentives.

We have some indirect evidence concerning individuals' preferences for pay procedures. Our results in column (7) of Table 3 suggest that preferences for having "income dependent upon your own productivity" correlate with the degree of incentive intensity in the group.[38] This relationship is consistent with the idea that preferences for certain pay procedures (as opposed to pay outcomes) may explain why some groups choose $\alpha = 1$ even when incentives are costly.

A third explanation for setting $\alpha = 1$ can be found in an expanded model of effort norms. Specifically, we can account for $\alpha = 1$ if we allow individuals to react differently to group members who work harder and less hard than they do.

[38] A linear probability model that regresses a dummy variable for $\alpha = 1$ against our measure of incentive preferences yields a coefficient of 0.15 (t-statistic = 6.05). This suggests the 0.4 point difference between the two groups would increase the probability of equal sharing rules by six percentage points. This result is quite substantial considering that roughly 20 percent of the sample choose $\alpha = 1$. These results are not likely due to a correlation between risk aversion and preferences for linking pay to performance. If we introduce our measure of risk aversion into the linear probability model, the relationship between α and pay preferences is unchanged. Similarly, if we introduce our measure of pay preferences into the equations in column (7) of Table 3, we find that it also has a statistically and economically significant effect on the incentive policies chosen by the group. Including this variable does not significantly alter the results in Table 3.

To implement this norms model, we write the utility of individual i as

$$E(U_i) = E(Y_i) - \frac{ce_i^2}{2} - \frac{\gamma}{n-1}\sum_{j\neq i}[h(e_j - e_i) + g(e_i - e_j)] \tag{13}$$

where $h(\,)$ is the disutility of working too little compared to others in the group (due to shame, guilt or peer pressure) and $g(\,)$ is the disutility of working harder than others in the group (due to anger at subsidizing weak performers). We give structure to this norms model by stipulating $h(x) = g(x) = 0$, if $x \leq 0$; $0 < h(x)$, $g(x)$ for $x > 0$; and $0 < h'(x)$, $g'(x)$ for $x \geq 0$.

All group members exert identical effort, and there are three possible equilibria. In the first, individuals supply the level of effort they would if there were no norms, $e = \alpha/c$. In the second equilibrium, effort norms cause individuals to supply more effort than they would in the absence of norms. In the third equilibrium, norms are dysfunctional and cause individuals to supply *less* effort than they would in the absence of norms. These last two effort equilibria are described in Eq. (14) below (recall that first best effort is $e^* = 1/c$):

$$e = \begin{cases} \dfrac{\alpha + \gamma h'(0)}{c} & \text{norms increase effort and groups achieve 1st best effort at } \alpha < 1. \\ \dfrac{\alpha - \gamma g'(0)}{c} & \text{norms reduce effort and groups do not achieve 1st best effort at } \alpha = 1. \end{cases} \tag{14}$$

The case where effort norms increase effort is a more general specification of the case we analyzed above. When social or psychological processes make individuals experience disutility when working less hard than others, the group can sustain higher levels of effort than would otherwise be possible given the value of α.[39] For this reason, groups can reach 1st best effort with $\alpha < 1$.

In the last case, people work *less* hard in the presence of effort norms. This equilibrium describes a setting in which group members are so strongly averse to feeling “cheated” by working harder than others that the group gets stuck at a low level of effort.[40] In this third equilibrium group, members say, in effect, “why should I put in extra effort if nobody else is.” The group can try to offset the dysfunctional norms by increasing α, but they cannot achieve first best effort levels even when $\alpha = 1$. Thus, under dysfunctional effort norms, firms might find it beneficial to set $\alpha = 1$, even when incentives are costly.

4. Conclusion

What accounts for the variation in incentive pay in the medical groups we study? We consider four possible explanations: (1) risk aversion, (2) income norms, (3) effort norms, and (4) mutual help activities. The last three models describe the various informal interactions between group members that make up the “sociology” of the group.

Our empirical findings indicate that none of the four candidate models of incentive pay can individually account for all the patterns in the data. The relationship between group size and

[39] Define $e^{**} = \alpha/c$ as equilibrium work effort in the absence of effort norms. Now consider a group in which every member works $e > e^{**}$. The effect on individual i’s utility of deviating below the group equilibrium is $-[\alpha - ce + \gamma h'(0)]$. Thus the group can sustain $e > e^{**}$ as long as $-[\alpha - ce + \gamma h'(0)] = 0$ or, equivalently, $e = (\alpha + \gamma h'(0))/c$.

[40] Define $e^{**} = \alpha/c$ as equilibrium work effort in the absence of effort norms. Now consider a group in which every member works at $e < e^{**}$. The effect on individual i’s utility of deviating above the group equilibrium is $-[\alpha - ce - \gamma g'(0)]$. Thus the group can sustain $e < e^{**}$ so long as $-[\alpha - ce - \gamma g'(0)] = 0$ or, equivalently, $e = (\alpha - \gamma g'(0))/c$. It is clear from this effort supply function that groups will not achieve 1st best effort even with $\alpha = 1$.

incentive pay is consistent with income norms or effort norms, but not risk aversion or mutual help activities. The relationship between incentive pay and work effort suggests that effort norms may explain "equal sharing rules" in the smallest groups, but not in larger ones. The relationship between incentive pay and intra-group consultations is consistent with mutual help activities, while the sorting of physicians across groups suggests that risk aversion also plays a role in setting pay policies. None of the four explanations can account for groups choosing $\alpha = 1$, although this result may be consistent with some expanded models of norms.

The conclusion we draw is that even in the very simple organizations we study, the determination of incentive pay in groups is too complex to be fully captured by any of the candidate models taken individually. The data may be, however, consistent with a richer model in which the sociology of the group and the risk aversion of individual members together determine the group's incentive pay policy. Indeed, combining sociological and economic models of behavior may prove generally fruitful for understanding the many settings (in medicine and elsewhere) where individuals both produce and are paid in groups.[41]

Acknowledgments

We would like to thank Daron Acemoglu, Ashish Arora, Linda Babcock, George Baker, Erik Brynjölfson, John Engberg, Rob Fairlie, Reuben Gronau, Bengt Holmström, Steve Jones, Bentley MacLeod, John Mullahy, John Penrod, Steve Pischke, David Sappington, Scott Stern, Lowell Taylor, Bill Vogt, and participants in seminars at Carnegie Mellon, Harvard, Illinois, MIT, Princeton, UC Santa Cruz, the NBER conference "What Do Employers Do?" and the World Congress on Health Economics for helpful comments and discussion. We are grateful to Liz Fowler for editorial assistance. The usual caveat applies.

Appendix A

Variables	Mean	Std. Dev.	Notes
Average years experience	19.319	7.236	Average years since medical school for physicians in the group
Average tenure in group	13.293	7.022	Average years with medical group for physicians in the group
Equal sharing	0.384	0.487	A dummy variable equal to 1 if Incentive Pay II = 1/Group Size
Frequency of intra-group consult	1.485	0.836	Group average of the number of consults per day that an individual physician reports having with other physicians in the group about medical problems of his or her patients
Importance of regular income	2.826	0.616	Group average of individual's assessment of the importance of regularity of income (lack of fluctuation) in deciding to move to a new practice. Scoring ranges from 1 = no importance to 4 = very important

[41] For an interesting recent experimental study, see Fehr et al. (2001). Berman (2000) demonstrates that economic analysis of groups or clubs can also be useful for understanding the sociology of religious groups.

Appendix A *Continued*

Variables	Mean	Std. Dev.	Notes
Importance of pay for performance	3.116	0.661	Group average of individual's assessment of the importance of basing pay on individual productivity. Scoring is the same as preceeding variable
Incentive Pay I	42.460	41.373	The percent of compensation (excluding fringe benefits) that the group distributes to owner physicians on the basis of individual productivity
Incentive Pay II	52.670	33.604	*Incentive Pay I* + (100 − Incentive Pay I)/group size D30
Inverse group size	0.157	0.091	Inverse of the number of full time equivalent physicians in the practice
Log of office hours	3.248	0.393	Log of the average of physician office hours worked in the week prior to the survey
Log of office visits	4.477	0.510	Log of the average of group office visits handled by physicians in the week prior to the survey
Multi-specialty group	0.574	0.495	A dummy variable equal to 1 if multi-specialty group and 0 if single specialty
Percent group board certified	75.221	29.131	
Percent group gen. practitioner	33.257	39.029	
Percent group general surgeons	8.115	11.991	
Percent group internal med.	24.126	31.608	
Percent group OB\GYN	11.643	25.059	
Percent group pediatricians	12.385	26.002	
Percent of revenues from HMO's	8.947	24.077	Percentage of your patient care revenues from prepaid or capitation payments? Data is collected in four categories (1 = under 25, 2 = 25–49, 3 = 50–74, 4 = 75+) and group is assigned mid-point of the category
Percent patient incomes $10–15 K	43.835	18.746	
Percent patient incomes $15 K+	30.889	20.020	
Percent patients under medicaid	10.801	10.001	
Percent patients under medicare	22.453	13.279	
Percent patients who are white	81.316	17.207	
Percent revenues from outside specialty referrals	13.736	18.472	What percent of this office's practice is made up of specialty referrals from physicians outside the group?
Within group variation in office hours	6.042	5.057	Within group standard deviation of physician office hours in the week prior to the survey
Within group variation in office visits	10.312	8.038	Within group standard deviation of patient office visits handled by physicians in the week prior to the survey

Appendix A *Continued*

Variables	Mean	Std. Dev.	Notes
Within group income variation σ_g	13818.720	9116.384	Within group standard deviation of income earned from group in 1977. Income data was recorded in 11 categories: 1 = \$0–\$9999; 2 = \$10 K–\$19,999; 3 = \$20 K–\$29,999; 4 = \$30 K–\$39,999; 5 = \$40 K–\$49,999; 6 = \$50 K–\$59,999; 7 = \$60 K–\$69,999; 8 = \$70 K–\$79,999; 9 = \$80 K–\$89,999; 10 = \$90 K–\$99,999; 11 = \$100 K+; 12 = did not practice with this group. Individuals were assigned mid-point of categories 1–10. In category 11 individuals were assigned income of 105 K. Mean income is \$58787.42

The survey of medical groups was conducted by Mathematica Policy Research. A group was defined as a medical practice having three or more full-time equivalent physicians. Information was collected at the group level by interviewing either the office manager or, if none were available, anyone else who had the necessary information. In addition, individual doctors practicing in the group were surveyed. In no group were more than 11 physicians interviewed. The final sample included 957 groups and 6353 physicians practicing in those groups. Five medical practice specialties were sampled: general practice, internal medicine, pediatrics, general surgery, and obstetrics/gynecology. Roughly 60 percent of all office-based physicians practiced in these specialties.

References

Babcock, L., Wang, X., Lowenstein, G., 1996. Social comparisons in negotiations that reflect a self-serving bias. Quarterly Journal of Economics 111, 1–19.

Baker, G.P., Jensen, M.C., Murphy, K.J., 1988. Compensation and incentives: practice vs. theory. The Journal of Finance 43, 593–616.

Baron, J.N., Pfeffer, J., 1994. The social psychology of organizations and inequality. Social Psychology Quarterly 57, 190–209.

Berman, E., 2000. Sect subsidy and sacrifice. Quarterly Journal of Economics 115, 905–954.

Encinosa, W.E. III, Gaynor, M., Rebitzer, J.B., 2000. The sociology of groups and the economics of incentives: theory and evidence on compensation systems. Carnegie Mellon University, Pittsburgh, March, Unpublished manuscript.

Farrell, J., Scotchmer, S., 1988. Partnerships. Quarterly Journal of Economics 103, 279–298.

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition and co-operation. Quarterly Journal of Economics 114, 817–868.

Fehr, E., Klein, A., Schmidt, K.M., 2001. Fairness, Incentives and Contractual Incompleteness. Working Paper No. 445, CESifo, Munich, Germany.

Frank, R.H., 1985. The demand for unobservables and other non-positional goods. American Economic Review 75, 101–116.

Frank, R.H., 1988. Passions Within Reason: The Strategic Role of Emotions. W.W. Norton and Company, New York.

Gaynor, M., Gertler, P.J., 1995. Moral hazard and risk spreading in medical partnerships. RAND Journal of Economics 26, 591–613.

Gaynor, M., Pauly, M.V., 1990. Compensation and productive efficiency in medical partnerships: evidence from medical group practice. Journal of Political Economy 98, 544–573.

Havlicek, P.L., 1996. Medical Groups in the U.S: A Survey of Practice Characteristics. American Medical Association, Chicago.

Holmström, B., Milgrom, P., 1991. Multi-task principal-agent analyses: incentive contracts, asset ownership, and job design. Journal of Law, Economics and Organization 7, 24–52.

Huck, S., Kübler, D., Weibull, J., 2003. Social norms and economic incentives in firms. University College London, March, Unpublished manuscript.

Kahneman, D., Knetsch, J.L., Thaler, Richard, 1986. Fairness as a constraint on profit seeking: entitlements in the market. American Economic Review 76, 728–741.

Kandel, E., Lazear, E.P., 1992. Peer pressure and partnerships. Journal of Political Economy 100, 801–817.

Kreps, D.M., 1997. Intrinsic motivation and extrinsic incentives. American Economic Review 87, 359–364.

Landers, R.M., Rebitzer, J.B., Taylor, L.J., 1996. Rat race redux: adverse selection in the determination of work hours in law firms. American Economic Review 86, 329–348.

Lang, K., Gordon, P., 1995. Partnerships as insurance devices: theory and evidence. Rand Journal of Economics 26, 614–629.

Lazear, E.P., 1989. Pay equality and industrial politics. Journal of Political Economy 97, 561–580.

Levine, D.I., 1993. Fairness, markets and ability to pay: evidence from compensation executives. American Economic Review 83, 1241–1259.

Nalbantian, H.R., Schotter, A., 1997. Productivity under group incentives: an experimental study. The American Economic Review 87, 314–341.

Prendergast, C., 1999. The provision of incentives in firms. Journal of Economic Literature 37, 7–63.

Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: a reference dependent model. Quarterly Journal of Economics 106, 1039–1061.