The Journal of Socio-Economics 39 (2010) 171–179

Contents lists available at ScienceDirect

The Journal of Socio-Economics

journal homepage: www.elsevier.com/locate/soceco

Incentive pay: Productivity, sorting, and adjacent rents

Noah Patrick Stefanec*

Miami University (OH), 114 University Hall, Hamilton, OH 45056, United States

ARTICLE INFO

Article history:
Received 16 July 2009
Received in revised form 26 October 2009
Accepted 11 February 2010

JEL classification:
J3: Wages
Compensation
Labor costs

Keywords:
Incentive pay
Rent
Incentive-altering preferences

ABSTRACT

I model and empirically test the hypothesis that higher-quality workers prefer performance pay to time-rate based pay because they realize rent upon two different dimensions: Explicit and implicit rents. First, higher-quality workers are outright more productive than their lower-quality counterparts, earning them explicit rent (Curme and Stefanec, 2007). Second, these same factors of production facilitate the unobserved heterogeneity for incentive workers, earning them implicit rent because they can produce a given level of output with less effort. I find strong empirical evidence to confirm that these implicit rents exist and I measure them at 1.5–3.4 percent of average real hourly earnings.

1. Introduction

The provision of incentives within a firm via the compensation design is a topic on which canonical labor economic theory is rather silent. The classical worker is assumed to be paid their marginal product in real terms, and effort is generally taken as exogenous. More realistically, employees are of heterogeneous ability and function in endogenous effort environments, making the contract design instrumental in allocating incentives and attracting the right workers.

The compensation design problem can most generally be thought of as a tradeoff; the degree to which an employee's pay is fixed and based upon inputs versus the degree to which an employee's pay is variable and based upon their individual output. Output-based pay is commonly referred to as incentive pay because it motivates the worker to increase production. It is well documented that incentive workers earn more than similar time-rate workers; for instance, workers paid on individual performance earn roughly 7–14 percent more than workers paid by the hour, *ceteris paribus* (Pencavel, 1977; Seiler, 1984).

In addition to motivating production, incentive pay attracts more productive workers to the firm. When offering fixed-pay, firms suffer a lemons problem. Offering a salary (for example) decreases the overall quality of the workforce but properly designed incentive schemes can act as an "invisible hand" to systematically attract the right types of workers (Akerlof, 1970; Miller, 1992). Empirically speaking, this creates a problem when estimating the returns to incentive pay as sorting effects related to ability cause an omitted variable bias. Specifically, the returns to incentive pay are overstated because more productive workers tend to gravitate towards work environments in which pay more accurately reflects their individual output. Employing longitudinal data is the most common way to remove the heterogeneity bias which contaminates and overstates the cross-sectional estimates of the returns to incentive pay, but this method only differences-out average productivity disparities under the assumption that heterogeneity is time-invariant (Parent, 1999).

Less addressed by empirical researchers is that after accounting for average productivity differences (i.e. ability), high-quality workers are still able to produce a given level of output with less effort than low-quality workers, implying they may be able to earn some additional underlying rents under incentive pay schemes (Lazear, 2000). These excess returns should be empirically measurable, and the documentation of such rents is the aim of this particular paper. Section 2 addresses the prior research regarding both canonical and behavioral approaches to measuring the returns to performance pay in greater detail.

In the vein of behavioral economics literature, the constituents of worker quality have been the subject of a recent debate. For instance, one-dimensional cognitive skill most commonly enters the production function generically as *ability*, but economists have recently begun to place more emphasis on the overall *quality* of workers. The literature has already begun to emphasize the importance of such approaches principally because employers have a

* Tel.: +1 513 785 7714; fax: +1 513 785 7766.
E-mail addresses: stefannp@muohio.edu, noahstefanec@hotmail.com.

1053-5357/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.socec.2010.02.010

vested interest in the behavioral preferences of their employees (Bowles et al., 2001b). In the words of Bowles et al. (2001a): "One cares about the [incentive-altering] preferences of those with whom one interacts in part because these preferences affect. . .the effectiveness of the incentives that one may deploy" (p. 155).

In theory, both skill and "incentive-altering preferences" may be becoming more accepted as entering the production function directly. Empirically, skill can be controlled for by employing Armed Forces Qualification Test Scores. Measures of non-skill traits, such as the Rotter Scale of Externality and the Rosenberg Self-Esteem Scale, can proxy incentive-altering preferences. The contention here is that these factors may enter the production function directly (generating explicit rent) as well as indirectly via an endogenous effort function, which may also produce some subjacent, implicit rents.

Section 3 provides the theoretical justification for such an argument whereas Section 4 outlines the empirical procedure designed to test the theory. The results found in Section 5 offer strong support for the main hypothesis of this paper; namely, that pay-for-performance workers earn subjacent rents which range from an additional 1.5–3.4 percent of average hourly earnings depending upon which factor of production is being discussed. Section 6 concludes.

2. Literature review

The theory of incentive pay is straightforward: Tying a worker's compensation to output rather than input supplies individuals with the incentive to increase productivity. With regards to the body of research on the subject, the literature has recently begun to gravitate away from the employment of classical approaches to the study of incentive pay and begun to focus more on the behavioral aspects of such schemes. Both of these veins of literature will be summarized in turn.

Canonically, there exists a large, seminal body of both theoretical and empirical literature regarding the effects of individual incentive pay. Theoretically speaking, the literature has reached a consensus that performance pay increases the earnings and productivity of workers (Booth and Frank, 1999; Lazear, 2000) and the so-called "ratchet effect," or the disincentive caused when firms intertemporally adjust either the performance standard or the pay rate, can be offset by the appropriate multi-period piece-rate scheme (Lazear, 1986; Gibbons, 1987; Carmichael and MacLeod, 2000). These outcomes are Pareto-efficient in that they maximize profits, induce workers to provide high effort, and stimulate quality sorting. Empirically, Pencavel (1977) finds that piece-rate workers earn roughly 7 percent more than similar time-rate workers whereas Seiler (1984) finds that individuals paid upon performance earn about roughly 14 percent more than their similar time-rate counterparts, *ceteris paribus*. Moreover, Oettinger (2001) finds that worker effort varies positively with the commission rates of Major League Baseball (MLB) stadium concession vendors.

While time-rate pay is based upon inputs such as hours, it is important to note that incentive pay is based upon output, and therefore places greater emphasis on issues such risk-preference, unobserved worker heterogeneity, monitoring costs, and quality concerns. In the interest of setting up a "straw man" model as a demonstration, it was first shown that a firm can maximize profit by paying risk-neutral workers a commission rate equal to one hundred percent of net revenue (Armstrong and Lorentzen, 1982). Unfortunately, transaction costs associated with paying on performance will generally abate the efficiency of such a scheme.

To exemplify, it is rather difficult to make employees full residual claimants because of hidden action (moral hazard) and hidden information (adverse selection). First, Holmstrom (1979) discusses the tradeoff associated with risk and incentives; risk-averse workers prefer variable pay schemes less than risk-neutral workers because of the uncertainty associated with variation in output. Holding effort constant, an individual's output may vary simply due to exogenous factors, such as the state of the market. Second, regardless of risk, low-ability individuals generally prefer a higher proportion of their total compensation to be from fixed-pay, and will typically leave output-based pay voluntarily. Third, measuring output is another hurdle. It is often too costly and unprofitable to observe the output of individual workers, and/or the measure of an individual's output may be too noisy for incentive pay to be effective. Finally, there would appear to be a tradeoff between quality and quantity as the incentive to produce more implies less time may be given to quality considerations (Lazear, 1986, 1995).

This is not to say that an incentive pay scheme cannot be successful to some degree if these transaction costs are minimized. In his classic study, Lazear (2000) observes the Safelite Glass Corporation and finds evidence of both incentive and sorting effects associated with a switch the company made from input-based pay to output-based pay. The classic draw scheme was designed such that workers were offered a fixed-rate of pay if they fell below a certain productivity threshold, but were paid by the piece beyond a particular point.

As a result of the switch, productivity rose by 44 percent. Due to the incentive effect, the average worker increased output from 2.6 units per day to 3.1 units per day. As a result of the sorting effect, the system reduced the turnover of high-effort workers and attracted more productive workers to the firm. Sorting should play a large role if one believes that better workers know their own "type" and self-select into incentive pay firms; "better workers" have more of a chance of differentiating themselves and increasing their earnings when pay more accurately reflects output.

In a related vein, a debate has begun in the literature as to what exactly constitutes a "worker quality." In the past, unobserved worker-heterogeneity has been vaguely labeled as "ability" and this ambiguity has spawned a more behavioral approach to earnings determination and labor market sorting. With respect to earnings determination, most researchers traditionally limit their focus to one-dimensional cognitive aptitude (ability), but more recent empirical findings confirm that non-skill traits (e.g. ambition, pessimism) are also priced in the labor market and can be just as valuable to workers as hard skill (Bowles et al., 2001a,b). For lack of a better term, these researchers have labeled non-skill traits as being "behavioral" in nature (Bowles and Gintis, 2000). Put differently, both ability *and* behavioral preferences may constitute what can be more cohesively thought of as "worker quality" rather than generic "ability."

In more economic terms, the insights of Bowles et al. improve upon the canonical assumptions established by Leon Walras (1874): Specifically, that worker effort is exogenous and transactions are costless. Walras' assumptions were long accepted by economists until Joseph Schumpeter (1934) observed that different personality types (i.e. entrepreneurs) are better suited to deal with disequilibrium. Moreover, Ronald Coase (1937) famously noted the importance of endogenous effort, incomplete contracts, and the comparative advantage firms enjoy in reducing transactions costs.

The theoretical application of behavioral traits to various labor market outcomes has been somewhat limited, however. Bowles et al. (2001a) formally model the impact of non-skill traits on wages in what they dub the contingent renewal model of the employee relationship. Their main underlying assumption is sociological: Human beings prefer to interact with others whose preferences favor the pursuit of one's own objectives. In particular, they argue employ-