



---

Compensating Wage Differentials and Unobserved Productivity

Author(s): Hae-shin Hwang, W. Robert Reed, Carlton Hubbard

Source: *The Journal of Political Economy*, Vol. 100, No. 4 (Aug., 1992), pp. 835-858

Published by: The University of Chicago Press

Stable URL: <http://www.jstor.org/stable/2138690>

Accessed: 17/08/2009 11:46

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Political Economy*.

# Compensating Wage Differentials and Unobserved Productivity

---

Hae-shin Hwang

*Texas A&M University*

W. Robert Reed

*University of Oklahoma*

Carlton Hubbard

*Texas A&M University*

It is well known that the inability to observe workers' full labor market productivity can bias estimates of compensating wage differentials. This paper attempts to determine how serious this bias is likely to be. It adopts a stochastic framework of workers' tastes over job attributes and models their equilibrium wage-job attribute choices. Workers' productivity is assumed to consist of observed and unobserved components. Applying the standard estimation methodology, we find that the degree of bias can be surprisingly large. On the basis of our analysis, we conclude that contemporary labor market studies are likely to severely underestimate workers' willingness to pay for job attributes. This has implications for a number of applications of compensating wage differentials, including value of life studies.

We have benefited from discussions with Badi Baltagi, Donald Deere, John Lott, Doug Marcoullier, Carol Moore, Rick Muehler, Kevin Murphy, Brooks Pierce, and Eric Schansberg. We are very grateful to an anonymous referee whose comments served to substantially improve earlier versions of this paper. Any errors that remain are the sole responsibility of the authors.

[*Journal of Political Economy*, 1992, vol. 100, no. 4]

© 1992 by The University of Chicago. All rights reserved. 0022-3808/92/0004-0007\$01.50

## I. Introduction

It is well known that the inability to observe workers' full labor market productivity can bias estimates of compensating wage differentials derived from cross-sectional labor market data (Brown 1980; Duncan and Holmlund 1983; Garen 1988). Unfortunately, techniques for correcting the effects of unobserved productivity heterogeneity are not generally applicable.<sup>1</sup> As a result, most researchers continue to estimate compensating wage differentials using the standard wage equation framework, implicitly assuming that this bias is not significant.<sup>2</sup> This paper investigates the validity of this assumption.

The analysis proceeds as follows. Section II provides a brief explanation of why unobserved productivity biases estimates of compensating wage differentials. Section III investigates the size of this bias for the case in which workers face a linear hedonic wage constraint and preferences are Cobb-Douglas. It is analytically demonstrated that the bias is a function of three factors: (i) the average share of total hourly remuneration taken in the form of wages, (ii) the proportion of wage dispersion due to the differing tastes of workers, and (iii) the degree of unobserved productivity heterogeneity. This result pro-

<sup>1</sup> Brown (1980) attempts to correct for unobserved productivity heterogeneity by using longitudinal data. He uses a fixed-effects model to control for unobserved productivity differences across individuals that do not change over time. His proposed correction technique does not substantially improve estimates that are judged to be "wrong-signed." Duncan and Holmlund (1983) also use longitudinal data and follow a similar estimation procedure, with better results. While studies of this sort represent improvements over standard cross-sectional studies, their applicability is restricted by the availability of longitudinal data sets that include the relevant nonwage job attribute variables. In most cases, this is a binding constraint. Garen (1988) develops an innovative approach for correcting for unobserved productivity heterogeneity by using instrumental variables. His technique requires that the number of instrumental variables equal the number of nonwage amenities included in the hedonic wage equation. Unfortunately, it is difficult to obtain appropriate instruments. In the Garen study, two variables are used: a proxy for risk aversion and nonlabor income. The proxy for risk aversion is composed of variables such as marital status, number of children, and house value. A substantial portion of nonlabor income is composed of income of spouse, obviously related to marital status. The usefulness of these instruments is vitiated to the extent that they are correlated with unobserved human capital. As evidence that this is a legitimate concern, consider that the well-known "marriage premium" in earnings equations is widely interpreted as representing unobserved human capital (Becker 1981, 1985; Kenny 1983; Nakosteen and Zimmer 1987; Korenman and Neumark 1991).

<sup>2</sup> Recent works include Butler and Worrall (1983), Low and McPheters (1983), Topel (1984), Viscusi and Moore (1987), Hamermesh and Wolfe (1990), Hersch and Viscusi (1990), and Kostiuik (1990). Valuable summaries of the compensating wages literature are contained in Linnerooth (1979), Smith (1979), Brown (1980), Blomquist (1981), and Rosen (1986). All the empirical studies listed above are subject to the unobserved productivity bias discussed here. The problem of unobserved productivity bias is also relevant for cite-specific attributes, such as locational amenities. Blomquist, Berger, and Hoehn (1988) and Roback (1988) are recent examples of studies that use wage equations to estimate workers' marginal willingness to pay for locational amenities.

vides the key for determining the likely size of the bias when estimating compensating wage differentials with real data. Substituting a range of values for the three factors that, in our opinion, conservatively characterize actual labor market data sets, we find that the size of the bias is large—large enough to cause estimates to underestimate true compensating differentials by a factor of 50 percent or more, and even to result in wrong-signed coefficients. Section IV generalizes this analysis by allowing the hedonic wage constraint to assume a quadratic functional form. While the properties of the bias cannot be analytically derived, Monte Carlo analysis yields identical results: In particular, the size of the bias is the same when the three factors identified above are held constant. Section V demonstrates that the wage equations estimated from our simulated data “look” very much like regressions commonly reported in labor market studies.

On the basis of these findings, we conclude that contemporary labor market studies may severely underestimate workers’ marginal willingness to pay for job attributes. This has implications for a number of applications of compensating wage differentials, including value of life studies.

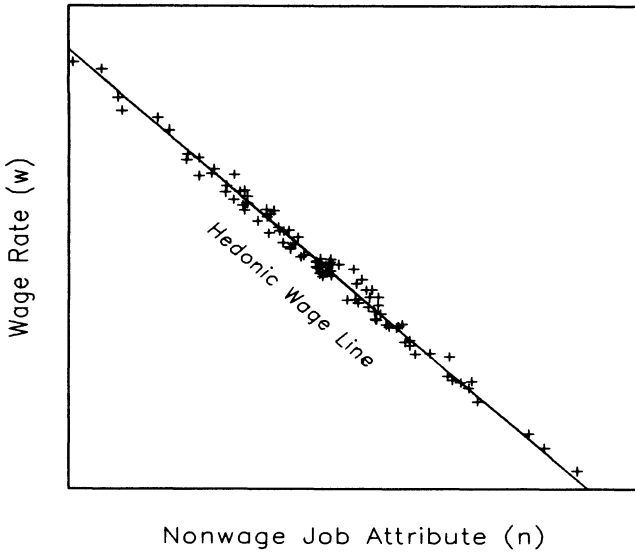
## II. The Nature of the Bias Due to Unobserved Productivity

The problem that arises when estimating compensating wage differentials in the presence of unobserved productivity is easily illustrated in figure 1. Here we have drawn the true hedonic wage line (corresponding to a given stock of human capital) for the case in which jobs consist of two dimensions, the wage rate ( $w$ ) and one (desirable) nonwage job attribute ( $n$ ). In figure 1*b* we have also included an expansion path, which identifies an average worker’s optimal  $w$ - $n$  combinations as his human capital varies. The expansion path slopes upward, indicating that as a worker’s human capital increases, he will choose jobs that are characterized by larger values of *both* the wage and the nonwage job amenity.<sup>3</sup>

Consider first a population of workers with differing tastes, identical stocks of observed human capital, and no unobserved productivity differences across workers. This population of workers will sort themselves along the given hedonic wage line. Accordingly, the econometrician, observing the  $w$ - $n$  choices of the workers, perhaps with some measurement error in the wage, is able to consistently estimate the

<sup>3</sup> While the “normality” assumption of nonwage job characteristics is widely accepted, the only empirical analysis of this subject concerns job safety. Viscusi (1978), Biddle and Zarkin (1988), and Garen (1988) all report that safety on the job is positively related to workers’ nonlabor income.

a



b

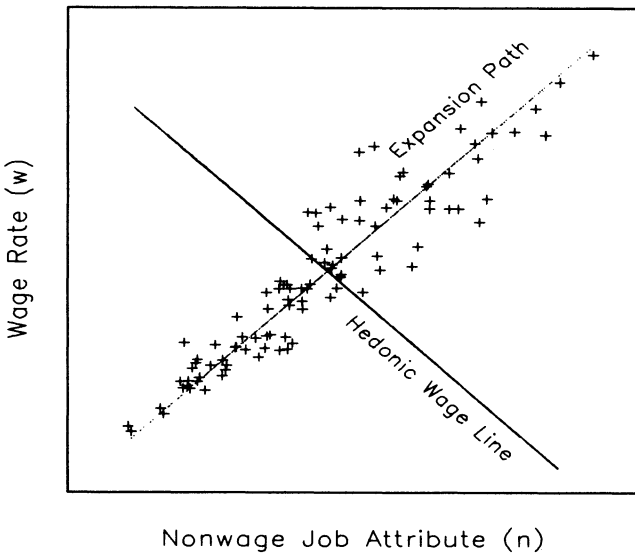


FIG. 1.—Effect of unobserved productivity heterogeneity on estimates of compensating wage differentials: *a*, no unobserved productivity heterogeneity; *b*, unobserved productivity heterogeneity.

wage trade-off—or compensating wage differential—associated with a given increase of the job attribute variable. This case is illustrated in figure 1a.

If we introduce productivity differences across workers, the observed  $w$ - $n$  combinations become more diffuse. More productive workers choose jobs further along their expansion paths than less productive workers. When these differences are very large, worker types having identical preferences will be located at far ends of their respective expansion paths. Sufficient productivity heterogeneity can cause sorting along the expansion path to dominate sorting along the hedonic wage lines. If these productivity differences are not controlled for in estimation, the  $w$ - $n$  scatter will lead to an estimated hedonic wage line that converges toward the expansion path. This situation is illustrated in figure 1b.

In both cases, the compensating wage differential that the econometrician wants to estimate is given by the slope of the true hedonic wage line, which is identical in both graphs. In the first case, the  $w$ - $n$  scatter allows him to correctly identify this differential. In the second case, the presence of unobserved productivity heterogeneity induces a positive slope to the  $w$ - $n$  scatter that causes the econometrician to incorrectly estimate the true hedonic wage line. The next two sections are concerned with determining the size of this bias.

### III. A Model of the Labor Market in Which Job Choices are Represented by a Linear Hedonic Wage Function

Let jobs be characterized by their hourly rate of compensation, consisting of a wage component ( $w$ ) and a single nonwage component ( $n$ ). The nonwage component may be thought of as representing some desirable job attribute, such as safe working conditions, flexible hours, or fringe benefits.<sup>4</sup>

Competition in the product and labor markets causes compensation packages to be characterized by trade-offs between these two components. Jobs that have low values of the nonwage amenity compensate by offering high wages. Jobs that have higher values of the amenity offer lower wages. Let  $p_n$  denote the hedonic wage, or compensating wage differential, associated with the nonwage job attribute. It represents the forgone wages associated with an additional unit of  $n$ . Finally, we assume that more productive workers are able to choose

<sup>4</sup> The case of an *undesirable* job attribute is discussed below. The extension to the multiattribute case is straightforward, provided that the respective taste parameters are distributed mutually independently.

among jobs that are characterized by larger rates of hourly total compensation; the  $i$ th worker's labor market productivity is denoted by  $K^i$ . These ideas are represented in the following linear hedonic wage function:

$$w = K^i - p_n n. \quad (1)$$

Equation (1) identifies the set of jobs, characterized by their wage and nonwage components, that are available to a worker with labor market productivity  $K^i$ .

In our analysis, we shall think of  $K$  as including the kinds of human capital variables usually included in empirical labor market studies—variables such as education, age, and labor market experience—as well as other variables that are difficult if not impossible for the econometrician to observe and measure—such as intelligence, perseverance, and the ability to work well with others. We assume that both the worker and prospective employers in the labor market are able to distinguish the worker's true labor market productivity. The econometrician, on the other hand, is assumed to measure only a portion of the worker's total productivity. This difference between total and observed productivity is critical for empirical estimates of  $p_n$ .

In addition to choosing a job, each worker must determine his optimal level of product market consumption ( $X$ ) and leisure ( $L$ ). Let the  $i$ th worker's preferences be represented by

$$U^i(X, L, n) = a_X^i \ln X + a_L^i \ln L + a_n^i \ln(T - L)n, \quad (2)$$

where  $T - L$  represents hours of work.<sup>5</sup> The worker is assumed to maximize utility subject to (1) and the usual financial budget constraint relating product market expenditures to labor market earnings,

$$p_X X = (T - L)w. \quad (3)$$

Given this framework, the resulting demand equations for the wage and nonwage job attributes are given by

$$w = \left( \frac{r}{1 + r} \right) K \quad (4a)$$

and

$$p_n n = \left( \frac{1}{1 + r} \right) K, \quad (4b)$$

<sup>5</sup> We assume without loss of generality that all the arguments in the utility function are positively valued by workers, i.e.,  $a_X$ ,  $a_L$ , and  $a_n > 0$ .

where  $r = a_x/a_n$  and the  $i$  superscript is suppressed, here and subsequently, for notational convenience.

Consider now a population of workers of differing tastes and productivities, each choosing his optimal job as described above. Suppose that an econometrician was able to exactly measure individual workers' job choices. Suppose further that he was interested in estimating the hedonic wage function given by (1) and, in particular, the value of the compensating wage differential,  $p_n$ . We now derive the bias that will result when important labor market productivity variables are omitted from the empirical analysis.

Let each worker's total labor market productivity be given by the sum of two variables,  $K_o$  and  $K_u$ , representing the worker's observed and unobserved (to the econometrician) labor market productivity. Let the regression equation be

$$w = \psi + \alpha K_o + \beta n + \epsilon, \tag{5}$$

where  $\epsilon = \alpha(K_u - \bar{K}_u)$ , and the true values of  $\psi$ ,  $\alpha$ , and  $\beta$  are (i)  $\psi = \alpha\bar{K}_u$ , (ii)  $\alpha = 1$ , and (iii)  $\beta = -p_n$ .<sup>6</sup> Let  $\hat{\beta}$  be the least-squares estimator of  $\beta$  in (5). Then the asymptotic bias of  $\hat{\beta}$  is given by

$$\text{bias} = \frac{\text{cov}(n, K_u)}{\text{var}(n) - \{[\text{cov}(n, K_o)]^2/\text{var}(K_o)\}}. \tag{6}$$

If we assume that the ratio of taste parameters,  $r = a_x/a_n$ , and market productivity parameters,  $K_o$  and  $K_u$ , are distributed mutually independently across the population of workers, then this bias term can be conveniently expressed in terms of three factors. This result will allow us to make some inferences concerning the size of the bias in actual labor market data.

**PROPOSITION.** Define  $\omega$  as the average share of total hourly remuneration taken in the form of wages,  $\omega = E[w/K]$ . Let  $\tau$  represent the proportion of wage dispersion due to the differing tastes of workers,  $\tau = E[\text{var}(w|K)]/\text{var}(w)$ . Finally, let  $\gamma$  identify the degree of unobserved productivity heterogeneity, defined by  $\gamma = \text{var}(K_u)/\text{var}(K)$  (note that  $0 \leq \omega, \tau, \gamma \leq 1$ ). Then

$$\text{bias} = -\beta \frac{\gamma(1 - \tau)(1 - \omega)}{\tau\omega^2 + \gamma(1 - \tau)(1 - \omega)^2}. \tag{7}$$

(A proof of this proposition is given in the Appendix.)<sup>7</sup>

<sup>6</sup> The subsequent analysis of the bias remains the same if  $K_o$  is replaced by  $\sum \rho_i X_i$  in eq. (5) and the  $\rho_i$ 's are estimated along with  $\beta$  in the regression equation, provided that  $K_u$  is uncorrelated with the  $X_i$ 's. In this case, the  $X_i$ 's and  $\rho_i$ 's may be thought of as representing the worker's human capital variables and their respective rates of return.

<sup>7</sup> The Appendix also demonstrates that the expression for the bias in eq. (6) applies when there is more than one attribute, if the taste parameters for the attributes have



Since  $\beta$  is negative, the estimate  $\hat{\beta}$  will be positively biased.<sup>8</sup> It is straightforward to demonstrate that the size of this bias is increasing in  $\gamma$  and decreasing in  $\tau$ . In terms of figure 1, the greater the extent of unobserved productivity heterogeneity (represented by increasing  $\gamma$ ), the more severe the sorting will be along the respective expansion paths, and the greater the associated bias. Further, the greater the dispersion along the respective hedonic wage lines due to differing tastes (represented by increasing  $\tau$ ), the smaller the biasing effect of a given degree of unobserved productivity heterogeneity. The effect of the third factor,  $\omega$ , the average wage share, is ambiguous.<sup>9</sup>

It will prove convenient to express equation (7) in terms of the ratio of estimated to true  $\beta$  values. Since  $\text{plim } \hat{\beta} = \beta + \text{bias}$ , rearrangement of (7) yields

$$\text{plim} \left( \frac{\hat{\beta}}{\beta} \right) = \frac{\tau\omega^2 - \gamma(1 - \tau)\omega(1 - \omega)}{\tau\omega^2 + \gamma(1 - \tau)(1 - \omega)^2}. \quad (8)$$

Three observations are in order concerning equation (8). First, when  $\gamma = 0$ , that is, when all the workers' productivity is observed,  $\text{plim } \hat{\beta} = \beta = -p_n$  (in terms of eq. [7], bias = 0). Second, when  $\gamma$  is sufficiently large,  $\hat{\beta}$  will take a "wrong" positive sign; sufficiently large is defined by  $\gamma > \tau\omega / [(1 - \tau)(1 - \omega)]$ . Noteworthy here is that this condition is independent of the absolute size of the compensating wage differential,  $p_n$ . And third, if the values of  $\omega$ ,  $\tau$ , and  $\gamma$  are known, then they can be substituted into equation (8) to obtain  $\beta$  from its estimated value,  $\hat{\beta}$ .

Equation (8) is especially valuable for our purposes because it allows one to calculate what the expected estimated value of  $\beta$  would be, given a data set characterized by particular values of  $\omega$ ,  $\tau$ , and  $\gamma$  and a given true value of  $\beta$ . In particular, it would be interesting to estimate the degree of bias that would likely result when using the kinds of large, national, cross-sectional data sets that are commonly em-

---

an identical distribution across workers. That is, *each* of the nonwage job attribute coefficients will be biased to the same degree as in the single nonwage job attribute case, provided that  $\omega$ ,  $\tau$ , and  $\gamma$  are the same.

<sup>8</sup> Suppose that the nonwage job attribute is a disamenity, defined by  $a_n < 0$ , and that the hedonic wage equation is given by  $w = K^i + p_n n$ , where  $p_n$  now represents the increase in wages for each unit of  $n$  the worker must consume on the job. Let  $\omega = E[w/K]$  as before, but now  $\omega > 1$  (the wage share of total compensation is greater than one because total compensation now includes the disutility from consuming the undesirable job attribute). Then eq. (7) remains the same except that  $1 - \omega$  is replaced by  $\omega - 1$ . Note that  $\omega - 1$  in the amenity case equals  $1 - \omega$  in the disamenity case, since both expressions equal  $p_n n / K$ . Also, since  $\beta$  is positive in the disamenity case, the estimate  $\hat{\beta}$  will now be negatively biased.

<sup>9</sup> Note, however, that in the range of parameter values used in the analysis below, decreases in  $\omega$  are unambiguously associated with an increase in the bias.

ployed in empirical studies of compensating wage differentials. This is the subject of the remainder of this section.

The first step in this line of inquiry requires settling on some reasonable values for the three determinants of the bias,  $\omega$ ,  $\tau$ , and  $\gamma$ . According to the Bureau of Labor Statistics, fringe benefits accounted for some 27.6 percent of total compensation in 1980 (U.S. Department of Labor 1980).<sup>10</sup> Furthermore, this does not include the value of nonpecuniary job characteristics, such as working environment, length of commute from home to work, and risk of injury. As a result, we shall look at two values of average wage shares ( $\omega$ ): 75 percent and 65 percent.<sup>11</sup>

Determining a representative set of values for the average share of total wage variance due to taste differences across workers ( $\tau$ ) is somewhat more subjective. We suspect that most labor researchers would argue that relatively little wage dispersion is due to differences in tastes. As a result, we look at the following three values for  $\tau$ : 10 percent, 20 percent, and 30 percent.

With regard to the degree of unobserved productivity ( $\gamma$ ), we report the effect of increasing unobserved productivity heterogeneity as it ranges from zero to 90 percent. Even so, it will be necessary to focus in on a range of values that might be expected to characterize actual labor market data. In this matter, we are guided by the fact that the  $R^2$ 's associated with wage equations estimated from national, cross-sectional data sets rarely rise above .50. Accordingly, we suggest that 30–50 percent of total worker productivity variance remains uncaptured by the usual set of labor market productivity variables—variables such as age, labor market experience, and formal schooling.

Figure 2 reports the results of substituting these values into equation (8). Figures 2*a* and 2*b* hold constant the average wage share ( $\omega$ ) at 65 and 75 percent, respectively. The three dotted lines in the figures each hold constant the given values of  $\tau$ . The vertical axis reports the ratio of the estimated to the true value of  $\beta$  ( $\text{plim}[\hat{\beta}/\beta]$ ), and the horizontal axis reports differing values for the percentage of unobserved productivity variance,  $\gamma$ . By moving along any dotted line in the figures, one can identify the bias associated with changes in unobserved productivity heterogeneity for predetermined values of  $\omega$  and  $\tau$ . The solid horizontal line in the middle identifies the point at which

<sup>10</sup> Fringe benefits are defined as including “pay for leave time,” “private pension plans,” “life, accident, and health insurance,” “government-required contributions to social security, workers’ compensation, and unemployment insurance.”

<sup>11</sup> We have also analyzed the bias when the average wage share equals 55 percent. As can be confirmed from eq. (8), the bias is generally larger than for the two cases reported below. We omit these results in the interest of brevity.

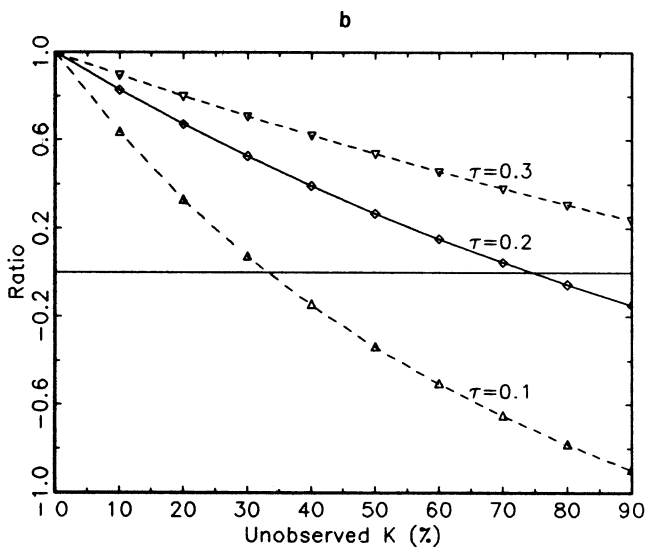
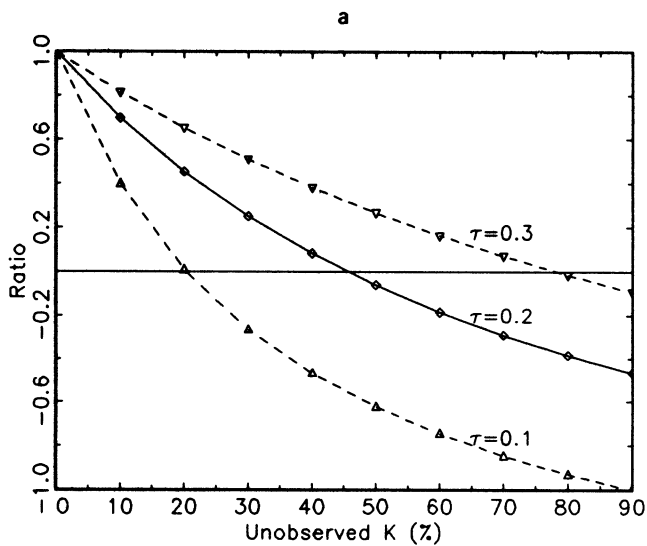


FIG. 2.—Ratio of estimate to true wage differential: linear hedonic wage function: *a*, wage share = 65 percent; *b*, wage share = 75 percent.

$\text{plim}(\hat{\beta}/\beta)$  is zero. Below this line,  $\hat{\beta}$  becomes positive, taking the opposite sign of  $\beta$ .

Consider first the case in which the percentage of unobserved productivity variance is equal to zero; that is,  $K_o = K$ . In this case, workers still differ in their stock of observed human capital, but unobserved differences across individuals are due solely to differences in tastes. As a result, the econometrician is able to accurately estimate the compensating wage differential associated with the nonwage job amenity. Both figures reflect this by displaying that, when  $\gamma = 0$ ,  $\text{plim} \hat{\beta} = \beta$  ( $\text{plim}[\hat{\beta}/\beta] = 1$ ), no matter what the value of  $\tau$ .<sup>12</sup> We next consider the effect of increasing the degree of unobserved productivity heterogeneity.

Suppose that wages average 65 percent of total hourly compensation for workers, and differences in tastes account for 20 percent of total wage variance. As the middle line in figure 2a reveals, an increase in the proportion of unobserved productivity variance from zero to 10 percent results in a decrease in the corresponding compensating wage estimate of approximately 30 percent (corresponding to a decrease in the value of  $\text{plim}[\hat{\beta}/\beta]$  from 1.00 to approximately .70). Moving further along the middle line in figure 2a, one sees that additional increases in unobserved productivity heterogeneity have a further substantial impact on estimates of the hedonic wage.

Consider now the case in which the unobserved productivity variance is between 30 and 50 percent of total productivity variance—values we suggest conservatively represent contemporary labor market data. Not only is the compensating wage differential severely underestimated, but it can even be wrong-signed (as indicated by the fact that the value for  $\text{plim}[\hat{\beta}/\beta]$  lies below the solid horizontal line).

To attach a monetary value to this result, suppose that 50 percent of total productivity heterogeneity is unobserved (let  $\omega = .65$  and  $\tau = .20$  as before) and that the job amenity is valued in the market for labor at \$1.00 a unit. That is, jobs with an additional unit of this attribute pay workers \$1.00 less per hour than they could earn elsewhere. Using micro data that contain information on workers' job choices and their personal characteristics, the econometrician would incorrectly estimate a wage *premium* for these jobs. According to figure 2a, this premium would amount to a little less than 5 cents an hour. Thus while the truth of the matter was that workers viewed the respective job attribute as a "good," the econometrician would

<sup>12</sup> There are other sources of bias, such as measurement error in the nonwage job amenity, that could bias the estimate of the compensating wage differential, even when there is no unobserved productivity heterogeneity. The present analysis omits consideration of these other factors.

incorrectly conclude from the sign of the estimated coefficient that workers judged this attribute to be a disamenity. Alternatively, if the coefficient was insignificant, he might incorrectly conclude that workers were indifferent toward this attribute.

Reviewing the range of estimates that correspond to  $.65 \leq \omega \leq .75$ ,  $.10 \leq \tau \leq .30$ , and  $.30 \leq \gamma \leq .50$ , we see that a large subset of these values results in estimates of compensating differentials that are wrong-signed. Among those that are right-signed, few approach even half of their true values.

As they stand, these results cast doubt concerning the validity of labor market studies that attempt to estimate compensating wage differentials. Even so, one wonders to what extent these results generalize for different hedonic wage and utility functions.

#### IV. A Model of the Labor Market in Which Job Choices Are Represented by a Quadratic Hedonic Wage Function

This section investigates the role of unobserved productivity heterogeneity for estimates of compensating wage differentials when workers' job choices are represented by a nonlinear hedonic wage function. When the hedonic wage function is linear, changes in a worker's productivity have the effect of inducing a pure "income" effect, since the price of the job amenity is held constant. In contrast, Rosen (1974) argues that "linearity is unlikely" if there are increasing marginal costs across firms in providing the job attribute. Accordingly, suppose that the implicit price of the job amenity increases as additional amounts of the attribute are demanded by the worker. One wonders if the bias associated with unobserved changes in the worker's productivity might be tempered by a corresponding substitution effect due to the attribute's changing price.

In a related vein, the combination of a linear hedonic wage function and a Cobb-Douglas utility function results in a linear expansion path, or constant income elasticity with respect to the demands for wage and nonwage compensation. Suppose that the expansion path was nonlinear. Would the same three factors identified in the preceding section still be sufficient to explain the bias?

Let the set of job choices available to a worker with productivity  $K^i$  be given by the quadratic hedonic wage function

$$w = K^i - \theta n - \delta n^2, \quad (9)$$

where  $\theta > 0$  and  $\delta > 0$ . The corresponding compensating wage differential associated with  $n$  is

$$p_n(n) = \theta + 2\delta n. \quad (10)$$

Note that the restrictions on  $\theta$  and  $\delta$  guarantee that the price of the attribute is positive and rises as more of it is consumed, behaviors consistent with increasing marginal costs of supplying the attribute.

If everything else related to workers' job choice decisions remains unchanged from the previous section, then the resultant wage and nonwage amenity demand equations are given by

$$w = \frac{2\theta^2r(1+r) - 2\theta r[\theta^2(1+r)^2 + 4\delta(1+2r)K]^{1/2} + 8\delta r(1+2r)K}{4\delta(1+2r)^2} \tag{11a}$$

and

$$n = \frac{-\theta(1+r) + [\theta^2(1+r)^2 + 4\delta(1+2r)K]^{1/2}}{2\delta(1+2r)}. \tag{11b}$$

As equations (11a) and (11b) demonstrate, one implication of the quadratic linear hedonic wage function is that workers' demand functions are no longer linear in  $K$ . The expansion path is now described by

$$w = r(\theta n + 2\delta n^2). \tag{12}$$

Equation (12) reveals that as workers' labor market productivity increases, so does the wage proportion of their total compensation. Increases in  $\theta$  and  $\delta$ , representing increases in the price of the nonwage amenity, have the effect of increasing the proportion of total compensation taken in the form of wages.

Once again, suppose that each worker's total labor market productivity and taste parameters are distributed as assumed above. Let the regression equation now be specified by

$$w = \psi + \alpha K_o + \beta_1 n + \beta_2 n^2 + \epsilon, \tag{13}$$

where  $\epsilon = \alpha(K_u - \bar{K}_u)$ , and the true values of  $\psi$ ,  $\alpha$ ,  $\beta_1$ , and  $\beta_2$  are (i)  $\psi = \alpha\bar{K}_u$ , (ii)  $\alpha = 1$ , (iii)  $\beta_1 = -\theta$ , and (iv)  $\beta_2 = -\delta$ . Let  $\tilde{n} = (n, n^2)$  and  $\tilde{\beta} = (\beta_1, \beta_2)'$ . Then the asymptotic bias of the least-squares estimate of  $\tilde{\beta}$  can be shown to be

$$\text{bias} = \left[ \text{cov}(\tilde{n}) - \frac{\text{cov}(\tilde{n}, K_o)\text{cov}(\tilde{n}, K_o)'}{\text{var}(K_o)} \right]^{-1} \text{cov}(\tilde{n}, K_o). \tag{14}$$

Ideally, we would like to solve for this bias as a function of model parameters, as in the previous case. Unfortunately, this is not possible since we are unable to derive the analytical distributions of  $w$  and  $n$ , given the distributions of  $r$  and  $K$ .

Our solution is to proceed with Monte Carlo analysis as a means of numerically evaluating the regression bias. In particular, we are interested in answering two questions. First, are the three determi-

nants of the regression bias in the linear hedonic wage case sufficient to explain that bias when the hedonic wage function is quadratic in the job amenity? Second, when these three factors are held constant, is the numerical size of the regression bias the same?

The Appendix describes a numerical routine that was used to identify model parameters that would generate predetermined values for  $\omega$ ,  $\tau$ , and  $\gamma$ . Unfortunately, the relationship is not unique, which leads to a potential problem. If different sets of model parameters lead to different biases, with  $\omega$ ,  $\tau$ , and  $\gamma$  held constant, what then can be learned from this exercise? Remarkably, this is not the case. After extensive Monte Carlo analysis, we have come to the conclusion that the size of the bias is "very close," if not identical, to the linear case in which the three factors identified above are held constant. We first illustrate what we mean by "very close" and then explain our uncertainty about whether these results are, in fact, identical to those of the preceding section.

Figure 3 reports the ratio of the estimated to the true value of the compensating wage differential, evaluated at the mean level of the nonwage job attribute,  $\bar{n}$ , for various model parameters, with  $\omega = .65$  and  $\tau = .20$  held constant: ratio =  $\hat{p}_n(\bar{n})/p_n(\bar{n}) = (-\hat{\beta}_1 - 2\hat{\beta}_2\bar{n})/(\theta + 2\delta\bar{n})$ . Eighteen different calculations of the bias were computed for each value of  $\gamma$ , corresponding to the 18 different combinations of the standard deviation of  $r$  (.05 or .10) and the values of  $\theta$  and  $\delta$  (1, 10, or 20).<sup>13</sup> Each calculation is represented in figure 3 as a tick mark plotted at the appropriate value of  $\gamma$ . The solid line represents the corresponding degree of bias in the linear hedonic wage case. As can be easily seen, changing model parameters results in little change in the bias associated with estimates of compensating wage differentials, provided that  $\omega$ ,  $\tau$ , and  $\gamma$  are held constant.

While figure 3 depicts only one set of outcomes, they are representative of a large number of Monte Carlo analyses. In every instance, comparisons of the bias in the quadratic hedonic wage case to those from the linear hedonic wage case resulted in similar results. On the basis of results such as those depicted in figure 3, we conclude that the bias associated with the quadratic hedonic wage case is "very close" to that of the linear hedonic wage case.

In fact, they may be identical. The reason for this uncertainty has to do with an approximation employed in the numerical routine used to calculate  $\tau$  for the Monte Carlo analysis. In essence, the difference in the two cases may simply be "approximation error."<sup>14</sup> Whatever

<sup>13</sup> We calculated 25 estimates for each combination of  $\sigma$ ,  $\theta$ ,  $\delta$ , and  $\gamma$  values. We then averaged these to obtain the associated estimate of  $\text{plim}[\hat{p}_n(\bar{n})/p_n(\bar{n})]$ .

<sup>14</sup> The problem arises in calculating  $\tau = 1 - \{\text{var}[E(w|K)]/\text{var}(w)\}$ . As described in the Appendix, the numerical routine calculates the value of the wage evaluated at the

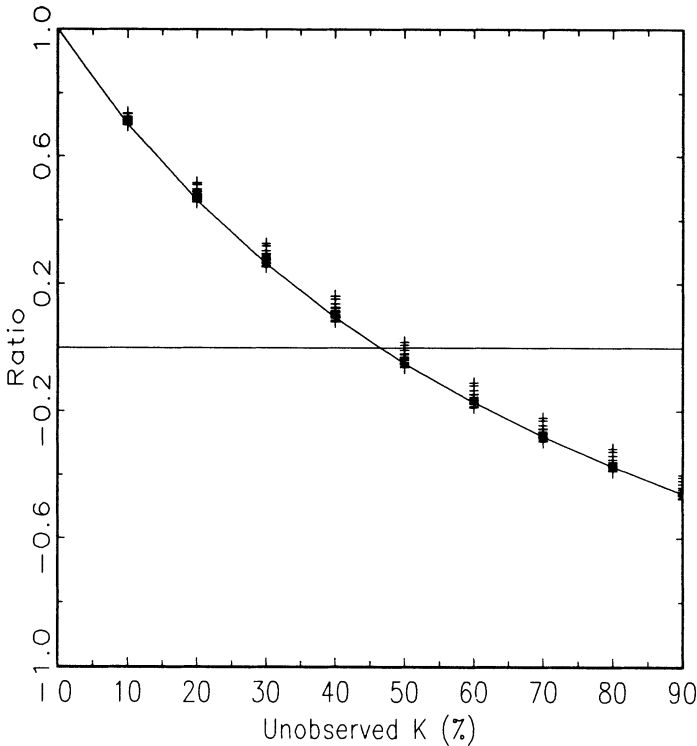


FIG. 3.—Ratio of estimate to true wage differential: nonlinear hedonic wage function.  $\tau = .2$ , wage share = 65 percent,  $\sigma_r = \{.05, .1\}$ ,  $\theta = \{1, 10, 20\}$ ,  $\delta = \{1, 10, 20\}$ .

the source of the difference, these results confirm the fact that the large biases associated with unobserved productivity heterogeneity in the linear hedonic wage case also hold when the hedonic wage function is quadratic. This raises the possibility that they may hold for other cases as well.

**V. But Does It Look like a Real Data Set?**

The previous two sections have demonstrated that, given either a linear or a quadratic hedonic wage constraint and worker preferences

---

mean taste parameter,  $\bar{r}$ , conditional on  $K$ ,  $w(\bar{r}|K)$ , and substitutes this value for  $E(w|K)$ . It then calculates the sample variance of  $w(\bar{r}|K)$  and uses this for  $\text{var}[E(w|K)]$ . In the linear case, this approximation yields bias values that are essentially identical to those calculated analytically using eq. (8). In the quadratic case, there will be a greater divergence because  $w$  is no longer a simple linear function of  $K$  (compare [11a] with [4a]). This approximation is adopted because the computational time associated with directly calculating  $\text{var}[E(w|K)]$  was found to be excessively large.



that are Cobb-Douglas, the bias in estimating compensating wage differentials due to unobserved productivity heterogeneity can be related to the three factors  $\omega$ ,  $\tau$ , and  $\gamma$ . Even so, one wonders what a data set generated by this underlying framework would look like. In particular, one wonders if this structure can generate the kind of estimated wage equations that are commonly reported in studies of compensating wage differentials.

In this section we attempt to replicate the regression results reported in a seminal study of compensating wage differentials undertaken by Thaler and Rosen (1976). We choose their study for a number of reasons. First, it attempts to estimate "the demand price for a person's own safety." Estimates of the value of saving a life are one of the most common and important policy-related applications of the theory of compensating wage differentials. Second, it estimates a linear hedonic wage function much like equation (1), even rejecting the hypothesis that the hedonic wage function is quadratic. Third, it is one of the most cited studies of its kind, in part because of the careful attention it devotes to employing only the highest-quality data about job risks.

The linear hedonic wage function estimated by Thaler and Rosen uses workers' weekly wage rate as its dependent variable. Independent variables include the human capital variables age and the square of age, as well as a formal education variable. The equation also includes one (unattractive) nonwage job attribute, risk of death. By observing the increase in wages with which a worker must be compensated to accept additional job risk, Thaler and Rosen are able to estimate how much workers would be willing to pay in order to reduce job-related deaths.

Column 1 in table 1 presents the coefficient values estimated by Thaler and Rosen.<sup>15</sup> These estimates suggest that an additional year of age is associated with a \$3.89 increase in weekly wages, minus approximately \$0.10 times the worker's age. An extra year of education is estimated to increase workers' weekly wages by \$3.40. Of particular interest for our purposes is the variable RISK, which measures the probability of an extra death per year at the worker's job, multiplied by  $10^5$ . The corresponding coefficient estimate is .0352. On an annualized basis, this implies that 1,000 people would together be willing to pay \$176,000 to have one fewer workplace fatality within

<sup>15</sup> Their equation also includes dummy variables for geographical region, race, occupation, full-time worker, and union status, as well as a variable measuring hours worked in previous week. The associated coefficient estimates are omitted from table 1 in the interest of brevity.

TABLE 1  
REGRESSION ESTIMATES OF THE VALUE OF SAVING A LIFE

VARIABLE	THALER-ROSEN (1)	SIMULATION I		SIMULATION II	
		Estimate (2)	True (3)	Estimate (4)	True (5)
A. Regression Results ( $N = 907$ )					
AGE	3.89 (.80)	3.72 (.64)	4.50	3.86 (.52)	4.50
(AGE) <sup>2</sup>	-.0479 (.0092)	-.0695 (.0075)	-.0851	-.0820 (.0061)	-.0965
EDUCATION	3.40 (.55)	3.39 (.52)	4.95	3.58 (.42)	4.87
RISK	.0352 (.0210)	.0352 (.0222)	.4027	.0352 (.0192)	.3020
R <sup>2</sup>	.41		.31		.55
B. Value of Saving a Life					
Estimated	\$176,000	\$176,000		\$176,000	
True	Unknown	\$2,013,500		\$1,510,000	
C. Unobserved Sample Characteristics					
$\omega$	Unknown	.750		.800	
$\tau$	Unknown	.198		.106	
$\gamma$	Unknown	.640		.395	

NOTE.—The dependent variable is the weekly wage rate. The Thaler and Rosen equation is taken from table 3, eq. 1 in Thaler and Rosen (1976). Also included in their regression equation are region, race, occupation, full-time, and union dummy variables, as well as a variable measuring hours worked in previous week (a constant term was presumably included but the corresponding coefficient is not reported). The equations reported in cols. 2 and 4 also include a constant term. Standard errors are in parentheses.  $\omega$ ,  $\tau$ , and  $\gamma$  are defined in the text.

their group.<sup>16</sup> This serves as the basis for Thaler and Rosen's conclusion that the market-implied value of saving a life is equal to \$176,000 (in 1967 dollars).

The Thaler and Rosen study also includes information about the means and standard deviations of the variables used in their analysis. We report below the summary sample statistics contained in their table 2:

<sup>16</sup> Thaler and Rosen give the following explanation: "Suppose 1,000 men are employed on a job entailing an extra death risk of 0.001 per year. Then, on average, one man out of the 1,000 will die during the year. The regression indicates that each man would be willing to work for \$176 per year less if the extra death probability were reduced from 0.001 to 0.0 [ $0.001 \times (\$0.0353 \times 10^5)$  per week  $\times$  50 weeks = \$176 per year]. Hence, they would together pay \$176,000 to eliminate that death: the value of the life saved must be \$176,000" (Thaler and Rosen 1976, p. 292).

Variable	Mean	Standard Deviation
AGE (years)	41.8	11.3
EDUCATION (years)	10.11	2.73
WEEKLY WAGE (dollars)	132.65	50.80
RISK (probability $\times 10^5$ )	109.8	67.6

Our objective is to determine whether there exist model parameters that (i) are able to generate a data set that conforms to the sample characteristics and regression results reported by Thaler and Rosen and (ii) are close to those values that we conjecture are most likely to represent actual labor market data.

For the sake of comparison, we modify the model of Section III by letting  $K_o = \rho_1 \text{AGE} + \rho_2 (\text{AGE})^2 + \rho_3 \text{EDUCATION}$ , where the respective  $\rho$ 's represent the market rates of return associated with each of the human capital variables. Accordingly, the regression equation of equation (5) becomes<sup>17</sup>

$$\begin{aligned} \text{WEEKLY WAGE} = & \text{CONSTANT} + \beta_1 \text{AGE} + \beta_2 (\text{AGE})^2 \\ & + \beta_3 \text{EDUCATION} + \beta_4 \text{RISK}. \end{aligned} \quad (15)$$

The first simulated sample that we shall discuss consists of 907 observations, the same as the Thaler-Rosen sample, and has essentially the same sample characteristics: The means of all the regression equation variables are identical to those reported above, as are the standard deviations for AGE and EDUCATION. The standard deviations of WEEKLY WAGE and RISK are very close at 50.84 and 66.9, respectively.

Turning now to the coefficient estimates reported in column 2, we see that the estimated coefficients for AGE and EDUCATION are very similar to their analogues in column 1. So are the associated standard errors. A comparison of these estimates with their "true" values (reported in col. 3) reveals that the respective coefficients are underestimated between 20 and 30 percent. This provides the first illustration in this context of the biasing effect of unobserved productivity heterogeneity.

Consider now the estimate of the RISK coefficient. The point estimate of .0352 is identical to the value reported by Thaler and Rosen, and the associated standard error is very close (.0222 vs. .0210). The

<sup>17</sup> Since the underlying model is based on a job amenity, the actual regression equations included the desirable job attribute,  $-1 \times (\text{probability of fatality}) \times 10^5$ . The risk coefficient reported in cols. 2 and 4 is the negative of the coefficient estimated in those regressions. Note 7 discusses how the concept of the wage share carries over to the disamenity case.

true coefficient value for our simulated sample is .4027. Thus while the RISK coefficient from the estimated wage equation implies that the market value of saving a life is \$176,000, the same figure reported by Thaler and Rosen, the real market value is \$2,013,500, over 11 times larger than the amount estimated.<sup>18</sup>

Panel C of table 1 reports the unobserved characteristics of the simulated data set. Note that the  $\omega$  and  $\tau$  values of .750 and .198 are both within the range of values that we conjecture are most likely to typify real labor market data. While the  $\gamma$  value of .640 is somewhat larger, it does result in an  $R^2$  value that is very plausible for studies of this kind.<sup>19</sup>

Columns 4 and 5 of table 1 present similar information from a second simulated sample. This sample of 907 observations is also characterized by means and standard deviations for the regression variables that are essentially the same as those reported by Thaler and Rosen.<sup>20</sup> A comparison of columns 4 and 1 reveals that both coefficient values and their associated standard errors closely resemble their analogues in the Thaler-Rosen sample. In particular, the estimated RISK coefficient is identical. This time, the true market value for saving a life is \$1,510,000, almost nine times larger than the estimated value of \$176,000. The biggest difference is that the  $R^2$  of the estimated regression equation is somewhat larger than the value associated with the first simulated sample. The reason is that the underlying degree of unobserved productivity heterogeneity,  $\gamma$ , is substantially smaller than it was before (.395 vs. .640).

In summary, the first simulated sample demonstrates that it is pos-

<sup>18</sup> One might find it surprising that the estimated rates of the return on the human capital variables are approximately 70–80 percent of their true values, but the estimated compensating wage differential is only 8–12 percent of its true value. In fact, this is predicted from the linear hedonic wage framework. Solving for  $\text{plim}(\hat{\alpha}/\alpha)$  in the same fashion as we solved for  $\text{plim}(\hat{\beta}/\beta)$  demonstrates that

$$0 \leq \text{plim}\left(\frac{\hat{\alpha}}{\alpha}\right) = \frac{\tau\omega^2}{\tau\omega^2 + \gamma(1 - \tau)(1 + \omega)(1 - \omega)} \leq 1.$$

Comparison with eq. (8) demonstrates that  $\text{plim}(\hat{\alpha}/\alpha) \geq \text{plim}(\hat{\beta}/\beta)$ .

<sup>19</sup> One objection to this analysis is that all the wage variation in our simulated samples comes from unobserved productivity and taste heterogeneity. No doubt some of the wage variation in the Thaler-Rosen study comes from non-productivity-related determinants of wages, such as regional differences, and left-hand-side measurement error. Accordingly, one could argue that the simulated samples should adjust for this by allowing smaller wage variation than reported by Thaler and Rosen. We explored this possibility and were able to replicate their regression results with model parameters that lay within the range of values identified in the text.

<sup>20</sup> As before, the means of all the variables are identical, as are the standard deviations for AGE and EDUCATION. The standard deviations of WEEKLY WAGE and RISK are 51.02 and 66.4, respectively, and differ only slightly from Thaler and Rosen's values of 50.80 and 67.6.

sible to generate a data set that looks like the data used by Thaler and Rosen, while suffering from a substantial degree of unobserved productivity bias. The second simulated sample demonstrates that there is more than one possible set of parameter values that will accomplish this. It also demonstrates that very different underlying model parameters (cf. panel C of table 1) can generate data that yield very similar estimated hedonic wage equations.

We emphasize that the simulated samples used in our replication analysis are observationally equivalent to the sample used by Thaler and Rosen, as measured by the means and standard deviations of the variables included in the regression equation. Further, they give rise to estimated wage equations that have exactly the same RISK coefficient that Thaler and Rosen estimate. The other coefficients and associated standard errors are also very close. Finally, both artificial data sets are characterized by (unobserved) values of average wage share, taste dispersion, and unobserved productivity heterogeneity that either fall within or are close to the range of values that we conjecture are likely to represent existing labor market data. As a result, we conclude that model parameters consistent with a severe degree of unobserved productivity bias *are* capable of generating observations that correspond to those employed in actual studies of compensating wage differentials.

## VI. Conclusion

The empirical study of compensating wage differentials has been applied to a host of subjects. Prominent among these have been attempts to measure the implicit valuation of job-specific characteristics—such as hazardous work conditions, layoff probabilities, flexible work schedules, pensions, vacations, and other fringe benefits—and location-specific characteristics—such as climate, crime, pollution, and crowding. Particularly interesting, and perhaps most important for public-policy purposes, have been attempts to derive market valuations of a human life from estimated compensating wage differentials for the risk of fatality.

If our results are representative of what one may expect in actual labor market data, then two important implications follow. First, point estimates reported in existing studies are likely to seriously underestimate the true compensating wage differentials they are intending to measure.<sup>21</sup> For example, given our optimizing framework,

<sup>21</sup> This raises the question, Why don't we see a large frequency of wrong signs in existing empirical studies of compensating wage differentials? In fact, we do. Brown (1980), in summarizing the empirical findings from other studies, writes that "the

if (i) wages constitute 65 percent of total compensation, (ii) differences in tastes account for 20 percent of total wage variation, and (iii) unobserved productivity variance equals 40 percent of total productivity variance, then workers' true valuations of life will be 10 times greater than valuations calculated from the estimated wage differentials (cf. fig. 2*a*). This means that job safety, in particular, and non-wage job attributes, in general, may be much more important to workers than previous studies have indicated.

Second, greater efforts to measure workers' labor market productivities are unlikely to reduce unobserved productivity until it is no longer a serious problem. We draw this conclusion from an inspection of figure 2: even 10 or 20 percent of unobserved productivity can cause large biases.

Where then should future research on the estimation of compensating wage differentials be directed? In our opinion, significant progress is most likely to come from new econometric methods that are better able to address the problem of unobserved productivity heterogeneity. We hope that this study will stimulate renewed efforts along this line.

**Appendix**

*A. Proof of Proposition*

Let  $K = K_o + K_u$ , where  $K_o$  and  $K_u$  are the observed and unobserved human capital, respectively. Assume that  $K_o$  and  $K_u$  are independently distributed across workers with means  $\bar{K}_o$  and  $\bar{K}_u$  and variances  $\sigma_o^2$  and  $\sigma_u^2$ , respectively. Let  $\gamma$  denote the proportion of unobserved human capital variance, so that  $\sigma_u^2 = \gamma\sigma_K^2$  and  $\bar{K}_u = \gamma\bar{K}$ . Finally, define  $\phi_w = r/(1 + r)$ ;  $\phi_w$  is assumed to be independent of the distribution of  $K_o$  and  $K_u$ , distributed with mean  $\bar{\phi}_w$  and variance  $\sigma_w^2$ .

Under the distributional assumptions about  $\phi_w$ ,  $K_o$ , and  $K_u$ , it is straightforward to verify that the equilibrium solutions in (4a) and (4b) imply

$$\text{var}(w) = E[\text{var}(w|K)] + \text{var}[E(w|K)] = \sigma_w^2(\bar{K}^2 + \sigma_K^2) + \sigma_K^2\bar{\phi}_w^2, \tag{A1}$$

$$\text{var}(n) = \frac{\sigma_w^2(\bar{K}^2 + \sigma_K^2) + \sigma_K^2(1 - \bar{\phi}_w)^2}{p_n^2}, \tag{A2}$$

$$\text{cov}(n, K_o) = \frac{(1 - \bar{\phi}_w)\sigma_o^2}{p_n}, \tag{A3}$$

$$\text{cov}(n, K_u) = \frac{(1 - \bar{\phi}_w)\sigma_u^2}{p_n}. \tag{A4}$$

---

overall pattern . . . is one of mixed results: some clear support for the theory but an uncomfortable number of exceptions" (p. 118). Further, studies that find insignificant and wrong-signed values of compensating wage differentials have a more difficult time getting published. This "publication selection" has been noted elsewhere (Tullock 1959; Denton 1985).

Substitution of these results into equation (6) in the text yields

$$\text{bias} = \frac{p_n(1 - \bar{\Phi}_w)\sigma_u^2}{\sigma_u^2(1 - \bar{\Phi}_w)^2 + \sigma_w^2(\bar{K}^2 + \sigma_K^2)} > 0. \tag{A5}$$

Note that (A1) implies that  $\sigma_w^2(\bar{K}^2 + \sigma_K^2) = [\tau/(1 - \tau)]\sigma_K^2\bar{\Phi}_w^2$ . Substituting this and  $\sigma_u^2 = \gamma\sigma_K^2$  into (A5), we obtain

$$\text{bias} = \frac{p_n(1 - \bar{\Phi}_w)\gamma(1 - \tau)}{\tau\bar{\Phi}_w^2 + \gamma(1 - \tau)(1 - \bar{\Phi}_w)^2} = -\beta \frac{\gamma(1 - \omega)(1 - \tau)}{\tau\omega^2 + \gamma(1 - \tau)(1 - \omega)^2}. \tag{A6}$$

This proves the proposition.

The expression for the bias in (A6) also applies when there is more than one job attribute, if the taste parameters for the attributes have an identical distribution across workers. To show this, let  $n_i$ ,  $a_i$ , and  $p_i$ ,  $i = 1, 2$ , be the  $i$ th job attribute, its taste parameter, and the corresponding hedonic price, respectively. Then, under the Cobb-Douglas utility function of equation (2), the equilibrium demand equations are given by  $w = \phi_w K$  and  $n_i = \phi_i K/p_i$ , where  $\phi_w = a_x/(a_x + \Sigma a_j)$  and  $\phi_i = a_i/(a_x + \Sigma a_j)$ ,  $i = 1, 2$ . Consider a linear regression equation

$$w = \psi + \alpha K_o + \beta_1 n_1 + \beta_2 n_2 + \epsilon = \psi + \alpha K_o + N\beta + \epsilon, \tag{A7}$$

where  $N = (n_1, n_2)$ ,  $\beta' = (\beta_1, \beta_2)$ , and  $\epsilon = \alpha(K_u - \bar{K}_u)$ . The asymptotic bias of the least-squares estimator  $\hat{\beta}$  is given by

$$\text{bias} = \{\text{cov}(N) - \text{cov}(N', K_o)[\text{var}(K_o)]^{-1}\text{cov}(K_o, N)\}^{-1}\text{cov}(N', \epsilon). \tag{A8}$$

In addition to the previous assumption of the statistical independence of the human capital variables and the taste parameters  $\phi_w$  and  $\phi_i$ , we assume that the  $\phi_i$ ,  $i = 1, 2$ , are distributed with the same mean  $\phi_n$ , the same variance  $\sigma_n^2$ , and covariance  $\sigma_{ij}$ . Under these assumptions, one can show that the variance of  $w$  is given by (A1) and that

$$\text{var}(n_i) = \frac{\sigma_n^2(\bar{K}^2 + \sigma_K^2) + \sigma_K^2\bar{\Phi}_n^2}{p_i^2}, \tag{A9}$$

$$\text{cov}(n_i, n_j) = \frac{\sigma_{ij}(\bar{K}^2 + \sigma_K^2) + \sigma_K^2\bar{\Phi}_n^2}{p_i p_j}, \tag{A10}$$

$$\text{cov}(n_i, K_o) = \frac{\bar{\Phi}_n \sigma_o^2}{p_i}, \tag{A11}$$

$$\text{cov}(n_i, K_u) = \frac{\bar{\Phi}_n \sigma_u^2}{p_i}. \tag{A12}$$

Substitution of these results into (A8) and rearrangement of the terms lead to the bias of  $\hat{\beta}_i$ :

$$\text{bias}(\hat{\beta}_i) = \frac{p_i \bar{\Phi}_n \sigma_u^2}{2\sigma_u^2 \bar{\Phi}_n^2 + (\sigma_n^2 + \sigma_{ij})(\bar{K}^2 + \sigma_K^2)}. \tag{A13}$$

Assumptions about the distributions of the variables indicate that  $\sigma_u^2 = \gamma\sigma_K^2$ ,  $\bar{\Phi}_w = 1 - 2\bar{\Phi}_n$ ,  $\bar{\Phi}_n = (1 - \bar{\Phi}_w)/2$ ,  $\sigma_w^2 = 2(\sigma_n^2 + \sigma_{ij})$ ,  $\omega = E(w/K) = \bar{\Phi}_w$ , and  $\sigma_w^2(\bar{K}^2 + \sigma_K^2) = [\tau/(1 - \tau)]\sigma_K^2\bar{\Phi}_w^2$ . Substituting these results into (A13) and simplifying terms, one can easily show that

$$\text{bias}(\hat{\beta}_i) = -\beta_i \frac{\gamma(1 - \omega)(1 - \tau)}{\tau\omega^2 + \gamma(1 - \tau)(1 - \omega)^2}. \tag{A14}$$

## B. Description of the Numerical Algorithm for Figure 3

## Step 1

Generate a random vector  $\mathbf{r}$  of size 50,000 with mean  $\bar{\tau}$  and variance  $\sigma_r^2$ , where  $\mathbf{r}$  represents the ratio of taste parameters,  $\alpha_x/a_n$ . The mean and variance of  $\phi_w = \mathbf{r}/(1 + \mathbf{r})$  are approximately

$$E(\phi_w) \approx \frac{\bar{\tau}}{1 + \bar{\tau}} - \frac{\sigma_r^2}{(1 + \bar{\tau})^3}$$

and

$$\text{var}(\phi_w) \approx \frac{\sigma_r^2}{(1 + \bar{\tau})^4}.$$

We further approximate  $E(\phi_w)$  by  $\bar{\tau}/(1 + \bar{\tau})$  and set  $\bar{\tau}$  such that  $\bar{\tau}/(1 + \bar{\tau}) = \omega$  for a given value of the wage share  $\omega$ ;  $\sigma_r^2$  is chosen to make the approximation of  $E(\phi_w)$  close (e.g.,  $\sigma_r^2 = .05$ ).

## Step 2

Let  $K = \sum_{i=1}^{10} K_i$ . For a predetermined common value  $\bar{K}$ , find through an iterative procedure the common standard deviation  $\sigma_i$  of  $K_i$  such that the sample value of  $\tau$  is the predetermined value .10, .20, or .30. To do this, for each  $\sigma_i$ , generate  $K_i$ ,  $i = 1, 2, \dots, 10$ , and corresponding  $K$ . Compute vectors  $\mathbf{w} = \phi_w K$  and  $\mathbf{n} = (1 - \phi_w)K/p_n$  and their sample means and variances. Compute  $\tau$  by  $1 - \{\text{var}[E(w|K)]/\text{var}(w)\}$ , where  $\text{var}[E(w|K)]$  is approximated by the sample variance of a vector  $[\bar{\tau}/(1 + \bar{\tau})]K$ , where  $\bar{\tau}$  is the sample mean of  $\mathbf{r}$ . Once  $\tau$  is computed for each  $\sigma_i$ , the procedure repeats by varying  $\sigma_i$  until the sample value of  $\tau$  is equal to the predetermined value of  $\tau$  within the interval .0001.

## Step 3

The samples of  $w$ ,  $n$ ,  $K_i$ , and  $K$  generated above satisfy the conditions that the average wage share is (approximately)  $\omega$  and the average degree of wage dispersion due to taste is  $\tau$ . For each choice of the proportion  $\gamma$  of unobserved  $K$ , the least-squares estimates of  $\beta_1$  and  $\beta_2$  in equation (13) in the text are computed. From these estimates, we compute  $(-\hat{\beta}_1 - 2\hat{\beta}_2\bar{\pi})/(\theta + 2\delta\bar{\pi})$  for the ratio of the estimate to the true wage differential.

## References

- Becker, Gary S. *A Treatise on the Family*. Cambridge, Mass.: Harvard Univ. Press, 1981.
- . "Human Capital, Effort, and the Sexual Division of Labor." *J. Labor Econ.* 3, no. 1, pt. 2 (January 1985): S33–S58.
- Biddle, Jeff E., and Zarkin, Gary A. "Worker Preferences and Market Compensation for Job Risk." *Rev. Econ. and Statis.* 70 (November 1988): 660–67.
- Blomquist, Glenn C. "The Value of Human Life: An Empirical Perspective." *Econ. Inquiry* 19 (January 1981): 157–64.
- Blomquist, Glenn C.; Berger, Mark C.; and Hoehn, John P. "New Estimates of Quality of Life in Urban Areas." *A.E.R.* 78 (March 1988): 89–107.
- Brown, Charles. "Equalizing Differences in the Labor Market." *Q.J.E.* 94 (February 1980): 113–34.
- Butler, Richard J., and Worrall, John D. "Workers' Compensation: Benefit



- and Injury Claims Rates in the Seventies." *Rev. Econ. and Statis.* 65 (November 1983): 580–89.
- Denton, Frank T. "Data Mining as an Industry." *Rev. Econ. and Statis.* 67 (February 1985): 124–27.
- Duncan, Greg J., and Holmlund, Bertil. "Was Adam Smith Right after All? Another Test of the Theory of Compensating Wage Differentials." *J. Labor Econ.* 1 (October 1983): 366–79.
- Garen, John E. "Compensating Wage Differentials and the Endogeneity of Job Riskiness." *Rev. Econ. and Statis.* 70 (February 1988): 9–16.
- Hamermesh, Daniel S., and Wolfe, John R. "Compensating Wage Differentials and the Duration of Wage Loss." *J. Labor Econ.* 8, no. 1, pt. 2 (January 1990): S175–S197.
- Hersch, Joni, and Viscusi, W. Kip. "Cigarette Smoking, Seatbelt Use, and Difference in Wage-Risk Tradeoffs." *J. Human Resources* 25 (Spring 1990): 202–27.
- Kenny, Lawrence W. "The Accumulation of Human Capital during Marriage by Males." *Econ. Inquiry* 21 (April 1983): 223–31.
- Korenman, S., and Neumark, D. "Does Marriage Really Make Men More Productive?" *J. Human Resources* 26 (Spring 1991): 282–307.
- Kostiuk, Peter F. "Compensating Differentials for Shift Work." *J.P.E.* 98, no. 5, pt. 1 (October 1990): 1054–75.
- Linnerooth, Joanne. "The Value of Human Life: A Review of the Models." *Econ. Inquiry* 17 (January 1979): 52–74.
- Low, Stuart A., and McPheters, Lee R. "Wage Differentials and Risk of Death: An Empirical Analysis." *Econ. Inquiry* 21 (April 1983): 271–80.
- Nakosteen, Robert A., and Zimmer, Michael A. "Marital Status and Earnings of Young Men: A Model with Endogenous Selection." *J. Human Resources* 22 (Spring 1987): 248–68.
- Roback, Jennifer. "Wages, Rents, and Amenities: Differences among Workers and Regions." *Econ. Inquiry* 26 (January 1988): 23–41.
- Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *J.P.E.* 82 (January/February 1974): 34–55.
- . "The Theory of Equalizing Differences." In *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland, 1986.
- Smith, Robert S. "Compensating Wage Differentials and Public Policy: A Review." *Indus. and Labor Relations Rev.* 32 (April 1979): 339–52.
- Thaler, Richard, and Rosen, Sherwin. "The Value of Saving a Life: Evidence from the Labor Market." In *Household Production and Consumption*, edited by Nestor E. Terleckyj. New York: Columbia Univ. Press (for NBER), 1976.
- Topel, Robert H. "Equilibrium Earnings, Turnover, and Unemployment: New Evidence." *J. Labor Econ.* 2 (October 1984): 500–522.
- Tullock, Gordon. "Publication Decisions and Tests of Significance: A Comment." *J. American Statis. Assoc.* 54 (September 1959): 593.
- U.S. Department of Labor. Bureau of Labor Statistics. *Handbook of Labor Statistics, 1980*. Bull. no. 2070. Washington: Government Printing Office, 1980.
- Viscusi, W. Kip. "Wealth Effects and Earnings Premiums for Job Hazards." *Rev. Econ. and Statis.* 60 (August 1978): 408–16.
- Viscusi, W. Kip, and Moore, Michael J. "Workers' Compensation: Wage Effects, Benefit Inadequacies, and the Value of Health Losses." *Rev. Econ. and Statis.* 69 (May 1987): 249–61.