A Discrete Choice Approach to Estimating Workers' Marginal Valuation of Fringe Benefits

Anne Beeson Royalty

Indiana University Purdue University Indianapolis
and
Stanford University

May 2000

This paper offers a new way of estimating workers' valuation of fringe benefits using data on workers' choices among fringe benefits packages offered by the employer. This approach overcomes both the omitted variable problem and the identification problem that bias estimates of compensating differentials and supply and demand parameters for fringe benefits from the traditional hedonic model most frequently used to analyze this problem. With this approach, the observed choice among offered fringe benefits packages which require different employee contributions and receive different employer subsidies conveys information about how much in wages workers are willing to give up to obtain additional firm dollars in the form of fringe benefits. That is the valuation of fringe benefits that we want to estimate. The comparison among alternatives implicit in the discrete choice method differences away fixed unobservable individual productivity differences that are believed to be the main problem in estimating compensating differentials and lessens the endogeneity problems that arise in estimating hedonic demand parameters. Variation in the offered wage-fringe price across firms identifies workers' valuation of fringe dollars, serving the same function as the sometimes arbitrary market boundaries that must be imposed in the hedonic model to achieve identification. Exploratory empirical results using grouped firm data on choices of alternative health insurance plans provide support for the proposed approach. Unlike most estimates of compensating differentials for fringe benefits, the estimates are of the correct sign. The results suggest that families value health benefits substantially more than singles and that that valuation of fringe benefits dollars is substantially less than one-for-one with wage dollars.

Many thanks to Tim Bresnahan and Julie Schaffner for helpful conversations, to Dan Hamermesh, Tom Nechyba, John Pencavel, and workshop participants at Brown, Yale, Duke, University of Maryland, BLS, West Virginia University, SMU, the New York Federal Reserve Board, IUPUI, UC Davis, UC San Diego, the Ninth Annual Health Economics Conference, and the Bay Area Labor Economists Conference for valuable comments, and to Laura Carstenson and Chun Zi Zhang for research assistance. This material is based upon work supported by the National Science Foundation under Grant No. 0096094. Funding from the American Compensation Association Emerging Scholar Program is also gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF) or the American Compensation Association. A previous version of this paper was circulated under the title "A Discrete Choice Estimator of Workers' Valuation of Fringe Benefits."

As nonwage compensation has become an increasingly important portion of workers' compensation — nonwage benefits now comprise as much as 40% of compensation by some estimates (U.S. Chamber of Commerce 1992) — fringe benefits have come to have a significant effect on workers' well-being. Numerous public policies affect the nonwage compensation or fringe benefits received by workers. Pension regulation, tax incentives for retirement savings or health insurance provision, proposals for mandated employer-provided health insurance, and minimum wage legislation, as well as many other policies, either are specifically designed to influence nonwage benefits or are expected to do so through indirect channels. Policy makers require credible estimates of the demand and supply of fringe benefits as they analyze the effect of these policies on total compensation, savings, insurance status, or employment. Unfortunately, researchers have identified some somewhat daunting obstacles to obtaining unbiased estimates of the compensating differentials or market wage-fringe tradeoff and the supply and demand parameters for fringe benefits from the traditional hedonic model frequently used to analyze this problem (Epple 1987, Kahn and Lang 1988, Hwang, Reed and Hubbard 1992).

Many researchers have tried to estimate market compensating differentials for positive and negative pecuniary and non-pecuniary nonwage job attributes, including fringe benefits, as the first step to estimating the underlying supply and demand parameters associated with such job attributes. Yet even that first step has proven quite difficult. The theoretical importance of the theory of compensating differentials and its prediction of a wage-fringe tradeoff is rarely debated[1] yet the history of empirical work on the subject might lead one to wonder whether its predictions

[1] Rosen (1986) says, "On the conceptual level [the theory of equalizing differences] can make legitimate claim to *the* fundamental (long-run) market equilibrium construct in labor economics."

could ever be verified or its parameters estimated with available data.[2] In addition to the difficult issues associated with the first stage compensating differentials estimations, the matching of workers and firms implied by the hedonic model raises an unusual and difficult-to-solve set of identification issues in estimation of the structural supply and demand equations for fringe benefits. These obstacles and the lack of confidence in empirical estimates that these obstacles engender suggest that an entirely different approach to this problem is needed.

In this paper, I take a new approach to the problem of estimating how much workers value fringe benefits. I use a discrete choice method with data on choices among fringe benefit packages by employees to estimate workers' valuation of fringe benefits. The choice among various fringe benefits options that carry different prices to the employee conveys information about how much workers are willing to pay in terms of foregone wages for additional benefits. The comparison among alternatives within firms implicit in the discrete choice method differences away the unobservable individual productivity fixed effects that are believed to be the main problem in estimating compensating differentials and lessens the endogeneity problem in estimating demand parameters. Empirical results are in accordance with theory, in contrast to many results in the hedonic literature, and indicate that after tax wages and benefits are strong substitutes for families.

In the next section, I describe in detail the problems of the hedonic model. Following that, I describe the discrete choice model used in this paper and discuss how and to what extent it

[2] Brown (1980) writes of "an uncomfortable number of exceptions" when considering empirical support for the theory and Duncan and Holmlund (1983) write of its "surprising resistance to empirical confirmation."

solves the problems associated with the hedonic model. Last, I describe the data analyzed and present empirical results and possible extensions of the model.

Traditional Hedonic Model

The supply and demand for job attributes that are bundled together in defined jobs has been frequently analyzed in the hedonic framework of Rosen (1974). Rosen develops formally the idea that the value of a bundled good such as a job derives from the value of its attributes. For example, a particular make of car is valued for its horsepower and its seating capacity or a particular job is valued for its wage, its nonwage compensation, and its working conditions. Compensating differentials are the implicit prices of the job attributes that arise from the underlying demand for and supply of those attributes. The model also allows for heterogeneity of preferences for and production costs of such attributes. This theory has been quite influential in framing the way economists think about differentiated products and jobs.

In the hedonic model, workers' preferences are defined over job attributes. Workers value both wages and benefits and choose the exact combination of wages and benefits that maximizes utility by setting the (implicit) marginal price of each fringe benefit equal to its marginal utility. Standard microeconomic theory is applied to the market for job or product attributes with the recognition that the price of these goods is implicit rather than explicit since these goods are bundled together in a product or a job.

The first step in a hedonic estimation is to estimate the implicit prices of these job attributes. The usual way of estimating the relationship between wages and fringe benefits is first to estimate an hedonic wage equation of the form:

$$w_{ig} = \beta' X_i + \varphi \cdot f_{ig} + u_{ig} \qquad \textbf{(1)}$$

where w_{ig} is the wage or log wage of worker i at job g, X_i is a vector of characteristics of the worker, f_{ig} is the level of fringe benefits associated with the worker and the job and u_{ig} is the error term associated with the wage of worker i at firm g.[3] The parameter φ captures the market wage-fringe tradeoff. The relationship can, of course, be allowed to be nonlinear and interactions of f_{ig} and worker characteristics can allow for differences in the wage-fringe market price for workers of different productivity levels. If X includes all relevant determinants of compensation, then estimation of equation (1) provides unbiased estimates of the market tradeoff of wages and fringe benefits. This market tradeoff represents to individuals the compensating wage differential available in the market for various levels of fringe benefits. Theory predicts a market tradeoff — a worker with low levels of fringes benefits will be more highly compensated with wages than an otherwise identical worker with high fringe benefits. Ideally we would consistently estimate this market wage-fringe tradeoff and then use these estimates to estimate firm supply and worker demand as suggested in Rosen (1974). This second stage raises difficulties (to be described in greater detail below (Brown and Rosen (1982), Epple (1987))) that encourage many researchers to be content with simply estimating the market compensating differentials (Rosen (1986)). Yet even this has proven exceptionally difficult.

Previous studies of the tradeoff between wages and fringes or wages and other job characteristics have found at best only partial support of the theory of compensating differentials. The usual explanation for not finding the predicted results is a positive correlation between fringe benefits (or other job characteristics) and productivity variables such as ability that are

---

[3] I ignore job attributes other than fringe benefits here for ease of exposition.

unobservable but positively associated with wages. That is, we expect that in equation (1) $u_{ig} = \mu_i + v_{ig}$ where v_{ig} is iid across individuals and firms and uncorrelated with X_i and f_{ig} but μ_i is correlated with fringe benefits. Since the omitted variable is positively correlated with the included explanatory variable, f, the estimated coefficient $\hat{\varphi}$ is biased upwards. Researchers have always assumed this problem to be at the root of the too frequent wrong-signed estimates obtained in studies of the wage-fringe tradeoff. Methods proposed to solve the omitted variable bias problem rely on panel data (Brown (1980), Duncan and Holmlund (1983)) or difficult-to-find instruments (Garen (1988), Biddle and Zarkin (1988)) and, in any case, have been only partially successful in finding the theoretically predicted empirical results. Recent work by Hwang, Reed and Hubbard (1992) establishes that under reasonable assumptions, traditional compensating differential estimates may be underestimated (in absolute value) by a factor of ten.

Even more important in terms of policy evaluation are estimates of the underlying utility and cost parameters that give rise to the demand and supply of fringe benefits and other job attributes, which in turn give rise to the compensating differentials observed in the labor market. We cannot evaluate public policies that affect compensation without reliable estimates of the demand and supply of fringe benefits. Even assuming we had consistent estimates of the compensating differentials for fringe benefits, difficult problems arise in estimating supply and demand parameters. The combination of attributes bundled into differentiated products and heterogeneous workers and firms reverses the standard result of anonymity wherein it does not matter who purchases from whom (or who works for whom). In the hedonic world, for example, workers with high preferences for fringes are matched with firms with a low cost of producing fringe benefits. Additionally, the bundling of attributes implies a lack of arbitrage amongst attributes, allowing attribute prices to depend on quantity. Therefore, price and quantity may be jointly determined even for price-taking individuals because the choice of where to locate on the hedonic market price function or compensating differentials function determines price and quantity simultaneously (Diamond and Smith (1985), Kaha and Lang (1988)). These results imply that the price and quantity in the supply and demand equations are jointly determined but that we cannot use instrumental variables techniques in the traditional way since worker and firm characteristics and unobservables will be correlated via the matching process implied by the model (Kahn and Lang (1988) and Epple (1987)).

The solution to this problem suggested by Kahn and Lang and used by many others that have followed (Montgomery et. al. (1992), Hersch and Viscusi (1990), for example) has been to assume the existence of different markets for workers. One might assume, for example, separate markets for workers by industry or region. If there exist separate markets, then the compensating differential or hedonic price function may differ across those markets due to exogenous differences in the distribution of worker and firm characteristics across markets. Market indicators are in this case valid instruments for price in the second stage when estimating the supply and demand equations.

Segmented markets are not always a reasonable assumption and even when such an assumption is seemingly valid, estimates from these models do not always produce sensible results.[4] Thus, the second stage of a typical hedonic model also runs into difficulties. I show below how the problems of both the first and second stage of hedonic estimation can be solved or

[4]See, for example, the wrong-signed price effect in the demand equation in Montgomery et al. (1992).

at least alleviated by transforming the problem to a discrete choice setting that implicitly controls for fixed differences across individuals and firms.

A Discrete Choice Model

With only a few exceptions, previous studies of the wage-fringe tradeoff or, more generally, supply and demand for job attributes have been done in this hedonic framework.[5] We can obtain the same valuation of fringes or willingness to pay estimates using discrete choice models although, as Mason and Quigley (1990) point out, these two strands of literature have often remained remarkably separate. For the most part, certain types of applications are traditionally done in the hedonic framework while others are done in the discrete choice framework. The literature estimating supply and demand for job attributes has stayed primarily in the hedonic world.

The closest discrete choice analog to the traditional hedonic wage equation would look at worker choices among alternative compensation packages offered to the worker across, or both

---

[5]Rosen (1986) uses a discrete choice model of job choice to illustrate the concept of and difficulties with compensating differentials models but with a different approach than that used here and he does not estimate the model presented. Woodbury (1983) identifies a demand system for fringe benefits based on variation in marginal tax rates across income levels. Woodbury and Hamermesh (1992) also use variation in benefits prices due to tax rates to estimate demand for benefits. The latter two papers circumvent the problems of the hedonic model and estimate demand parameters by assuming that the relative price of fringes and wages differs from unity only because of tax preferences for fringe benefits. They assume that by observing (or imputing) tax rates that we also observe the price of fringe benefits and they then use that price to estimate compensation share demand functions. Although this approach avoids some of the problems encountered in the hedonic model, it encounters problems of its own. For example, it is difficult to be confident that the estimated price effects identified only from marginal tax rates are not simply nonlinear income effects. Also, the assumption that the firm tradeoff between wages and fringes is 1-for-1 is troublesome if employees' tax rates are correlated with firm characteristics that affect the firm's wage-fringe offer curve as would be implied if workers choose jobs at least in part based on the benefits offered. Woodbury and Hamermesh address the former problem with a clever instrumental variables approach but they cannot avoid the latter problem and still maintain the assumption that allows them to "observe" a price of fringe benefits. Since I want to allow for more a general specification of the relative price of wages and fringe benefits than is possible with any method where relative wage-fringe prices must be observed, I will continue to make comparisons primarily with the hedonic model, which is also general in that regard.

within and across, firms, since the hedonic model implicitly assumes that workers are choosing among a continuous range of possible wage-fringe combinations. The closest discrete choice model would simply make that a discrete choice among such combinations. This approach is not viable because data on all compensation packages offered to a worker are simply not available.

In this paper I suggest an alternative type of data on fringe benefits and wages with which to estimate a discrete choice model of the wage-fringe tradeoff. Many employers offer employees a menu of fringe benefit choices of varying dollar values. Different choices may entail a different employee contribution. The employee contribution required for a particular choice can be thought of as the wage that an employee must forego to purchase that option from the menu of fringes. Different choices often also carry different employer contributions. The employer contribution for a particular choice can be thought of as the fringe benefit dollars associated with that choice. The observed choice between offered fringe benefits options conveys information about how much in wages workers are willing to give up to obtain additional firm dollars in the form of fringe benefits. That is the valuation of fringe benefits that we want to estimate.

To understand this idea more fully, let us first consider more formally the employee's choice among fringe benefit packages and then illustrate how that choice can be used to infer the wage-fringe tradeoff of employees. In recognition of the data available, consider only the choice among employer-provided health insurance plans. Suppose firm g offers its employees J_g possible health insurance packages. Each package is characterized by a price paid by the employee, p_j, and a vector q_j that fully represents the characteristics of plan j. The vector q_j includes characteristics of plan j that are observed by the econometrician as well as characteristics such as quality that are unobservable to the econometrician. So we have health plan j (or fringe benefits package j) defined as $f_j \equiv (q_j, p_j)$, and the set of choices available to individual i at firm g defined as $F_{ig} \equiv (f_1, \ldots, f_{J_g})$. The vector q includes all possible quality and quantity dimensions, including characteristics of the health plan, or more generally, fringe benefit, that may be valuable to the firm even if not to the employee. For example, q will include whether certain benefits are provided, the level of benefits included, and the perceived quality of those benefits given their level. Again, the vector q may contain both observed and unobserved measures of quality. Note that p_j is the portion of the premium of plan j paid by the employee. Often this will not be equal to the total premium since the employer will pay some portion of the premium. A worker i at firm g with attributes, a_i, will choose among health plans j the alternative that maximizes a random utility function

$$U_{igj} = U_i(q_j, m_i; a_i) + \varepsilon_{igj} = V_{ij} + \varepsilon_{igj} \qquad \textbf{(2)}$$

subject to $f_{ij} \varepsilon F_{ig}$ and $p_j + m_i = y_i$ where m is a composite outside good with price equal to one. Let the deterministic component of utility be represented as

$$V_{ij} = \gamma_p(y_i - p_j) + \delta' q_j \qquad \textbf{(3)}$$

The form of the probability of choosing a particular plan will depend on the assumption made about the distribution of the error term. The approach to the problem presented below and its relation to the hedonic model is general with respect to this distributional assumption but it will ease exposition to focus on one particular case. Therefore, assume that ε_{igj} follows the type I extreme value distribution.[6] The probability of a worker at firm g choosing plan j is:

[6]The independence assumption of the logit model and the embedded independence of irrelevant alternatives is, of course, a restrictive assumption. The logit specification is used at this point, however, just as an illustration. More general selection due to job choice than that that is allowed by this restrictive assumption is discussed in the section below on matching of workers to jobs.

$$\Pi_{gj} = \frac{e^{\gamma_p(y-p_j)+\delta' q_j}}{\sum_{k=1}^{J_G} e^{\gamma_p(y-p_k)+\delta' q_k}} = \frac{e^{-\gamma_p p_j + \delta' q_j}}{\sum_{k=1}^{J_g} e^{-\gamma_p p_k + \delta' q_k}} = \frac{e^{(-\gamma_p p_j + \delta' q_j) - (-\gamma_p p_m + \delta' q_m)}}{\sum_{k=1}^{J_G} e^{(-\gamma_p p_k + \delta' q_k) - (-\gamma_p p_m + \delta' q_m)}} \qquad \textbf{(4)}$$

where m is any alternative in J_g and the third equality is introduced simply to emphasize that the probabilities can always be expressed in terms of the differences in utilities of the various alternatives. I begin with the simplest possible model in order to introduce the way in which the choice among health plans can be used to estimate the worker's valuation of fringe benefit dollars. Individual fixed effects, heterogeneity of preferences for fringes across workers, and endogeneity of worker firm matches will be allowed below.[7]

The model of equation (4) is the standard conditional logit model. Note that δ/γ_p represents the employee's willingness to pay for health plan characteristics, or the employee's marginal valuation of various characteristics of health plans relative to dollars. In this paper, I am not interested in the value to employees of particular health insurance coverage provisions; I am interested in employee preferences for a dollar in fringe benefits as compared with a dollar in wages. Using this framework, we are able, however, to transform the equations to be estimated and obtain an estimate of the marginal value of fringe benefit dollars relative to wage dollars.

Next consider how the employee's choice among health plans can provide an estimate of the worker's relative valuation of wages and fringe benefits. The important point is that the total premium of health plan j, TP_j is the total price paid for this health plan, including both employee and employer portions, for the bundle of attributes, q. And, it is the dollars spent on the bundle

[7]Again, I could make alternative distributional assumptions. Berry (1994) discusses alternatives that would apply here and in the empirical work below.

of characteristics q that we want to value. This will allow us to value the actual dollars spent on health insurance benefits relative to wages.

TP_j is the market valuation in dollars of the characteristics, q, of plan j, but an employee may value one dollar of health insurance at either more or less than one dollar. In fact, economists would generally expect wage dollars to be valued more highly by employees than dollars spent by employers on in-kind benefits since the wage dollars offer the employee more choice. The wage dollars could be spent on the same package of in-kind benefits that the employer could have offered or any other amount of those benefits that offered the employee higher utility. Therefore, we might expect that firms would offer no such benefits.

Many factors affect employer offerings of benefits, however, and we expect these factors to induce a positive level of nonwage benefits. The first, of course, is the tax treatment of fringe benefits. Employer-provided fringe benefits are excluded from the employee's income and from the employer's payroll tax. The tax advantaged status of fringe benefits relative to wages, therefore, provides incentives for nonwage benefits. A second reason why employers offer benefits is that many benefits, in particular health insurance, can be purchased more cheaply through an employer group than individually. A third important reason is the incentives fringe benefits provide for inducing a given behavior by the employee that is valuable to the employer. For example, pension benefits that will not be received unless the employee stays at the firm for some minimum number of years are expected to reduce turnover and its associated costs to the employer. Alternatively, it may be that, rather than inducing a behavior change, such fringe benefits act as a screening mechanism, detering less desirable workers and attracting more desirable ones to the employer.

For all of these reasons, employers may offer fringe benefits as part of a compensation package and also for these reasons the wage-fringe tradeoff need not be equal to unity. An employer might be willing to offer more than one (present value) dollar of pension benefits in exchange for one dollar of wages if pension benefits reduce turnover costs. An employee might value that pension contribution at either more or less that its dollar cost depending on his or her demand for pensions relative to wages and the tax benefits of pensions relative to wages. Therefore estimation of the worker valuation of fringe benefits is not just an attempt to test or verify a theory predicting a one-for-one tradeoff (as is sometimes implied in the literature) but also to estimate the size of this tradeoff.

Since I want to estimate the value to the employee of the dollars spent by employers on wages relative to dollars spent on fringe benefits, I will estimate the effect on employees' health plan choices of employee price and of total dollars spent on those characteristics rather than estimating the effect of price and characteristics, q, on the employee's choice. The two prices, employee price and total price, are separately identified because some employers pay some or all of the premium and because some employers subsidize different alternatives by different amounts. That premium payment by the firm is, of course, the fringe benefit of interest and the price paid by the employee represents wages foregone.

To use workers' choices among health plans to estimate their valuation of dollars spent on health insurance benefits, I estimate a model where the probability of a worker at firm g choosing plan j is:

$$\Pi_j = \frac{e^{-\gamma_p p_j + \gamma_T TP_j}}{\sum_{k=1}^{J_g} e^{-\gamma_p p_K + \gamma_T TP_K}} \qquad \textbf{(5)}$$

where TP_j is the total premium of plan j and p_j is the employee's portion of the premium.[8] The ratio of the resulting estimates of γ_T and γ_p provides an estimate of workers' marginal valuation of fringe benefits relative to wage dollars.

To make this point more clear let us consider the transformation suggested more formally. In the model of health plan choice originally presented, $\delta' q$ represents the total utility value to the worker of all the attributes q. It follows that $\frac{\delta' q}{\gamma_p}$ is the total willingness to pay in dollars for the bundle of attributes q. What we want to know is how well TP, the total dollars spent on this bundle of attributes q, tracks workers' total willingness to pay for that bundle. For example, do workers value dollars spent on health insurance at more or less than their full market value? For workers who value benefits exactly at their market value, TP will equal $\frac{\delta' q}{\gamma_p}$. Suppose that all workers value benefits exactly at their market value. In this case, $\gamma_P TP = \delta' q$ where multiplying TP by γ_P converts dollars to the utility unit scale of $\delta' q$. Substituting TP in equation (4) for $\delta' q$ and estimating the model as implied by equation (5) will in this case produce an estimated coefficient on TP, $\hat{\gamma}_P$. In this case the coefficients on TP and P should not differ significantly; both are estimates of γ_P and we have considered the case where workers value benefit dollars and wage dollars equally.

[8]Note that with the assumptions about the form of the utility function made above, income (and other individual-specific variables) will fall out of the choice probabilities. I therefore omit these variables from the utility function from here on to simplify notation.

Now consider the general case where all workers do not value benefit dollars equally with wage dollars. We want to know how, on average, TP differs proportionately from $\frac{\delta' q}{\gamma_p}$ or, equivalently, how γ_PTP differs from $\delta' q$. By substituting TP for $\delta' q$ in equation (4) and estimating the model in equation (5), we can obtain an estimate of the value to workers of fringe benefit dollars. The estimated coefficient on TP, γ_T, from the transformed model of equation (5), will be an estimate of $\gamma_V \gamma_P$ where γ_V is a proportionality factor that represents the average proportion by which total dollars spent on benefit differs from the total willingness to pay for benefits. We can then obtain an estimate of γ_V, $\hat{\gamma}_V = \frac{\hat{\gamma}_T}{\hat{\gamma}_P}$ using the estimated coefficient on the employee-paid premium, $\hat{\gamma}_P$.

Notice that if we observed all elements of q we could estimate the original model of equation (4) and obtain an estimate of γ_V from a regression of $\frac{\hat{\delta}' q}{\hat{\gamma}_p}$ on TP. Since some characteristics q (such as quality) are not observed, this method cannot be implemented by simply estimating equation (4). Note that the omission of elements of q in this case is a problem even if the omitted attributes are uncorrelated with the included attributes and price since we need an estimate of the total valuation of the bundle derived from all of its attributes, both observable and unobservable.

Another way to understand the parameters to be estimated in the transformed model of equation (5) is simply to consider TP_j and p_j attributes of plan j and γ_T and γ_P the valuation of those attributes by workers. TP_j is dollars spent on insurance benefits but the valuation of those dollars by the worker is incorporated in γ_T[9]. That is, γ_T will depend on the value to the worker of

[9] Using TP_j directly in this equation is somewhat analogous to an indirect utility function. I simply include the dollars that it takes to purchase the bundle q_j in the estimation equation to obtain an estimate of the worker's valuation of total fringe benefit dollars relative to wage dollars. As Small and Rosen (1981) point out, the discrete choice model is traditionally set up as what they call a conditional indirect utility function because utility depends both on the price of an

the package of services that TP_j buys.[10] I estimate workers' valuation of health insurance fringe benefits using the estimated coefficients on the total premium and the employee-paid portion of the premium from this model. γ_T represents the value to the employee of increasing the dollars spent on a health insurance plan while γ_p represents the effect on the employee's utility of an increase in the wages foregone or cost of the plan. Therefore, estimates of γ_T/γ_p obtained from this model produce an estimator of the worker's marginal valuation of fringe benefits relative to wages.[11]

The parameters of interest are identified because both employee and employer contributions to health insurance premiums can differ across the plans offered to the employee within and across firms. For example, an employer may subsidize a more expensive plan more heavily but also require a larger employee contribution for that plan. The estimates obtained from the discrete choice model will reflect the extent to which a worker is willing to pay for the extra fringe benefit dollars contributed by the firm.

The model can be easily extended to account for differences in preferences for fringe benefits across observable groups by interacting total premium, TP, with demographic or other variables expected to influence tastes. Such interaction terms can overcome many deficiencies of the stripped down model presented above. For example, perhaps there are systematic differences in how workers at small and large firms value firm dollars spent on health benefits due to the fact

---

alternative and its characteristics. This formulation pushes that notion further. The employee's utility derives from the attributes q that TP will buy but in this case we want to value the dollars spent on health benefits.

[10]Note also that plan characteristics are not included in the equations in (1) since we want to value the total dollars spent on fringes, not total dollars given coverage levels.

[11]Of course, the more general model can allow for a nonlinear relationship.

that insurance loading fees are larger at small firms. Interactions of firm size with total premium will allow for differences in the valuation of benefits for workers across firms of differing sizes. Similar interactions can account for job- or firm-related differences in workers' valuation of fringe benefits. For example, suppose jobs that entail a high health risk induce workers in those jobs to value health benefits more. An interaction of total premium and some measure of job risk can account for such differences.[12] Additionally, nonlinearities can be incorporated by entering employee-paid premium and total premium as a quadratic or other more flexible form. These possibilities will be allowed for in the empirical work below. The methodology of Berry (1994) and Berry et. al. (1995) could also be extended to allow for unobservable differences across workers in their tastes for fringe benefits but this extension is left to future work.

Next, let us consider how and why the proposed discrete choice model has advantages over the traditional hedonic model, how the proposed model compares to more common discrete choice problems, how the hedonic identification problem manifests itself in the discrete choice model, and what situations are not addressed by the proposed discrete choice procedure.

Comparison with the Hedonic Model

Surprisingly few comparisons of hedonic and discrete choice models have been made in the literature. Cropper et. al. (1993) simulate the pricing of housing attributes in both frameworks and find that the two methods do equally well in pricing the marginal value of

[12] I discuss below when such situations also produce an endogeneity problem due to workers' choice of jobs.

attributes but that the discrete choice model is superior in valuing non-marginal changes. Mason and Quigley (1990) also use Monte Carlo simulations to compare these two methods. They find that estimates from the hedonic model are closer to the "truth" when the random errors are small but that the discrete choice model does better when errors are large. Both Cropper et. al. and Mason and Quigley simulate straightforward applications that do not include unobservables that are correlated with the explanatory variables, the most serious problem for the hedonic model, and both compare estimation procedures rather than the structural differences and similarities of the two approaches. Trajtenberg (1990) compares some of the assumptions implicit in the two models and emphasizes the importance of using the discrete choice model when choices are, in fact, discrete and the continuity assumption of the hedonic model cannot be empirically justified. Nechyba and Strauss (1998) compare studies of housing and community choice, noting hedonic versus discrete choice methods as one way such studies differ. These studies provide useful comparisons of these two familiar methodologies but do not address the most important issues that arise in the study of compensating differentials nor the reasons that different types of problems arise in estimating the two different models.

CONTROLLING FOR INDIVIDUAL AND FIRM FIXED EFFECTS

Due to the different nature of the problem, none of the above-cited studies points out one difference between hedonic and discrete choice estimation techniques, a difference that is a substantial advantage for the problem presented here. In the estimation of compensating wage differentials, the major problem arises from individual-level unobservable omitted variables such as unobservable productivity. In the hedonic model, cross-section comparisons across individuals will inevitably pick up these individual-level unobservables and correlation of such unobservable fixed effects with included variables will cause biased estimates. In the discrete choice model, each alternative is evaluated relative to the others and therefore purely individual-level (additive) unobservable fixed effects play no role. They are differenced out by the very nature of the discrete choice problem.

Suppose fixed unobservable characteristics of the individual or firm, μ_i or φ_g, affect the value of each alternative plan available to an employee such that

$$U_{igj} = \gamma_p (y_i - p_j) + \delta' q_j + \mu_i + \varphi_g + \varepsilon_{igj}. \quad \textbf{(6)}$$

For example μ_i might be large if individual i values being covered by any type of health insurance more than the average worker. Or φ_g might be negative due to a firm effect if a small business with unhealthy workers can negotiate only below-average quality at a given price. Note first that the potential individual-specific unobservables in the discrete fringe benefit choice problem are different from those in the traditional hedonic wage equation and are expected to be much less troublesome in general. Second, if the individual- and firm-specific fixed effects enter as in equation (6) above, they will present no problem at all since, in the comparison of one alternative with another, they affect both alternatives equally. Neither pure individual effects nor pure firm effects affect the choice probabilities; allowing explicitly for μ_i and φ_g, and writing the probability in terms of differences across alternatives, the probability of a worker at firm g choosing plan j is:

$$\Pi_{gj} = \frac{e^{(-\gamma_p p_j + \delta' q_j + \mu_i + \varphi_g) - (\gamma_p p_m + \delta' q_m + \mu_i + \varphi_g)}}{\sum_{i=1}^{J_g} e^{(-\gamma_p p_k + \delta' q_k + \mu_i + \varphi_g) - (-\gamma_p p_m + \delta' q_m + \mu_i + \varphi_g)}}. \quad \textbf{(7)}$$

By moving to the discrete choice setting and using data on multiple alternatives available to the worker rather than just the chosen alternative, I easily circumvent the major problem that previous work has had in successfully implementing even the first stage of estimating compensating differentials in the hedonic framework.

Several studies of compensating differentials have used a fixed effects model intended to accomplish the same purpose as is effected by the discrete choice model. Brown (1980) was the first to introduce this method in this context but his results were not substantially more in line with theory using the fixed effects model than without it. Duncan and Holmlund (1983) followed this approach with more success. Although the idea on which the fixed effect hedonic wage equation is based is in large part the same as the differencing across alternatives here, the data across which differences are taken are more informative and exhibit more variation in the model I present. In Brown's and in Duncan and Holmlund's models, wages are differenced for a particular worker observed in different years. Since within firm year-to-year changes in the composition of compensation are quite small, this method relies largely on workers who switch jobs between interviews. Since the job change is itself endogenous and job changers are unlikely to be representative of the entire population, in practice this method is less satisfactory than it is in theory.[13] The "fixed effects" discrete choice model that I propose, on the other hand, controls for within-worker and within-firm fixed effects at a given time. This is much more effective control, I believe, for the type of individual and firm fixed effects with which we are concerned.

[13] Woodbury and Hamermesh (1992) also use a firm fixed effects model to estimate demand for fringe benefits, identifying variation in benefits prices using differences in tax rates estimated for the median faculty member at each academic institution in their panel. Their approach meets with more success than the hedonic models described in the text but, as discussed in footnote 5, a new set of issues is raised by their assumption that the relative price of wages and fringe benefits depends only on tax rates.

Endogeneity of job choice can affect the estimates from the discrete choice model but I argue below that the endogeneity problem that arises in the discrete choice model due to job choice is less severe than that which arises in the hedonic model, again due to the "fixed effects" nature of the discrete choice method.

UNOBSERVABLE ATTRIBUTES AND BIASED PRICE EFFECTS IN THE DISCRETE CHOICE MODEL

The problem of individual-level unobservables that plagues attempts to estimate compensating differentials stands in contrast to the more common problem that troubles consumer demand studies. Consumer demand studies must often confront the problem of choice-level unobservables such as product quality (Berry (1994)). Unobservable product quality that is correlated with price will produce upwardly biased price coefficient estimates in discrete choice models of product choice. This is not a problem in the model presented here because, since I want to value the dollars spent on benefits, I am using the total plan premium, TP, in the utility of each alternative rather than the observed attributes of each plan. Total plan premium reflects the market price of both observable and unobservable components of q, the attributes of the plan. Indeed, that is the very problem that was just described. If only observable elements of q are included in the estimation and unobservable attributes are correlated with price as we expect, then the price coefficient will be biased. Since we want to value the dollars spent on both observable and unobservable attributes of the health benefit, we want TP to pick up the value of both unobservable as well as observable components of q. So what is often a problem is actually an advantage in this case. The estimated wage-fringe tradeoff takes account of the total value of the fringe benefit package to the employee.

One can imagine, however, alternative discrete choice models that might be used in the ways described above to solve some of the problems encountered in estimating a hedonic wage equation model that would face the problem of unobservable attributes. For example, if we did have data on all jobs in an individual's choice set and the job that was chosen, we might estimate a discrete choice model of job choice and estimate the tradeoff between wages and fringe benefits by allowing the utility of each job to depend on the offered wage, nonwage benefits, and other job attributes. In this case, the discrete choice model would suffer from the same problem as the consumer demand studies described above. If some job attributes or nonwage benefits are unobservable and correlated with wages or observed benefits, then the coefficients estimated in this way will be biased. Note, however, that the hedonic method suffers from this problem and the problem of individual-level unobservables. Even in the context where job-level unobservables are a potential problem, the discrete choice model has an advantage in terms of differencing out individual-level fixed effects.

Endogenous Prices and the Hedonic Identification Problem

In the second stage of a typical hedonic estimation procedure, the econometrician uses the implicit prices or compensating differentials estimated in the first stage in an estimation of supply and demand equations. Demand is determined by workers maximizing utility subject to the given market (implicit) price function. We want to know the effect of the (implicit) price on demand for fringe benefits but we do not actually observe differences across individuals in the prices that they face. Instead we use the wage-fringe benefits combinations observed in the market to estimate the wage-fringe tradeoff that workers face. In the simplest case, the only variation in the estimates of price would be from any nonlinearities in the compensating differentials function itself: individuals observed consuming different levels of fringe benefits pay different prices. Price is obviously determined simultaneously with quantity when the price function is nonlinear and we must instrument for price in order to obtain unbiased price effects.

The natural solution to this problem might seem to be to use supply variables as instruments but since workers and firms are matched in this equilibrium, supply variables associated with the chosen job are not valid instruments. The characteristics of matched workers and firms will be correlated in equilibrium so firm characteristics of the firm with which the worker is matched may be correlated with the worker's demand (Kahn and Lang (1988) and Epple (1987)). There is no obvious source of exogenous price variation with which to estimate price effects on demand since everything that belongs in the supply equation may be correlated with demand via the choice that workers make of where to locate themselves on the compensating differentials function.

In the proposed model, we observe the price and other attributes of each alternative available to the worker. We assume that the alternative chosen offers the highest utility. Utility to workers or consumers depends on the attributes of the product as well as its price but demand is defined over the bundle of attributes identified as a particular product, in this case health plans. Since we do not need implicit prices to estimate this demand system, we do not have to estimate prices as we must in the first stage of the hedonic approach. The price-quantity relationship can, of course, still be nonlinear but since this nonlinear relationship varies across individuals who face different price-quantity combinations, relative prices are not systematically related to quantities due to this nonlinearity as is the case in the hedonic model; therefore, we do not have the same endogeneity problem.

Having data on prices for alternatives other than that chosen and variation in the relative prices of those alternatives across individuals in different firms or different markets is essentially like observing more than one point on the hedonic price locus (market wage-fringe tradeoff) and observing variation in the market tradeoff faced by individuals. That variation allows identification of worker preferences.

Figure 1 in Appendix A may make this idea more clear. Figure 1 graphs the wage-fringe tradeoff available at three different firms in the sample I will use to estimate the proposed model. The y-axis is the negative of the worker's portion of the premium or the wage gain from choosing a less expensive plan. The x-axis is the firm-paid portion of the health plan premium or the fringe benefit dollars associated with each plan. Utility to the worker increases in the NE direction. The points representing the set of health plans offered by a particular firm are connected by a line. Three firms are represented in this graph. As is clear from this figure, there is variation in the relative price of wages and fringe benefits across firms. It is that variation that will be used to identify worker preferences for wages versus health benefits.

The problem in the hedonic model is finding enough exogenous variation in the data to identify separately the implicit prices and the effect of price on demand. The most common solution to this problem is that suggested by Kahn and Lang (1988) wherein they instrument price with market indicator variables. Differences across markets imply that the entire market wage-fringe tradeoff function (or set of compensating differentials available at different levels of wages and fringe benefits) shifts with something exogenous to demand. These shifts are used to identify the effect of price on demand. In the model presented here, we observe price variation across individuals and use that to identify price effects. There is no need to estimate implicit prices nor to make what can sometimes be arbitrary market boundaries in order to assure that there is some source of identifying price variation in the model. I reformulate the problem in a way such that data are available on within-worker, within-firm choices and where those within-worker within-firm price differences vary across workers and firms. The hedonic and discrete choice solutions are really very similar, but the additional information on nonchosen alternatives that is used in the discrete choice model and the variation across workers or consumers in the attributes of the alternatives makes it easier to overcome the obstacles to consistent estimates in the discrete choice model.

Additionally, if there are other reasons that price is endogenous, as there may well be and as will be discussed below, in this model we can instrument for the price of each of the alternatives faced by the individual. What is not valid in the hedonic model is using instruments in an equation for demand of a particular product that correspond only to the supply cost of the chosen alternative. Instrumenting price for each of the available alternatives (see, for example, Berry et. al. 1995) essentially amounts to the hedonic equivalent of instrumenting the entire market wage-fringe tradeoff function rather than only the chosen point on that function. Again the solution is more practical in the discrete choice model.

THE MATCHING EQUILIBRIUM OF THE HEDONIC MODEL AND THE STRUCTURE OF THE PROPOSED MODEL

The central issue that remains is whether a worker's choice among jobs and the implied matching of workers to firms, the problem so clearly highlighted by the hedonic framework, induces a correlation between the explanatory variables and the error terms in the individual's utility function in the proposed model. In the previous section, I argued that the use of data on nonchosen alternatives and variation in the attributes of the alternatives across workers provides information with which to identify workers' valuation of fringe benefits in the discrete choice model I propose. Observed prices are not systematically related to quantities as they are when implicit prices must be estimated. Variation in within-firm relative prices is observed across workers. The question that remains is whether that observed variation is correlated with the errors in the discrete choice model.

Let us again begin by making the comparison with the hedonic demand equation where the level of benefits chosen by individual i employed at firm g, f_{ig} depends on demand variables Z_i and the implicit prices estimated in the first stage, $\hat{p}_{ig}$:

$$f_{ig} = \gamma' Z_i + \alpha \cdot \hat{p}_{ig} + \tau_{ig}. \qquad \textbf{(8)}$$

The problem that I now isolate is that the estimated implicit price, $\hat{p}$, faced by an individual is correlated with the error τ_{ig} since workers with, say, high unobservable levels of demand for fringe benefits can choose jobs with low prices for fringe benefits. The problem arises when workers differ in their demand for benefits and choose jobs on that basis, thereby creating a correlation between the explanatory variables and the unobservables in this equation.

Now I consider how that problem may manifest itself in the proposed model. As before, we must remember that workers are comparing amongst the alternatives in the choice set so the issue is not whether, for example, TP_1 is correlated with ε_1 across firms but whether TP_1 - TP_2 is correlated with ε_1 - ε_2. To know whether this is a problem, let us consider the matching process described by the hedonic model. Workers and firms may be matched based in some part on the fringe benefit package offered and demanded. If firm k is a "high fringe" firm, a worker with a high demand for fringes may be more likely to accept a job with firm k. However, if the firm heterogeneity in fringes across which workers choose jobs based on demand for fringes is fixed across all plans of the firm then those unobservable firm fixed effects will difference out and we will not expect TP_1 - TP_2 to be correlated with ε_1- ε_2. In other words, selection on levels of fringe benefits across firms (for example, φ_g in equation (7) or average TP for the firm), does not produce biased coefficient estimates in this model. Differencing across alternatives rids us of the first order part of the endogeneity problem that is due to job matching on benefit levels. The differencing of the discrete choice model will circumvent the job matching problem whenever the choice among jobs is made on the basis of something that is fixed across the menu of fringe benefit choices, such as job-level generosity of benefits.

The problem is not solved by this differencing, however, if matching of workers to firms also occurs on the price differences of plans within firms or any factor that is not fixed across the menu of fringe benefit choices. That is, we need to know, do workers choose firms based on the difference in TP_1 and TP_2 or p_1 and p_2 at a particular firm? In other words, is there selection across jobs on benefit price differences within firms? Some such selection may be important but it is more plausible that a more significant component of workers' choice of firm is the levels of benefits offered in that firm and the relative prices across firms. This probably more likely source of endogeneity, firm heterogeneity in price levels, is differenced away in the discrete choice model.

Even if job selection on the basis of within firm benefit price differences is relatively unimportant, there may still be occasions where the endogeneity of job choice creates a problem in obtaining consistent estimates from the discrete choice model. This problem is most likely to occur when we believe there are job- or firm-related differences in workers' valuation of benefits. As discussed above, we might want to allow for the possibility that workers' valuation of benefits varies with the job. For example, in jobs that entail a large health risk, workers may value health benefits more highly than they would in jobs with a lower health risk. We run into a problem, however, if the choice of such jobs is correlated with workers' individual-level deviations from average fringe valuations. That is, if, for example, workers with higher than average valuations of fringe benefits tend to choose jobs with low health risks, then we cannot consistently estimate the effect of job health risk on workers' valuations. In this case, we would want to use a job risk * (TP_1 - TP_2) interaction to capture differences in wage-fringe preferences by job risk. But, by assumption, the job risk variable is correlated with unobservable individual differences in the valuation of benefits via the worker's choice among jobs. The problem in this case relative to the benefit levels problem discussed above is that the variable that is correlated with job choice (job risk * (TP_1 - $TP_{2)}$) and thereby individual heterogeneity in valuation of fringe benefits is not fixed across alternatives (since it enters as an interaction with total premium differences) and therefore is not differenced away in the comparison across alternative fringe benefit choices. The model as it stands cannot therefore allow for a situation where workers' valuation of benefits vary with firm characteristics, if those firm characteristics enter into workers' job choice decisions <u>and</u> that dimension of the job choice is correlated with individual heterogeneity in the valuation of benefits. Note, however, that the discrete choice approach again has an advantage over the hedonic model to the extent that job choices are made on job characteristics that are fixed across fringe benefit alternatives within firms.

If it is thought that endogeneity of price differences or interactions with price differences are a serious concern, there are potential solutions to this problem, although estimation of such models is left to future work. The methodology of Berry (1994) could be adapted to deal with the endogeneity problem that could be induced by the matching of workers and firms. Berry suggests a method whereby instrumental variables can be used in a discrete choice model based on market share data. It is possible in this model to instrument for the price of each of the alternatives faced by the individual, not simply the alternative actually chosen, as is necessary in the hedonic model. The endogeneity problem here and thus the instruments are different in Berry's problem but a similar method can be used here. The same process of converting the problem to a linear share equation is already used below. The problem then becomes simply finding appropriate instruments given the correlation of worker and firm characteristics induced by the worker-firm matching process.[14]

Data

Data on individual choices of health insurance plans for workers in different firms is uncommon. The National Medical Care Expenditure Survey (NMCES) of 1977 and the National Medical Expenditure Survey (NMES) of 1987 both asked respondents and their employers about

---

[14]One possible instrument for total price differences is an indicator for whether the plan is self-insured. Uninteracted workforce variables representing the diversity of the work force may also be suitable for instruments for both total and worker premium differences across plans. Variety in workforce composition gives incentives for employers to offer different types of plans with different payment structures; therefore, we expect that workforce composition may be correlated with premium differences. Although the workforce variables interacted with total premium may affect aggregate market demand as argued below, the uninteracted workforce variables would 1) fall out of the individual level equations since they do not vary across plans and 2) would not be expected to enter individual level utility anyway; therefore, we expect these variables to be uncorrelated with the unobservables affecting demand for each plan. It is more difficult to think of potential instruments for interactions with premium differences although some of those just suggested might be appropriate. In the case where the model includes interactions with premium differences, a more promising strategy, were data available, might be an adaptation of the Abowd et. al. (1994) approach where worker, firm, and worker-firm effects are identified with panel data on workers and firms.

their health insurance options and the employee and employer contributions for each available alternative. These two datasets are the best available data on such choices currently available at the individual level. More recent data are preferable, however, in light of the very rapid changes in the health care market in recent years, particularly with respect to employer offerings to employees of choice among health plans. It is only very recently that a significant fraction of workers were offered the choice among two or more health insurance plans. The successor to NMCES and NMES, the Medical Expenditure Panel, will be available next year with current individual data of the sort needed to estimate this model.[15] More recent data are available from the Robert Wood Johnson Employer Health Insurance Survey (RWJ-EHIS) of 1993. These data are aggregated to the firm level and include detailed information on each health insurance plan offered by the firm, including the share of families and individuals enrolled in each plan and the employee and employer share of the premium for each plan. Following Berkson (1953) and Theil (1969), the logit model described above can be aggregated and these share data can be used to estimate the model. This aggregation and the resulting econometric model is described in the following section. As will be discussed below, more detailed individual data has some important advantages over aggregate firm-level data; however, the currently available RWJ-EHIS data allow testing of the proposed model and provide preliminary empirical estimates of its parameters.

The RWJ-EHIS surveyed approximately 2000 private employers and from 46 to 262

[15]Employers are also now more frequently offering a choice among fringe benefits options more generally. In the future, it should also be possible to estimate this model with data on choices among a menu of benefit choice encompassing pensions, other insurance, and time off, rather than having to rely only on health plan choices.

public employers in each of 10 states.[16] The data include information on 22,465 health insurance plans in 22,890 firms in these 10 states. The total number of firms includes those firms that offered no health insurance plans. Employers were asked about the total premium for each health plan for families and for singles and also the employer-paid portion of the premium for both groups. Firms were also asked to identify the share of enrollees by plan and the percentage of enrollees for each plan who were families, allowing one to identify the share of family enrollment in each plan and the share of single enrollment in each plan. Employers were also asked about their workforce. For example, data are available on the proportion of the firm's workforce that is in various low wage categories and breakdowns of the workforce by age and gender. Firm level data on firm size, industry, and number of years in business is also available.

The analysis will use only data on firms that offer two or more health insurance plans since only workers at such firms have a choice among plans.[17] I also eliminate plans with missing values for total or employee share of premium and plans with no share of the firm's market. After these selection criteria were imposed, I was left with 6827 single coverage observations in 2253 firms, 3865 of which reported that the employee-paid portion of the premium does not vary across workers within the firm. These selection criteria produced a family dataset of 6810 family coverage observations in 2532 firms, 3758 of which have

---

[16]The 10 states in which firms were surveyed are Colorado, Florida, Minnesota, New Mexico, New York, North Dakota, Oklahoma, Oregon, Vermont and Washington. Approximately 500 firms in four size categories (2-4, 5-9, 10-24, $\geq$25 employees) were sampled in each state. In a few cases, firms in particular geographic areas or SIC codes were oversampled to improve precision.

[17]The remaining firms may no longer be a representative sample of firms. I have investigated the effect of this selection criterion by comparing means of the data analyzed to the full sample. Most notably the full sample includes more smaller firms. There is little to be done about this problem since choices are made only by the subpopulation of workers who are offered more than one plan.

employee premiums that do not vary within the firm.[18] Means of the data are summarized in Table 1.

Table 1 Means		
Health plans per firm	3.02	
Firm size	857	
	Single Coverage	Family Coverage
Total premium*	151.36	396.65
Employee-paid portion of premium*	27.63	135.50
Median difference between highest and lowest total premium*	39.00	99.00
Median difference between highest and lowest employee-paid portion of premium*	14.46	66.01
*Premiums are monthly, in 1993 dollars.		

[18] Of the plans eliminated from the analysis, 75% in both datasets were eliminated because the firm offered only one health insurance plan.

Econometric Model

The logit model described above can be aggregated to the firm level in order to use the firm-level health plan share data available in the RWJ-EHIS survey. It should be emphasized again that the conceptual issues outlined above and the comparisons made between the hedonic and discrete choice approaches to the problem of estimating workers' valuation of fringe benefits are general and not specific to the logit model used as an illustration. The logit model is used, however, in the empirical work presented here because the available data are aggregated to the firm level, necessitating grouped data methods, of which the logit model is the most convenient. Also it should be noted that alternative distributions are not less restrictive than logit unless the independence assumption is relaxed. That extension is left to future work using individual level data when it becomes available.

The linear logit model estimated here is based on the fact that the probabilities represented by equation (5) can be combined and written as:

$$\ln \frac{\Pi_{jg}}{\Pi_{Jg}} = -\gamma_p(p_{jg} - p_{Jg}) + \gamma_T(TP_{jg} - TP_{Jg}). \quad \textbf{(9)}$$

Following Berkson (1953) and Theil (1969)[19], I can form groups of individuals facing the same $(p_j - p_J)$ and $(TP_j - TP_J)$ and estimate Π_j and Π_J based on the observed proportion of individuals in that group choosing each alternative.

For the problem at hand, each firm g will be a group facing the same prices and total

[19]See Ben-Akiva and Lerman (1985) for a clear exposition of this type of model.

premiums. The unit of observation is then the group of employees at each firm.[20] Let N_{jg} be the proportion of workers at firm g choosing plan j. The model to be estimated is simply a linear regression:

$$\ln \frac{N_{jg}}{N_{Jg}} = -\gamma_p (p_{jg} - p_{Jg}) + \gamma_T (TP_{jg} - TP_{Jg}) + \theta_{jg} \qquad \textbf{(10)}$$

The model can be estimated as an OLS regression where the error θ comes from having to estimate the probability of the choice of plan j for each group. The estimated parameters will be consistent as each group size gets large.

The model can be extended to include the effect of demographic variables when data on the composition of each group with respect to those demographic variables is available. Suppose, for example, we believe that the value of fringe benefits differs for high- and low-wage workers. In the individual-level model, one can allow for that by interacting total premium, TP_j, with an indicator for low-wage status:

$$V_{ij} = -\gamma_p p_j + \gamma_T TP_j + \gamma_{Tw} TP_j \, lowwage_i \qquad \textbf{(11)}$$

This in turn implies

$$\Pi_{ij} = \frac{e^{-\gamma_p p_j + \gamma_T TP_j + \gamma_{Tw} TP_j * lowwage_i}}{\sum_{k=1}^{J_g} e^{-\gamma_p p_k + \gamma_T TP_k + \gamma_{Tw} TP_k lowwage_i}} \qquad \textbf{(12)}$$

[20]Groups must be homogeneous in the explanatory variables of the model. For some plans, firms report that the price to the employee varies with the employee's age, income, wage/salary status or health habits. However, only one composite price is recorded in the data. The results reported below are from models that eliminate plans where employer price varies across employees. I have also estimated models with these plans included. As might be expected, in the latter case where the employee price is measured with error, the coefficient estimates are somewhat smaller in absolute value, probably reflecting a bias toward zero due to measurement error.

In the aggregated share model with large group sizes, we have:

$$\ln \frac{\Pi_{jg}}{\Pi_{Jg}} = \text{plim } \ln \frac{N_{jg}}{N_{Jg}}$$

$$= \text{plim } [(\ln \frac{N_{jg}}{N_{Jg}} \mid \text{lowwage} = 0) \cdot \text{prob}(\text{lowwage} = 0) + [(\ln \frac{N_{jg}}{N_{Jg}} \mid \text{lowwage} = 1) \cdot \text{prob}(\text{lowwage} = 1)])$$

(13)

$$= -\gamma_p (p_{jg} - p_{Jg}) + \gamma_T (TP_{jg} - TP_{Jg}) + \gamma_{TM} (TP_{jg} - TP_{Jg}) \cdot \text{prob}(\text{lowwage}).$$

Using the proportion of low wage workers at the firm, lowwageshare_g, as an estimate of prob(lowwage = 1), the estimating equation is, then:

$$\ln \frac{N_{gj}}{N_{Jg}} = -\gamma_p (p_{jg} - p_{Jg}) + \gamma_T (TP_{jg} - TP_{Jg}) + \gamma_{TM} (TP_{jg} - TP_{Jg}) \cdot \text{lowwageshare}_g + \theta_{jg}. \quad \textbf{(14)}$$

Results

The model was estimated separately for single and family coverage on differenced observations. Coefficient estimates for the samples where employee-paid premiums do not vary within the firm and for firms that offer fewer than 20 plans are reported in Tables 2 and 3. Heteroskedastic-consistent standard errors are reported since the use of estimated plan shares from groups of different sizes is expected to introduce heteroskedasticity in the errors. The first column of both tables reports a simple linear specification with only employee premium and total premium. The second column includes the linear employee premium and total premium terms as well as an interaction of total premium and firm size. The third column includes instead an interaction with number of plans per employee. Column four reports results from the model with quadratic effects of employee and total premium and column five adds to the linear model an interaction of total premium with a HMO dummy variable. The last two columns report results from a model with an interaction of total premium with the proportion of the firm's workforce earning less that $20,000/year and a linear model using only the observations for which the proportion less than $20,000/year is not missing.

Note first that the coefficients on employee and total premium have the expected signs. Higher total premiums have a positive effect on workers' utility and higher employee-paid premiums decrease utility. These coefficients are, with only the exception of the quadratic model on the single plan data, individually or, in the models that include interactions with the total premium variable, jointly statistically significant.[21] Unlike many of the wrong-signed results in the hedonic literature, the discrete choice model does yield estimates with signs as predicted by theory.

The implied estimates of willingness to pay for health benefits are reported in row nine of Tables 2 and 3. The simple linear specification of the first column produces estimates of the marginal valuation of health benefits to wages of 0.405 for families and 0.204 for singles. The estimates of the marginal valuation are reasonably stable across models until the interaction with the proportion of low wage workers is included in the model.

---

[21] The last row of the tables reports the F-statistic for the joint significance of all total premium variables and interactions with total premium for all specifications except the simple linear model reported in column one.

All of the regressions imply much smaller valuations of health benefits for singles than families. Although the estimates vary somewhat across specifications, in each case the single estimates are smaller than the equivalent estimate for families. This finding is consistent across other unreported specifications as well. While it may not be surprising to find that families value health benefits more than workers purchasing single coverage, the magnitude of the estimated difference appears surprisingly large.

The models presented in column two of the tables allow for differences in the valuation of benefits with firm size. Larger firms may be able to negotiate better benefits packages so a dollar spent on benefits at a large firm may be worth more to employees than a dollar spent at a small firm. On the other hand, small firms may be able to tailor their benefits more closely to the needs of employees, making small firm benefits dollars more valuable to their employees. An interaction of firm size and total premium may capture effects of this sort. For both singles and families the effect of firm size is negative and significant. For families, the marginal valuation of health benefits dollars is estimated to be 0.471 for a firm with 10 employees as compared to 0.410 for a firm with 1000 employees. The analogous estimates for singles are 0.241 and 0.209.

The idea that a similar number of plans at a small firms may better span the range of employee preferences might also be captured by allowing valuations to differ with the number of plans per employee. I do, in fact, find that plans per employee has a positive effect on employee valuations of health benefit dollars. More plans per employee increases a worker’s valuation of health benefits, although the effect does not appear to be of great enough magnitude to be very important as can be seen by the small effect of adding an additional plan at a firm of average size on marginal willingness to pay.

Quadratic terms in total premium and employee premium are also included in the models of column four. For families, the effect of total premium is increasing at a decreasing rate, as might be expected. For singles the effect of both the linear and quadratic terms is positive. The utility of both families and singles is found to decrease with the employee-paid portion of the premium but at a rate that decreases in absolute value as employee premium increases.

The fifth column of both tables reports results from a model that allows for differences in fringe valuations according to whether the plan is an HMO. Recent media reports would suggest that consumers are not happy with their HMO health coverage and that they would prefer to continue to have access to more flexible forms of coverage such as fee-for-service and preferred provider plans. But do workers actually value the dollars spent by employers on HMOs less than the dollars spent on other types of plans? The results reported in column five suggest that this is indeed the case. The interaction of total premium with the HMO dummy is negative and significant and also quite substantial. Whereas a family is estimated to value a dollar spent on a non-HMO plan at 0.386, a family is estimated to value a dollar spent on an HMO at only 0.126. The effect for singles is even greater. Singles value benefit dollars on a non-HMO at 0.462 as compared to dollars spent on an HMO at 0.045.

Alternative models including additional interactions of workforce and firm characteristics with total premium were also estimated. The effect of such variables was, for the most part, not significant. Column six of Tables 2 and 3 reports the results for an interaction of total premium with the proportion of the firm's workforce earning less than $20,000/year. Earning less than $20,000/year is found to reduce the effect of total premiums on utility of both families and singles although the effect is individually significant only for singles.[22] Notice that the estimated marginal valuations for both groups increases substantially in this specification. Some of this is due to the inclusion of the low-wage interaction and some is due to the loss of the observations for which the low-wage data are missing. The last column of these tables reports the results from the linear model using only the observations for which the low-wage data are not missing. For both singles and families we find substantially higher estimates of marginal willingness to pay for benefits in this sample.[23] We also find that within this sample workers earning more than $20,000/year have a substantially higher valuation of benefits than low-wage workers.

The estimates are somewhat sensitive to other ways of defining the samples used. For example, when I do not eliminate plans for which the employee-premiums differs (in unmeasured ways) within the firm, willingness to pay estimates implied by the linear specification falls to 0.36 for families and to 0.14 for singles. These differences may well be explained, however, by measurement error in the employee premium variable.

The aggregate data currently available provide support for the proposed method but also illustrate the need to estimate the model on individual-level data. Data that will be available in the future from the Medical Expenditure Panel will include information on workers and their health plan choices. Individual demographic and employment data will be available as well as data on all of the health plans available to the worker. These data will allow estimation of a

[22]The estimate for the effect of low salary on valuations come from a differenced regression that includes an interaction of the total premium difference with the share of the workforce at the firm that earns less than $20,000/year. As the discussed above, this specification captures the effect that would be modeled with individual-level data with an interaction of total premium and a dummy variable for earning less than $20,000/year.

[23]This appears to be due the loss of some plans with very small shares of the market. I am currently investigating these cases further.

discrete choice model that allows valuations of fringe benefits to vary across a variety of observed worker and firm characteristics. The aggregate data used here suggest the importance of such differences in a few dimensions but very little such data are available in the RWJ-EHIS data. Other important shortcomings of the aggregate data can also be avoided with individual data. For example, individual-level data on employee-paid premiums will avert the measurement error problem induced when firms report only one employee-paid premium even when employee premiums vary across workers within the firm.

Conclusions

This paper presents a discrete choice alternative to the hedonic model traditionally used to estimate compensating differentials and the demand for job attributes and fringe benefits. I use worker choices among alternative fringe benefit packages to identify worker preferences for fringe benefits versus wage dollars. I argue that the differencing across alternatives implicit in the discrete choice model offers a more effective individual-fixed effect control than standard fixed effect wage equations, thus allowing the econometrician to overcome the problem of unobservable productivity or other individual- or job-level fixed effects that bias compensating differentials estimates from hedonic wage equations. Observed variation in the offered wage-fringe price across firms identifies workers preferences, eliminating the need to impose arbitrary market boundaries to achieve identification as is necessary in the hedonic model. I also argue that the endogeneity problems that plague estimation of structural demand parameters in the hedonic model are less severe in the proposed discrete choice problem, again due to the "fixed effects" nature of the discrete choice model, since the matching of workers to firms will induce an endogeneity problem only when the job selection occurs on deviations from the firm-level means in the value of the offered fringe benefits alternatives.

Aggregate data at the firm level provides support for the method outlined in this paper. Coefficient estimates are of the predicted sign, unlike many estimates that emerge from hedonic wage estimations. Results from the aggregate data suggest that families value health benefits substantially more than singles but at less than one-for-one with wage dollars. The aggregate data currently available have some important disadvantages relative to individual-level data, however, and therefore the empirical results should be viewed cautiously. For example, previous empirical work on demand for health care and worker choices among health insurance plans indicate that the valuation of benefits does differ with observable worker and firm characteristics as does some of the evidence from these models; yet the available data contain little such information.

The insights developed in this paper about the advantages of the discrete choice model and its relationship to the hedonic model can also be translated into other settings, such as the valuation of urban amenities or the valuation of product attributes, when appropriate data are available. This method is most effective when a traditional hedonic problem can be reformulated in such a way such that data on more than just the chosen alternative are available and when troublesome omitted variables are held fixed across the observed alternatives.

Table 2
Linear Logit Model
Family Plan Data*
(Heteroskedasticity-Corrected Standard Errors in Parentheses)

Worker's share of premium in $ ÷ 100	-0.182 (0.035)	-0.176 (0.035)	-0.182 (0.035)	-0.279 (0.069)	-0.178 (0.036)	-0.146 (0.035)	-0.146 (0.035)
Worker premium squared ÷ 100	---	---	--	0.0002 (0.0001)	---	---	---
Total premium in $ ÷ 100	0.074 (0.028)	0.0083 (0.028)	0.061 (0.032)	0.104 (0.070)	0.069 (0.028)	0.090 (0.041)	0.085 (0.028)
Total premium squared ÷ 100	---	---	---	-0.00003 (0.00007)	---	---	---
Firm size* total premium ÷ 100	---	-0.00001 (0.000004)	---	---	---	---	---
Plans per employee * total premium ÷ 100	---	---	0.116 (0.100)	---	---	---	---
HMO dummy * total premium ÷ 100	---	---	---	---	-0.046	---	--
Share of workforce earning <$20,000/year* total premium ÷ 100	---	---	---	---	---	-0.011 (0.066)	---
Implied marginal value in $ of $1 in health benefits							
• Linear model	0.405						0.582
• At variable means		0.435	0.397	0.350			
• Firm size = 10		0.471					
• Firm size = 1000		0.410					
• At mean plans per employee			0.397				
• Add 1 plan for firm with mean number of employees			0.398				
• HMO = 0					0.386		
• HMO = 1					0.126		
• >$20,000 salary						0.620	
• <$20,000 salary						0.542	
_N_ (for differenced observations)	2158	2158	2158	2158	2158	2038	2038
F stat (P value) for joint significance of all total premium variables and interactions with total premium	---	7.55 (0.0005)	5.07 (0.0064)	3.55 (0.0289)	9.83 (0.0001)	4.80 (0.0083)	

* Observations for firms where employee-paid premium does not differ across employees and total plans offered by the firm < 20.

Table 3
Linear Logit Model
Single Plan Data*
(Heteroskedasticity-Corrected Standard Errors in Parentheses)

Worker's share of premium in $ ÷ 100	-0.628 (0.106)	-0.604 (0.105)	-0.642 (0.108)	-1.049 (0.183)	-0.628 (0.107)	-0.554 (0.103)	-0.555 (0.104)
Worker premium squared ÷ 100	---	---	---	0.0024 (0.0008)	---	---	---
Total premium in $ ÷ 100	0.128 (0.063)	0.146 (0.063)	0.071 (0.072)	0.046 (0.118)	0.115 (0.063)	0.256 (0.087)	0.145 (0.062)
Total premium squared ÷ 100	---	---	---	0.00015 (0.00023)	---		---
Firm size* total premium ÷ 100	---	-0.00002 (0.000007)	---	---	---	---	---
Plans per employee * total premium ÷ 100	---	---	0.344 (0.175)	---	---	---	---
HMO dummy * total premium ÷ 100	---	---	---	---	-0.082 (0.033)	---	---
Share of workforce earning <$20,000/year* total premium ÷ 100	---	---	---	---	---	-0.231 (0.142)	---
Implied marginal value in $ of $1 in health benefits							
• Linear model	0.204						0.261
• At variable means		0.222	0.168	0.100			
• Firm size = 10		0.241					
• Firm size = 1000		0.209					
• At mean plans per employee			0.168				
• Add 1 plan for firm with mean number of employees			0.169				
• HMO = 0					0.183		
• HMO = 1					0.053		
• >$20,000 salary						0.462	
• <$20,000 salary						0.045	
_N_ (for differenced observations)	2223	2223	2223	2223	2223	2104	2104
F stat (P value) for joint significance of all total premium variables and interactions with total premium		6.87 (0.0011)	4.56 (0.0106)	1.85 (0.1571)	4.96 (0.0071)	4.61 (0.0101)	

* Observations for firms where employee-paid premium does not differ across employees and total plans offered by the firm < 20.

BIBLIOGRAPHY

Abowd, John M. Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* Vol. 67 No. 2.

Ben-Akiva, Moshe and Steven R. Lerman. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge: MIT Press.

Berkson, J. 1953. "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response, Based on the Logistic Function." *Journal of the American Statistical Association* Vol. 48. pp. 565-599.

Berry, Steven. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *Rand Journal of Economics* Vol. 25, No. 2. pp. 242-262.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* Vol. 63, No. 4. pp. 841-890.

Biddle, Jeff E. and Gary A. Zarkin. 1988. "Worker Preferences and Market Compensation for Job Risk." *Review of Economics and Statistics* Vol. 70. pp. 660-667.

Brown, Charles. 1980. "Equalizing Differences in the Labor Market." *Quarterly Journal of Economics* Vol. XCIV, No. 1. pp. 113-134.

Brown, James N. and Harvey S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." *Econometrica* Vol. 50, No. 3. pp. 765-768.

Bureau of Labor Statistics. 1995. Employment Cost Index and Employee Benefits Survey.

Cropper, Maureen L., Leland Deck, Nalin Kishor, and Kenneth E. McConnell. 1993. "Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches." *Review of Economics and Statistics* Vol. 75, No. 2. pp. 225-232.

Diamond, Douglas B., Jr. and Barton A. Smith. 1985. "Simultaneity in the Market for Housing Characteristics." *Journal of Urban Economics*, Vol. 17, pp. 280-292.

Duncan, Greg J. and Bertil Holmlund. 1983. "Was Adam Smith Right After All? Another Test of the Theory of Compensating Wage Differentials." *Journal of Labor Economics* Vol. 1, No. 4. pp. 366-379.

Epple, Dennis. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy* Vol. 95, No. 11. pp. 59-80.

Garen, John. 1988. "Compensating Wage Differentials and the Endogeneity of Job Riskiness." *Review of Economics and Statistics* Vol. 70. pp. 9-16.

Hersch, Joni and W. Kip Viscusi. 1990. "Cigarette Smoking, Seatbelt Use, and Differences in Wage-Risk Tradeoffs." *Journal of Human Resources* Vol. XXV, No. 2. pp. 202-227.

Hwang, Hae-shin, W. Robert Reed, and Carlton Hubbard. 1992. "Compensating Wage Differentials and Unobserved Productivity." *Journal of Political Economy* Vol. 100, No. 41. pp. 835-858.

Kahn, Shulamit and Kevin Lang. 1988. "Efficient Estimation of Structural Hedonic Systems." International Economic Review Vol. 29, No. 1, pp. 157.166.

Mason, Carl and John M. Quigley. 1990. "Comparing the Performance of Discrete Choice and Hedonic Models." in <u>Spatial Choices and Processes</u>, (M.M. Fischer, P. Nijkamp and Y.Y. Papageorgiou, Eds.). Amsterdam: Elsevier Science Publishers. pp. 219-246.

Montgomery, Edward, Kathryn Shaw, and Mary Ellen Benedict. 1992. "Pensions and Wages: An Hedonic Price Theory Approach." *International Economic Review* Vol. 33, No. 1. pp. 111-128.

Nechyba, Thomas J. and Robert P. Strauss. 1998. "Community Choice and Local Public Services: A Discrete Choice Approach." *Regional Science and Urban Economics* Vol. 28, pp. 51-73.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82. pp. 34-55.

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." in <u>Handbook of Labor Economics</u>, Vol. 1, (Orley C. Ashenfelter and Richard Layard, eds.). Amsterdam: North Holland. pp. 641-692 .

Small, Kenneth and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* Vol. 49, No. 1. pp. 105-130.

Theil, Henri. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* Vol.10. pp. 251-259.

Trajtenberg, Manuel. 1989. "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners." *Journal of Political Economy* Vol. 97, No. 2. pp. 444-479.

U.S. Chamber of Commerce. 1992. <u>Employee Benefits</u>.

Woodbury, Stephen H. 1983. "Substitution Between Wage and Nonwage Benefits." *American Economic Review* Vol. 73, No. 1. pp. 166-182.

Woodbury, Stephen A. and Hamermesh, Daniel S. May 1992. "Taxes, Fringe Benefits and Faculty." *Review of Economics and Statistics* Vol. 74, No. 2. pp. 287-296.

Appendix A

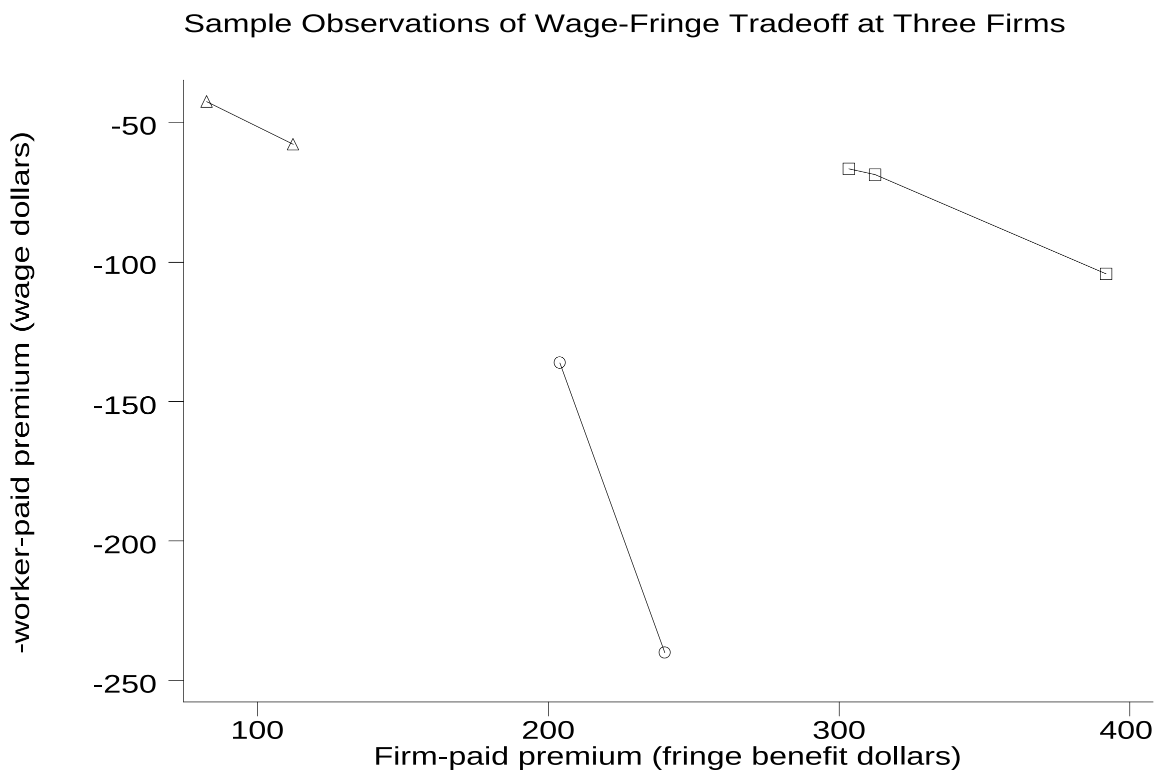

Figure 1