Accounting for wages and benefits using the ECI

Using the data set behind the Employer Cost Index to impute benefit values on the National Longitudinal Study of Youth and the Current Population Survey, this study finds that workers at the bottom part of the wage distribution exhibit a much stronger correlation between benefits and wages than those at the top

Jonathan A. Schwabish

Jonathan A. Schwabish is senior director of Research and Policy at the Partnership for New York City. E-mail: jschwabish@pfnyc.org. The views expressed here are those of the author and do not necessarily reflect the views of the Partnership for New York City.

Accounting for employee benefits as a form of real compensation for work has received much theoretical and empirical attention. The hedonic theory of compensating wage differentials, first made popular by Sherwin Rosen, contends that workers make tradeoffs between wages and benefits.[1] That is, in lieu of lower wages, workers are compensated by taking the greater benefits offered by employers. Empirical approaches to estimating the tradeoff however, have generally failed to correspond with theory. A slew of econometric difficulties are only the tip of the iceberg—unobserved worker and firm heterogeneity, measurement error, present discounted value issues (especially for pensions), and group discounts (especially for health insurance) complicate estimation. In addition, data sets often lack the necessary variables to construct and estimate hedonic models. There are few data sets that have enough variables to create such models but even fewer are nationally representative, containing employer and fringe benefit characteristics as well as employee (demographic) information. The hedonic model competes with what is commonly known as the "good jobs, bad jobs" story.[2] An observable feature of the labor market, the "good jobs, bad jobs" story asserts that workers with high wages also receive high benefits. However, perhaps *within* a job there are compensating wage differentials, which would put the hedonic and "good jobs, bad jobs" models in concert. Subsequently, workers at the bottom end of the income distribution may be forced to switch jobs in order to obtain a preferable benefit-wage mix and those at the top of the distribution may be better able to change the mix within their current job.

This study uses the Bureau of Labor Statistics' survey behind the Employer Cost Index (ECI) to impute the value of fringe benefits onto the National Longitudinal Survey of Youth (NLSY) household survey between 1990 and 1998 and formally test the hedonic theory of compensating wage differentials. As an additional check, this study performs the same procedure for the same set of years using the March Current Population Surveys (CPS). Because of concerns relating to endogeneity and measurement error—a common problem in this body of research—conclusions are generally confined to a discussion of the inequality of benefits and the degree of correlation between wages and benefits at different points in the income distribution. Although certain models point to compensating differentials for particular groups, these technical considerations generally lead to a discussion of correlations as opposed to tradeoffs.

In general, coefficients on health insurance, life insurance, and pensions are nontrivial and statistically significant. The sign on the benefit coefficients varies, depending upon the sample and estimation method. Pooling the NLSY observations and estimating a fixed effects model addresses the unobservable heterogeneity of workers; however, estimated coefficients are not markedly different than those from the ordinary least squares model. The CPS permits this study to have a cross-sectional approach and enables broader appeal to the model. The main contribution of the study is to utilize the BLS data in a hedonic framework. The very nature of the data permits employee benefits to enter not simply as dummy variables, as is usually done in the literature, but to enter as continuous, dollar-valued variables. The methodology continues to be troubled by endogeneity concerns but, for the time being, the ECI data sheds new light on nonwage forms of compensation and the distribution of such compensation.

The study begins by reviewing a selection of previous work in the wage-benefit literature, with particular focus on the distribution of employee benefits. It then discusses the various models used in the empirical work, and is followed by a discussion of the data itself. Summary results with decompositions of benefits and wages, and the full model results conclude the study.

Previous literature

The value of employee benefits is of increasing importance to researchers and policymakers alike. According to William Wiatrowski, total benefits make up almost 30 percent of all workers' total compensation from the employers' cost point of view.[3] An August 2001 study from the Employee Benefits Research Institute reports that employee benefits became an even greater proportion of total compensation, rising from 26.8 percent to 28.9 percent between 1987 and 1994.[4] Using Chamber of Commerce data, Masanori Hashimoto shows that between 1951 and 1994, employee costs for Social Security rose by 414 percent; workers' compensation costs increased by 66 percent; retirement costs by 76 percent; and health (and medical) insurance costs, 688 percent.[5] From the existing literature, it is clear that benefits play an important part of workers' compensation and thus the correlation between wages and benefits will have important consequences for public policy and distributional issues.

Several studies in the early 1980s attempted to estimate the *value* of benefits.[6] Additional work on hedonic price theory[7] generated mixed results from hedonic models; some negative coefficients, but not often statistically significant. Charles Brown's review article considers the research on the tradeoffs workers make between wages and job *disamenities*, such as poor working conditions, death rates, noise, heat, and so forth.[8] Brown's own empirical work (using the NLS Young Men's sample for the years 1966–71 and 1973) finds weak evidence for tradeoffs between wages and jobs that have the following characteristics:

- require employees to perform repetitive functions
- involve stressful conditions
- call for physical strength
- contain bad working conditions

Brown postulates several reasons why the hedonic model may not perform well in practice, including poorly measured job characteristic variables, omitted variable bias, and that "[l]abor markets are simply not as competitive as the theory of equalizing differences assumes." These issues continue to be challenging issues for researchers and strides are being continually made with improvements in methods and data sets.

Given the difficulties in estimating hedonic models, research in this area has spilled over into other areas of the labor market, such as mandated benefits and minimum wages. For example, Jonathan Gruber and Alan Krueger examine workers' compensation and mandated health insurance programs.[9] Arguing that looking at the tradeoff effects within entire population bias results, Gruber and Krueger select certain industries (trucking, carpentry, hospitals and plumbing) and show that employers largely shift the costs of providing workers' compensation insurance to the worker in the form of lower wages. Using variation in mandated maternity benefits as an instrumental variable, Jonathan Gruber identifies substantial shifting in benefit costs to workers.[10] Joseph G. Altonji and Christina H. Paxson investigate whether job quitters attempt to achieve more desirable benefit-wage mixes by accepting different jobs.[11] They find that workers typically demand additional compensation for jobs that offer "unattractive" hours (that is, jobs that do not permit workers to change or increase their hours). These papers are unique in the literature in that they are generally successful in finding tradeoffs between these mandated benefits and wages by using State variation in benefit laws as exogenous instruments.

Recent work has paid particular attention to the distribution of benefits and in conjunction with the wage distribution.[12] Brooks Pierce shows that wage and compensation inequality increased over the 1982–96 period, with those in the lower tail of the distribution experiencing the greatest decline.[13] Craig A. Olson uses an instrumental variables approach to estimate the tradeoff between health benefits and wages for wives who work full-time.[14] He contends that, "Husbands working in small firms or in non-union jobs are less likely to have health insurance through their jobs, and this increases the probability that their wives have a job with health insurance through their own employers." Problems with the instrumental variable approach used in Olson's study remain: assortative mating theory and quality of benefits, conditional on firm size and union status may bias the instruments.[15]

In all, research on the tradeoff between wage and non-wage compensation has encompassed several different areas of the labor market. Researchers have used different estimation methods and different data sets to explore these tradeoffs with varying degrees of success. Although earlier studies generally failed to find such tradeoffs, later studies have taken a different view of the model with more success.[16] Focus on distribution of benefits has received new attention in papers from Brooks Pierce and from William J. Carrington and others.[17] Although the results in this article fail to precisely identify the hedonic effect, the study brings these different approaches together while taking advantage of a rich but relatively unused data set.

Data

This study uses three data sets to estimate the various models in the next sections.

Employment Cost Index. Employee benefit cost data, along with wage and salary data, come from the survey used by the Bureau of Labor Statistics to generate the quarterly price index for employer costs, known as the ECI. Quarterly micro data are available back to the early 1990s and contains hourly dollar benefit costs for common benefits including health, dental, and life insurance, defined benefit and defined contribution pension plans,[18] sick and vacation days, overtime, bonuses, shift differentials ("other pay"), and legally required benefits such as Social Security and workers' compensation. Employers are asked the cost of employee benefits by occupation within the firm. Thus, a firm may be asked the cost of health insurance and days off for, say, managers inside the firm. The survey does not include Federal, private-household, or self-employed workers.[19]

Due to attrition in the ECI and the addition of new firms to account for these losses, external employment data from the BLS Web site were used to create consistent weights within each data set and to aggregate the quarterly observations to annual data sets.[20] Using these (weighted) observations, benefit costs were calculated by averaging more than 72 industry-occupation cells for each year between 1990 and 1998.[21] These averages were then merged onto the NLSY and CPS data sets using the same industry-occupation structure. In other words, if individual *n* works in industry-occupation cell *j*, then the benefit level imputed to the individual's record equals the average (weighted) benefit level from industry-occupation cell *j* if the individual receives (or is offered) the benefit, and zero otherwise. Ultimately, the average costs calculated in the study are close to published numbers by BLS.[22]

Several aspects of the ECI made the calculations of average benefit costs especially troublesome. In particular, the survey asks for full-time, part-time status not as a function of actual hours worked, but according to the practice of the employer (that is, benefit schedules). Thus, there are many observations for which the observation is categorized as a full-time worker but worked only 20 hours in the reporting period. It is therefore difficult to distinguish between actual full-time and part-time workers within the ECI to calculate benefits. This may be important because, as has been demonstrated in previous work, part-time workers often receive fewer (if any) benefits, compared with full-time workers.[23] Full-time dummy variables—based on the individual data—are included in some specifications to control for the full-time issue. Furthermore, the benefit averages from the ECI include some observations equal to zero, which could potentially downward bias the benefit-cell averages. However, the inclusion of these zeros may help identify the interfirm trade-off of the benefits at the job level. Additionally, the external weighting scheme may further help to mitigate this potential bias.

National Longitudinal Survey of Youth. The NLSY, beginning in 1979, surveyed individuals between 18 and 25 years old. Initially, more than 12,500 individuals responded to the survey. This study uses 7 of the recent years of data; 1990–1994, 1996, and 1998 (surveys were not conducted in 1995 or 1997). The discussion in the following sections is restricted to the results in 1998, although issues of unobserved worker heterogeneity are addressed by using the panel nature of the NLSY by pooling all 7 years. The demographic variables in the models are detailed in the next section; most importantly, from NLSY data one can find whether certain benefits are made available to workers. Benefit indicators are available for most years and include health, life, and dental insurance; sick, vacation and maternity leave; retirement programs; discounts, profit sharing, and flexible hours; and other indicators such as parking, training/education and child care are also available. This study focuses on health and life insurance, sick and vacation leave, and retirement because these are also included in the ECI data.[24]

Current Population Survey. The March CPS is a familiar data set to labor economists—it surveys approximately 50,000 households every year. The CPS asks several questions about health insurance and pension coverage, but other benefits (except some required benefits) are not covered in the survey. Thus the analysis using the CPS focuses on these two variables with the same set of demographic variables as in the NLSY (less actual work experience). The CPS is a large, cross-sectional data set and thus permits a more thorough analysis of the correlation between wages and benefits. The NLSY, however, is restricted both by the number of observations (although the average number of observations does range between 4,000 and 5,000 for the aggregate groups) and restricted age range of only 9 years.

Measurement concerns. Matching employer to employee data is becoming increasingly common and has the potential to yield large returns to research.[25] Linking these data sets however, introduces new econometric concerns as well. In the perfect world of matching employer information directly to data for their employees, measurement error and issues related to fixed effects (on several different levels) estimation are magnified in the matched data set case.[26] As Robert Elliott and Robert Sandy argue, workers who are dissatisfied with their pay may overstate workplace disamenities, and workers who are satisfied with their pay may understate the disamenities.[27] Such systematic responses to benefit surveys are shown to unambiguously bias hedonic model estimates downward. Panel data sets with matched employer-employee data would help solve some of these problems.

In the mid-1980s, a substantial literature on measurement error in this area was born. Papers by Greg J. Duncan and Daniel H. Hill, and Wesley Mellow and Hal Sider compared separate survey responses between employees and employers.[28] In the latter, the authors found that discrepancies for major industry and occupation responses were minor. Worker responses to number of hours worked tended to exceed employers' responses by about 4 percent, while their reports of wages were lower by almost 5 percent. Duncan and Hill find similar results and use the Panel Study of Income Dynamics Validation Study to show that answers to questions about benefits were relatively accurate. Error rates (percentage of workers whose responses did not match their employers) on fringe benefits were small: 1 percent on medical insurance and paid vacation days; 5 percent on dental benefits; 10 percent on life insurance; and 3 percent on pension coverage. Thus, this measurement error literature provides some evidence that the imputation procedure used in this study provides results that are not inconsistent with expected results.[29]

Given that the econometric issues related to matched data sets are not yet resolved, the imputation used in this study raises additional concerns regarding estimation. The ECI, NLSY, and CPS data sets do not make up a direct match data set but instead simply allow the imputation of benefit costs to individual demographic characteristics. General measurement issues are thus not the only concern; determining the proper matching procedure and testing for significance of the estimated coefficients is also important. First, the matching procedure may not be appropriate if one believes industry and occupation groupings do not adequately reflect employer costs. From the employer, one would expect benefits to be contingent upon other firms in that industry (and occupation) just as one would expect wages or any other form of compensation to be. The confidentiality restrictions placed on the ECI data also restricts further decomposition, such as by region, State, or even more specific industry/occupation groups.

Second, the model is estimated from within the household surveys and one might be concerned about what degrees of freedom to use in the hypothesis tests. The NLSY and CPS have about 5,000 and 55,000 observations in the 1998 sample, and the ECI has approximately 82,000 observations for the same year. Hence, in the subsequent sections, the household data sets are used as the yardstick for significance testing. The smaller number of observations is somewhat more restrictive although with these large numbers of observations, it should make little difference.

There are two other considerations that are important to the discussion of measurement error and are mentioned briefly. The first is the difference between employer cost and employee valuation. The cost of a benefit reported by the employer may differ from the value placed on the benefit by the worker. For example, if one individual in a married couple is capable of working, health insurance may have a greater value to the individuals than, say, an unmarried worker with no children. That said, while it is extremely difficult to measure the value an individual places on a benefit, the ECI data provides an easily quantifiable measure of benefits, namely the cost of the benefits to employers.[30] Second, the NLSY asks whether the respondent was offered a benefit, not whether the worker actually accepted the benefit.[31] Individuals who receive benefits were clearly offered benefits, but also realize that individuals who declined offered benefits still have a compensating differential. These differences are conceptually interesting and data on offered benefits are perhaps preferable to received benefits because receipt of benefits might miss the potential tradeoffs employees could make between benefits and wages. In that case, one would not be able to observe workers who refused such benefits and benefit packages.

Model

Three models are estimated to test the relationship between benefits and wages. After the imputation of average benefit values by industry-occupation cell, ordinary least squares, fixed effects, and quantile regression models are estimated. First, the ordinary least squares model asks what the average wage would be for the worker who has high (or low) benefits in any given industry-occupation combination. Because causality is difficult to identify, the ordinary least squares regression results are interpreted in terms of a wage-benefits correlation arising from across-industry/occupation group averages. This is carried out using both the NLSY and CPS—the dependent variable being the logarithm of the (individual) hourly wage with both wage and benefit data inflated to 1996 CPI-U adjusted dollars.[32] Average benefit values are interacted with the individual-level dummy variables that indicate whether a benefit is offered; thus, the primary variables of interest will be equal to zero if the respondent received no benefit and a (continuous) dollar value if the respondent received the benefit. Demographic controls include age; the number of children under age 6; actual experience and its square[33]; and dummy variables for marital status, urban status, race (black and Hispanic), firm size by number of employees (0–24, 25–99, and 100 or more), full-time status (35 hours or more worked per week), union status, and education level (high school graduate, some college, college graduate). Industry-occupation dummies are included in a second set of regressions; all regressions are weighted with the individual-level sample weights.[34]

The second model exploits the panel nature of the NLSY by pooling the 1990–98 samples and estimating a fixed effects model for the same groups as in the ordinary least squares, using the same covariates as previously specified. The fixed effects regressions identify the partial correlation between wages and benefits by allowing one to look at changes in benefits and wages for a given worker. The fixed effects approach addresses the unobserved heterogeneity of workers in the sample. Concerns remain, however, of firm-level unobserved characteristics but controlling for these factors is much more difficult.

Finally, a series of quantile regressions are estimated for the 10th and 90th percentiles. Quantile regression analysis is becoming increasingly popular with the main advantage being that, like the Least Absolute Deviation (LAD) estimator, estimated coefficients are not overly sensitive to outlier data points. Quantile regressions are very different from ordinary least squares regression and is best explained by analogy: regular ordinary least squares summarizes how the mean value of the dependent variable varies with some X regressor, whereas quantile regressions summarize how some quantiles (that is, median or 10th percentile) of the dependent variable vary with some X regressor. The extreme quantile regressions and the ordinary least squares results would differ if wage dispersion were very different in high-benefit cells than in low-benefit cells.[35]

Several studies have recognized that benefit variables, when measured as dummy variables indicating the existence of the benefit, are endogenous when included in an ordinary least squares framework. (See "Previous literature" section.) Olson notes that in an ordinary least squares model, the dummy variable for health insurance "is biased, and the positive sign suggests that unobserved factors affecting wages that are correlated with [the variable] more than offset the trade-off between wages and health insurance predicted by the theory."[36] This potential endogeneity is not eliminated by using the ECI data, although the variable of interest is no longer a latent variable and is instead a continuous dollar value. In addition, the fixed effects model from the NLSY addresses some of these concerns. In light of these issues, arguing causality based on these results may be difficult; however, the negative coefficients in some regressions at least suggest a compensating differential. Combining the ECI data with the NLSY and CPS allows for a unique look at the correlations between wages and benefits that may help our understanding of the issues relating to workers' compensation. The main contribution is to take advantage of the underutilized ECI data set to explore these correlations and tradeoffs and to generate dollar-valued benefit variables rather than simple indicator variables. As will be seen later in the article, however, the models ultimately produce few negative coefficients on the benefit variables.

Subgroups. Each of the models is estimated for different subgroups of the population. The first subgroup of interest consists of individuals at the extreme ends of the income distribution—the 10th and 90th percentiles—based on the individual level hourly wage rates. As shown by Pierce, workers at the lower end of the distribution take a larger percentage of their total compensation in terms of wages than those at the higher ends of the distribution.[37] (See sections, "Summary results" and "Model results.") Hence, one might also expect workers at different points in the distribution to correlate wages and benefits differently. And, as will be seen in the following sections, the payoff from such exploration will be significant. Selection into these different components of the wage distribution however, may be an issue of concern. To address this, a series of quantile regressions (discussed earlier) are estimated with some success. A separate look at these groups may also have important consequences for policy considerations of distribution and inequality.

Estimates were also generated for several other sub-groups, but due to space constraints and the fact that the results did not shed further light on the topic, those figures are not reported here.[38] The additional subgroups include, men and women; workers earning at least 25 cents above the minimum wage; workers earning at least 50 cents above minimum wage[39]; and four "at-risk" groups including single mothers, low education youths, single fathers, and working-age black women.

Summary results

As reported by Craig Copeland, the number of employers sponsoring pension plans grew by more than 5 percentage points between 1992 and 2000, and the number of employees participating in a plan grew from 47 percent to 52 percent over the same period.[40] Health insurance continues to be a highly valued form of compensation to workers in the United States. In 2001, 60 percent of those polled in an Employee Benefits Research Institute/Matthew Greenwald & Associates poll said health insurance was the most important benefit, followed by 23 percent saying retirement plans were the most important.[41]

Summary statistics of interest for the NLSY and the CPS 1998 samples used here (following imputation) are detailed in table 1, along with breakdowns by subgroups of interest, namely the 10th, 50th, and 90th wage percentiles. Summary statistics for other years are not dramatically different and, along with full regression results, can be obtained from the author. The summary statistics in table 1 illustrate some of the differences between the two data sets. CPS respondents earn about 50 cents less per hour and are a couple of years older than their NLSY counterparts.[42] Individuals in the NLSY are also much more likely to live in an urban area and belong to a union, although somewhat less likely to work full-time (defined as 35 hours or more per week). In terms of benefit variables, the NLSY contains a much richer set of variables although most are not included in the regression analysis. The benefit indicator variables show that NLSY respondents are somewhat more likely to be offered health insurance or pension plans, with other benefit variables included simply for reader interest.[43]

Table 2 breaks down the mean dollar values of the various benefit variables (following imputation) into the 10th, 50th, and 90th percentiles along with percentages of the total compensation. The patterns are as expected—people at the upper part of the distribution typically earn more in benefits, sometimes more than double workers at the 10th percentile. Health insurance and Social Security account for the biggest parts of the total benefit package, whereas some of the unemployment-type insurance benefits, such as supplemental and long-term disability, account for the smallest parts of total benefit compensation.

The variation in the distribution of benefits is interesting in its own right. For example, take the primary variables of interest (in table 2)—health and life insurance, and pension

Table 1. Summary statistics of the Current Population Survey and the National Longitudinal Survey of Youth, 1998

Variable	1998 CPS			1998 NLSY		
	Number of observations	Mean	Standard deviation	Number of observations	Mean	Standard deviation
Hourly wage (dollars)	55,439	15.509	6.517	4,735	16.068	6.453
Log(hourly wage)	55,439	2.470	.753	4,735	2.567	.610
Age	55,439	38.945	11.521	4,735	36.983	2.309
Black(0,1)	55,439	.118	.323	4,735	.135	.342
Hispanic(0,1)	55,439	.098	.298	4,735	.057	.232
High school graduate(0,1)	55,439	.324	.468	4,735	.436	.496
Some college(0,1)	55,439	.289	.453	4,735	.227	.419
College graduate(0,1)	55,439	.280	.449	4,735	.268	.443
Married(0,1)	55,439	.589	.492	4,735	.661	.473
Previously married(0,1)	55,439	.153	.360	4,735	.180	.384
Number of children	55,439	.237	.555	4,735	.224	.478
Urban(0,1)	55,439	.246	.431	4,735	.645	.479
Union(0,1)	11,497	.148	.355	4,584	.164	.370
Full-time[1]	55,439	.853	.354	4,735	.750	.433
Actual experience[2]	…	…	…	4,735	705.106	207.625
Benefit variables						
Health insurance(0,1)	55,439	0.606	0.489	4,735	0.812	0.391
Pension(0,1)	55,439	.631	.483	4,735	.708	.455
Life insurance(0,1)	…	…	…	4,735	.704	.457
Dental insurance(0,1)	…	…	…	4,716	.678	.467
Maternity coverage(0,1)	…	…	…	4,479	.690	.463
Flex-time(0,1)	…	…	…	4,731	.544	.498
Profit sharing(0,1)	…	…	…	4,694	.292	.455
Child care(0,1)	…	…	…	4,651	.073	.259
Number of sick days	…	…	…	4,479	23.620	63.124
Number of vacation days	…	…	…	4,602	13.748	29.888
Total number of days[3]	…	…	…	4,735	47.032	78.793

[1] Full-time defined as greater than or equal to 35 hours worked per week.

[2] Total number of weeks worked. See text, endnote 32.

[3] Includes sick and vacation days off. See text, endnote 24.

NOTE: Statistics are weighted using the CPS or NLSY individual weights.

Table 2. Average benefit values, total and 10th, 50th, 90th percentiles, Current Population Survey and National Longitudinal Survey of Youth, 1998

Panel A — CPS sample	Total		10th percentile		50th percentile		90th percentile	
Variable	Mean	Percent[1]	Value	Percent[1]	Value	Percent[1]	Value	Percent[1]
Sum	6.088	100.00	2.197	100.00	5.340	100.00	11.085	100.00
Vacations	.751	12.34	.257	11.69	.625	11.71	1.518	13.70
Holidays	.518	8.50	.189	8.59	.460	8.62	.985	8.89
Sick leave	.215	3.53	.037	1.70	.175	3.28	.395	3.56
Other paid leave	.073	1.21	.016	.73	.052	.98	.154	1.39
Shift differential	.067	1.11	.001	.05	.035	.65	.207	1.87
Nonproduction bonus	.289	4.75	.045	2.03	.213	3.99	.579	5.22
Severance pay	.027	.45	.001	.04	.019	.35	.071	.64
Supplemental unemployment	.009	.15	.000	.00	.000	.00	.014	.13
Life insurance	.054	.88	.013	.58	.047	.87	.106	.96
Health insurance	1.225	20.12	.520	23.68	1.377	25.79	1.916	17.28
Sickness and accident insurance	.037	.61	.013	.59	.028	.52	.080	.72
Defined benefit	.500	8.21	.089	4.07	.405	7.58	.926	8.35
Defined contribution	.340	5.59	.086	3.93	.258	4.82	.726	6.55
Social Security	1.006	16.53	.533	24.25	.883	16.54	1.530	13.80
Medicare	.254	4.18	.126	5.74	.209	3.92	.410	3.70
Federal unemployment insurance	.027	.44	.023	1.04	.027	.51	.035	.31
State unemployment insurance	.095	1.57	.079	3.57	.086	1.61	.123	1.11
Workers' compensation	.350	5.75	.127	5.79	.267	5.00	.642	5.79
Long-term disability	.029	.48	.005	.21	.024	.44	.050	.45

Panel B — NLSY sample	Total		10th percentile		50th percentile		90th percentile	
Variable	Mean	Percent[1]	Value	Percent[1]	Value	Percent[1]	Value	Percent[1]
Sum	6.234	100.00	2.226	100.00	5.317	100.00	11.109	100.00
Vacations	.764	12.26	.257	11.54	.625	11.76	1.518	13.67
Holidays	.527	8.45	.189	8.47	.460	8.65	.985	8.87
Sick leave	.215	3.45	.053	2.36	.154	2.90	.395	3.55
Other paid leave	.074	1.18	.018	0.79	.050	.93	.151	1.36
Shift differential	.066	1.06	.001	0.04	.036	.67	.207	1.86
Nonproduction bonus	.297	4.77	.040	1.81	.205	3.86	.579	5.21
Severance pay	.028	.45	.001	.04	.019	.35	.071	.64
Supplemental unemployment	.009	.15	.000	.00	.000	.00	.022	.20
Life insurance	.057	.92	.013	.57	.047	.88	.106	.96
Health insurance	1.250	20.06	.525	23.57	1.377	25.90	1.916	17.24
Sickness and accident insurance	.037	.59	.013	.59	.025	.48	.080	.72
Defined benefit	.513	8.22	.089	4.02	.410	7.71	.926	8.33
Defined contribution	.349	5.60	.086	3.88	.240	4.52	.726	6.54
Social Security	1.024	16.42	.533	23.93	.883	16.61	1.530	13.77
Medicare	.259	4.15	.126	5.66	.209	3.94	.410	3.69
Federal unemployment insurance	.026	.42	.023	1.02	.027	.50	.035	.31
State unemployment insurance	.098	1.58	.079	3.53	.087	1.64	.124	1.12
Workers' compensation	.379	6.08	.127	5.71	.267	5.02	.660	5.94
Long-term disability	.030	.47	.005	0.22	.024	.45	.050	.45

[1] Percent of total benefits.

[2] Called "pensions and retirements" by BLS prior to June 1995.

[3] Called "savings and thrift" by BLS prior to June 1995.

NOTE: Statistics are weighted using the CPS or NLSY individual weights. Life insurance, health insurance, defined benefit and defined contribution are primary variables of interest.

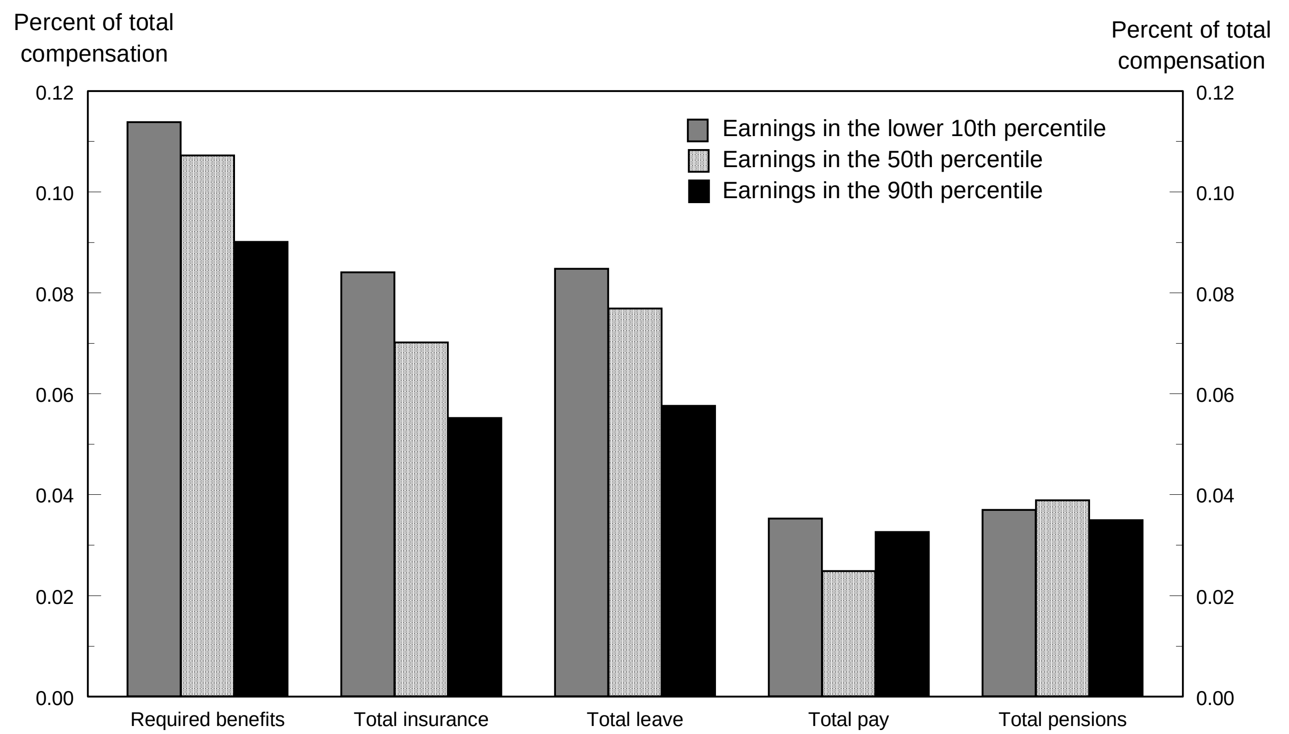
benefits (broken down into defined benefit and defined contribution plans). Health insurance clearly makes up the largest portion of total benefit compensation and is more important to those at the 10th and 50th percentiles than those at the upper tail of the distribution. Life insurance makes up less than 1 percent of benefit compensation, but is increasing through the distribution from roughly 0.5 percent for the 10th percentile to 1.01 percent for the 90th percentile. Pension benefits follow a similar pattern—those at the high end of the distribution take nearly *10 times* as much in pension benefits (around 90 cents per hour) than do those at the 10th percentile (around 9 cents per hour).

To further explore the differences in compensation across the distribution, chart 1 pictures the breakdowns in total compensation (for 1998) across the 10th, 50th, and 90th percentiles using the NLSY data set. As is clear from the chart, individuals whose earnings place them in the lower 10th percentile of the wage distribution take a higher percentage of their total compensation in the form of insurance (8 percent, including health, life, and other insurance) and leave (8 percent). Those in the upper part of the distribution however, take more in wages (72 percent versus 65 percent, not pictured) and less in required benefits (including Social Security), insurance, and leave. The differences between the percentiles are much less for total pay (including shift differential, nonproduction bonus, severance pay, and supplemental unemployment) and total pensions (including defined benefit and defined contribution). For the former, workers at the bottom part of the distribution take 3.5 percent of total compensation in the form of total pay, compared with those in the 90th percentile who take 3.3 percent. The differences are similar for pensions; 3.7 percent versus 3.5 percent. Visually, the differences in total compensation across the wage distribution are striking and, as will be shown in the following sections, estimating the correlations in an econometric model also produces different estimates across the distribution.

Model results

The most simple of the models used in this study is a standard ordinary least squares and is presented first. A fixed-effects model is estimated using the 7-year sample of the NLSY from 1990 to 1998. Finally, quantile regressions were estimated on both the CPS and NLSY to gain a better sense of what is happening within the 10th and 90th percentiles. Each model is

Chart 1. Benefits as a percentage of total compensation, National Longitudinal Survey of Youth, 1998



NOTE: Hourly wage not illustrated.

discussed in turn with some coefficient estimates converted to elasticities in order to better understand the magnitude of the regression results.

Ordinary least squares. Six ordinary least squares models are estimated for each year using the CPS and 12 ordinary least squares regressions are estimated using the NLSY—absence of the number of days off variable in the CPS accounts for the discrepancy. The base model, as noted earlier, includes 12 covariates plus the variables of interest namely, health and life insurance coverage, pension coverage, and number of days off (sick plus vacation). Additional controls include union status, full-time status, and firm size dummy variables. Each benefit variable is included separately and then together—industry and occupation dummies are included in an analogous set of regressions. The hedonic theory implies that there should be a negative coefficient on the variables of interest although in the sections that follow, positive (and statistically significant) coefficients will be the norm. These results demonstrate the degree of correlation between wages and benefits and how that correlation differs within the distribution although negative coefficients, at the minimum, suggest a compensating differential.

Tables 3 and 4 present the coefficients on the benefit variables from the 1998 ordinary least squares model results for the base models for all respondents and those below the 10th percentile and above the 90th percentile—the regressions are repeated with industry and occupation dummies and are found in the neighboring column. Panel A of tables 3 and 4 contains estimates from the whole sample and the estimated coefficients are, given the previous compensating wage differential literature, not surprising. In the regressions that do not include industry and occupation dummy variables, the coefficients are uniformly positive and are statistically significant. For health insurance, the coefficients are consistent between specifications and data sets, ranging between 0.11 and 0.22; life insurance (for the NLSY) is a bit higher with coefficients ranging between 0.24 and 0.50. Estimates for pensions differ slightly between the two data sets with NLSY point estimates ranging between 0.07 and 0.12 and coefficients from the CPS barely higher, between 0.09 and 0.14. All enter statistically significantly and the model fit (measured by the R^2) is fairly strong, ranging between 0.33 and 0.44. In the columns that include industry and occupation dummies, coefficients are uniformly smaller but of the same sign and significance.[44]

As expected—based on the previous literature—the regression coefficients on fringe benefit variables enter positively in the basic model. This is not completely surprising because we might expect that jobs with good pay would be accompanied by good benefits. The estimates in the base models for the entire population are indicative of this phenomenon with positive coefficients ranging roughly between 0.10 and 0.20 for health insurance implying nontrivial elasticities slightly larger than 0.16 for both data sets. Pension coefficients from both data sets imply a slightly smaller elasticity of around 0.08. Due to the endogeneity and measurement concerns noted earlier, conclusions regarding causality are difficult to make, although they are suggestive. On the other hand, the estimates do shed light on the degree of correlation between wages and benefits and suggest that health and life insurance are more strongly correlated with wages than pensions, a conclusion confirmed by the survey results found by the Employment Benefit Research Institute.[45]

The positive coefficients generated on the benefit variables are primarily indicative of an identification problem. Thus, although the "good jobs, bad jobs" story is not surprising (for example, CEO's have higher pay and better benefits than, say, janitors), the positive estimates in the tables imply that holding constant the set of covariates and industry/occupation dummies does not accurately hold job type constant. Note however, that workers in "good jobs" may be able to change the mix of their benefit package (for example, cafeteria-type plans) but workers in "bad jobs" may have to physically switch jobs to gain their preferred levels of wages and benefits.[46] Hence, identification of the tradeoff *within* job is not established but the *correlation* between wages and benefits is established.

Exhibit 1 (page 37) illustrates the compensating differentials theory diagrammatically with benefits on the horizontal axis and wages on the vertical axis. The isocost lines of two firms are pictured and workers' indifference curves are drawn tangent at different points to the firm's locus. For each firm, worker A prefers higher wages and fewer benefits, compared with workers B and C. If firms choose to offer wage-benefit packages so that line I is constructed, this will be the coefficient identified by the regressions in the top panels of tables 3 and 4. Similarly, if the higher wage firms offer relatively higher (lower) wages and fewer (more) benefits to their lower wage counterparts, the estimation will identify line II (III). The "hedonic line" is what these models aim to identify because the negative slope implies a tradeoff between wages and benefits. The implication of positive estimates from the model, therefore, is lack of *within job* identification. By segmenting the wage distribution into percentiles and focusing on the bottom and top deciles, we may be better able to detect those effects.

The top panels of tables 3 and 4 provide evidence for the issues discussed earlier, however, the results in the lower panels suggest otherwise. The estimates in panel B use the same models for those above the 90th percentile point in the income distribution.[47] The coefficients are *negative* and

Table 3. Regression results of benefits from the National Longitudinal Survey of Youth via ordinary least squares model, 1998

	Equation									
Panel A — whole sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	0.151	0.142	0.149	0.140	0.139	0.124	0.116	0.101	0.110	0.090
	(.014)**	(.016)**	(.014)**	(.016)**	(.014)**	(.016)**	(.014)**	(.016)**	(.015)**	(.016)**
Life insurance ..	.501	0.244	.502	0.248	0.539	.306	.521	.279	.519	.274
	(.106)**	(.113)*	(.106)**	(.113)*	(.102)**	(.108)**	(.101)**	(.107)**	(.101)**	(.107)*
Pension	0.120	0.087	.117	.084	.111	.069	.113	.071	.112	.066
	(.013)**	(.013)**	(.013)**	(.013)**	(.013)**	(.013)**	(.012)**	(.013)**	(.012)**	(.013)**
Total days off ...	…	…	.000	.000	.000	.000	.000	.000	0.000	.000
	…	…	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**
Constant	2.061	2.301	2.045	2.287	2.053	2.325	1.930	2.206	1.917	2.184
	(.146)**	(.147)**	(.146)**	(.147)**	(.143)**	(.143)**	(.142)**	(.143)**	(.142)**	(.143)**
Observations ...	4,735	4,735	4,735	4,735	4,584	4,584	4,584	4,584	4,584	4,584
R-squared	.37	.40	.37	.40	.39	.43	.40	.44	.40	.44
Panel B — 90th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	–0.097	–0.069	–0.101	–0.072	–0.057	–0.038	–0.063	–0.044	–0.049	–0.033
	(.04803)*	(.057)	(.047)*	(.057)	(.045)	(0.056)	(.045)	(.056)	(.047)	(.056)
Life insurance ..	–.279	–.394	–.263	–.361	–.208	–0.279	–.221	–.300	–.216	–.293
	(.206)	(.242)	(.205)	(.240)	(.196)	(0.227)	(.196)	(.227)	(.196)	(.227)
Pension	.033	.029	.028	.027	.017	0.007	.016	.006	.017	.007
	(.029)	(.031)	(.029)	(.031)	(.028)	(0.030)	(.028)	(.030)	(.028)	(.030)
Total days off ...	…	…	.000	.000	.000	0.000	.000	.000	.000	.000
	…	…	(.000)**	(.000)**	(.000)*	(0.000)*	(.000)*	(.000)*	(.000)*	(.000)*
Constant	3.295	3.352	3.195	3.264	3.092	3.206	3.018	3.106	3.022	3.115
	(.538)**	(.568)**	(.535)**	(.565)**	(.491)**	(0.516)**	(.493)**	(.519)**	(.494)**	(.519)**
Observations ...	474	474	474	474	478	478	478	478	478	478
R-squared	0.04	0.08	0.05	0.09	0.06	0.10	.06	.11	.07	.11
Panel C —10th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	–0.070	–0.046	–0.068	–0.047	–0.102	–0.085	–0.131	–0.119	–0.154	–0.143
	(.069)	(.074)	(.070)	(.074)	(.068)	(.073)	(.069)	(.074)	(.071)*	(.076)
Life insurance ..	–2.270	–1.918	–2.278	–1.912	–2.353	–2.015	–2.317	–1.980	–2.218	–1.889
	(.876)**	(.930)*	(.877)**	(.933)*	(.856)**	(.909)*	(.854)**	(.905)*	(.855)**	(.905)*
Pension	0.385	.343	.389	.342	.390	.340	.402	.349	.384	.328
	(.122)**	(.126)**	(.123)**	(.127)**	(.120)**	(.124)**	(.120)**	(.124)**	(.120)**	(.124)**
Total days off ...	…	…	–.000	.000	–.000	–.000	–.000	–.000	–.000	–.000
	…	…	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Constant	1.087	.845	1.080	.846	1.161	.916	1.080	.840	1.080	.872
	(.440)*	(.484)	(.441)*	(.484)	(.444)**	(.482)	(.445)*	(.481)	(.444)*	(.481)
Observations ...	476	476	476	476	464	464	464	464	464	464
R-squared	.09	.14	.09	.14	.10	.14	.10	.15	0.11	.16
Industry/ occupation dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Additional covariates	…	…	…	…	Union(0,1)	Union(0,1)	Union(0,1) Full-time(0,1) Firm size	Union(0,1) Full-time(0,1) Firm size	Union(0,1) Full-time(0,1) Firm size	Union(0,1) Full-time(0,1) Firm size

NOTE: Standard set of covariates include age; number of children under age 6; and indicator variables for black, Hispanic, high school graduate, some college, college graduate, married, previously married, urban residence, and actual experience. Data set—NLSY; Year: 1998; dependent variable: log (hourly wage). Standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent.

statistically significant with life insurance entering the largest in magnitude. The coefficient on health insurance in the NLSY is significantly smaller than its counterpart for the whole population (table 3, panel A) and is only statistically significant in the model without additional covariates. The same holds true for the CPS sample in which the magnitude of

Table 4. Regression results for Current Population Survey via ordinary least squares model, 1998

	Equation							
Panel A —whole sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health insurance	0.228	0.199	0.202	0.168	0.175	0.148	0.172	0.142
	(.004)**	(.004)**	(.008)**	(.009)**	(.008)**	(.009)**	(.009)**	(.009)**
Pension	.145	.109	.132	.089	.130	.088	.128	.080
	(.004)**	(.005)**	(.009)**	(.009)**	(.009)**	(.009)**	(.009)**	(.010)**
Constant	1.463	1.784	1.481	1.848	1.346	1.724	1.331	1.708
	(.013)**	(.019)**	(.027)**	(.039)**	(.029)**	(.041)**	(.030)**	(.041)**
Observations	55,439	55,439	11,497	11,497	11,497	11,497	11,497	11,497
R-squared	.30	.32	.32	.35	.33	.36	.33	.36
Panel B — 90th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health insurance	–0.054	–0.058	–0.028	–0.011	–0.020	–0.005	–0.012	0.001
	(.010)**	(.011)**	(.022)	(.024)	(.022)	(0.024)	(0.022)	(0.025)
Pension	–.020	–.037	–.012	–.010	–.012	–0.009	–0.003	–0.001
	(.009)*	(.010)**	(.019)	(.022)	(.019)	(0.022)	(0.020)	(0.022)
Constant	3.714	3.749	3.768	3.637	3.818	3.679	3.857	3.711
	(.056)**	(.065)**	(.137)**	(.158)**	(.139)**	(0.162)**	(0.140)**	(0.162)**
Observations	5,547	5,547	1,090	1,090	1,090	1090	1090	1090
R-squared	.02	.04	.04	.07	.04	0.07	0.05	0.07
Panel C — 10th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health insurance	0.056	0.075	0.006	–0.011	–0.003	–0.020	–0.002	–0.017
	(.021)*	(.022)**	(.039)	(.041)	(.040)	(.042)	(.040)	(.042)
Pension	–.028	.024	.029	.026	.028	.027	.024	.027
	(.028)	(.029)	(.054)	(.055)	(.054)	(.055)	(.055)	(.057)
Constant	1.166	.707	1.122	.959	1.096	.919	1.105	.936
	(.035)**	(.071)**	(.070)**	(.153)**	(.072)**	(.156)**	(.079)**	(.157)**
Observations	5,549	5,549	935	935	935	935	935	935
R-squared	.02	.04	0.03	.08	.03	.08	.03	.08
Industry/occupation dummies	No	Yes	No	Yes	No	Yes	No	Yes
Additional covariates	…	…	Union(0,1)	Union(0,1)	Union(0,1) Full-time(0,1)	Union(0,1) Full-time(0,1)	Union(0,1) Full-time(0,1) Firm size	Union(0,1) Full-time(0,1) Firm size

NOTE: Standard set of covariates include age; number of children under age 6; and indicator variables for black, Hispanic, high school graduate, some college, college graduate, married, previously married, and urban residence. Data set—CPS; Year: 1998; Dependent variable: log (hourly wage). Standard errors in parentheses; * significant at 5 percent; **significant at 1 percent.

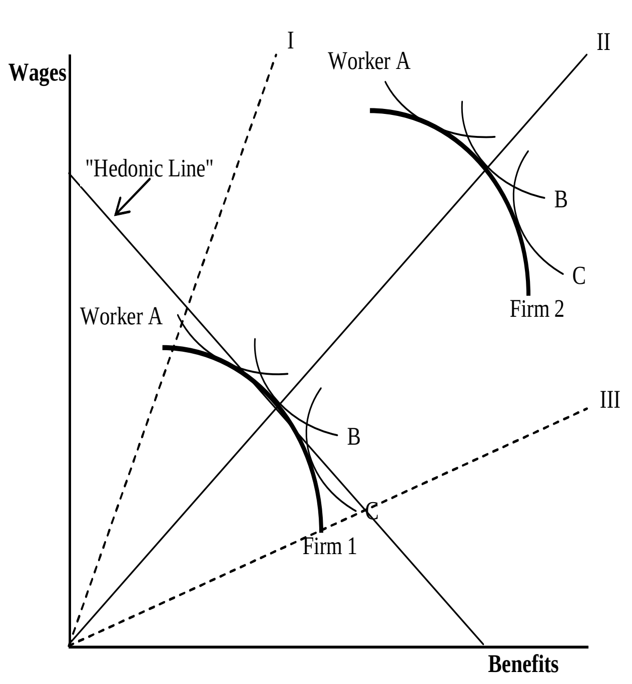
the health insurance coefficient is smaller than for the whole sample and the analogous result in the NLSY.

The estimates for the pension variable are somewhat more muddled. In the NLSY sample, pension coverage remains positive in panel B and is smaller than the analogous results in panel A; the estimates are not statistically significant however. For the CPS, the coefficient on pensions is *negative*, between -0.002 and -0.04, and statistically significant in the first two columns. The overall fit is markedly worse in the subgroup (R^2 around 0.05 for both samples) and the number of observations drops significantly in the NLSY, from around 4,700 in panel A to fewer than 500 observations in panel B.

The results for respondents at the upper end of the distribution suggest that wealthier individuals trade higher wages for less health and life insurance benefits. Intuitively, because individuals at the higher end of the distribution are better able to purchase insurance in the private market, the offset of health and life insurance is logical. The estimates also show that the correlation between wages and life insurance (and pensions) is much higher for workers at the low end of the distribution. When additional covariates are included in the regression (table 3, columns (6)–(10)), the same results also hold for health insurance. The results for pensions conflict each other between data sets in tables 3 and 4—estimates from other models (fixed effects and quantile regressions) are examined in the following sections and tend to be more in concert.[48]

Analogous to the sample for those at the upper end of the distribution, panel C estimates the model at the other extreme—those at and below the 10th percentile of the

Exhibit 1. Basic hedonic diagram



distribution. The results are strikingly different from those at the 90th percentile. In the NLSY specification, the health insurance coefficient is negative for specifications both with and without industry-occupation dummies but is statistically insignificant. In the CPS, which may be somewhat more reliable given the number of observations (5,500 and 935 observations compared to about 470 observations), the coefficient on health insurance is positive and statistically significant in the first three columns, between 0.007 and 0.08. The coefficient on health insurance is negative in the other five columns but of smaller magnitude and does not satisfy statistical significance tests. Regardless of the data set, the coefficient on health insurance is distinctly smaller in this sample than when all workers are included.

Again, the coefficient on the pension variable differs depending on the data set. In the NLSY, the coefficient on pensions is much larger than for the entire sample—between 0.34 and 0.41—and is statistically significant in all 10 columns. In the CPS however, the estimate on pensions is positive in nine of the ten regressions but is not statistically distinguishable from zero in any of the runs. Here also, the coefficients are smaller than their counterparts in panel A, different from the NLSY-based estimates. Also notice however, that there are more missing observations when unions, full-time status, and firm size are added to the CPS than in the NLSY. This may account for some of the differences across the columns in table 4.

Ultimately, the results for the bottom of the distribution are consistent in magnitude but are mixed in terms of statistical significance and sign. Estimates from the CPS imply that people at the bottom of the distribution correlate higher benefits with higher wages and people at the other end of the distribution correlate *lower* benefits with higher wages. The NLSY implies a negative correlation between wages and health and life insurance (though health insurance enters insignificantly) for workers at the lower end of the wage distribution but those workers also exhibit a positive correlation between wages and pensions. The differences not only raise interesting conceptual and distributional issues but also issues of estimation. The possibility of unobserved worker heterogeneity is one concern with these models. In the next section, the analysis is extended by estimating a fixed effects model, exploiting the panel nature of the NLSY.

Fixed effects. The ordinary least squares models, estimated earlier, neglect some important behavior of workers in the labor market. Specifically, ordinary least squares ignores the effects of unobserved heterogeneity among workers and thus in this section, a fixed effects model is estimated using the NLSY sample from 1990–98. (See table 5.)[49]

The entire sample from 1990–98 is pooled, resulting in around 30,000 observations. The coefficients on the variables of interest are similar in sign and statistical significance to the ordinary least squares results in table 3, but are smaller in magnitude. Again, the coefficients are all positive and regressions that include industry and occupation dummies result in slightly smaller coefficients on the variables of interest.

When the sample is segmented and the model is estimated for the 10th and 90th percentiles, the resulting estimates are not consistently statistically significant and the sign on health and life insurance is negative in most of the specifications. For the lower end of the distribution, life insurance enters negatively and is statistically significant in regressions that include the additional covariates. Pensions enter positively and are statistically significant in the other half; the regressions *without* the additional covariate set. For the 90th percentile, the model fit is poor with almost no coefficients entering statistically significantly, although the sign matches those for the 10th percentile. Overall, accounting for unobserved worker heterogeneity does not help identify the hedonic effect but does reinforce the results found in the previous tables, although the magnitude is much smaller.

Quantile regressions. It is apparent from the results in the first three tables that the relationship between wages and benefits are contingent upon the selected sample. Estimates for workers at the top of the earnings distribution suggest a tendency for workers to forego higher wages in lieu of more pension benefits, whereas workers at the bottom of the

Table 5. Regression results for NLSY via fixed effects

	Equation									
Panel A–Whole sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	0.049	0.044	0.055	0.044	0.033	0.014	0.033	0.014	0.031	0.013
	(.005)**	(.006)**	(.005)**	(.006)**	(.009)**	(–.010)	(.009)**	(–.010)	(0.009)**	(–0.010)
Life insurance	(.208)	(.190)	(.210)	(.193)	(.197)	(.139)	(.174)	(.114)	.172	.114
	(.072)**	(.076)*	(.072)**	(.076)*	(.091)*	(–.094)	(–.091)	(–.094)	(–.091)	(–.095)
Pension	(.050)	(.050)	(.050)	(.050)	(.046)	(.042)	(.046)	(.041)	.046	.041
	(.006)**	(.006)**	(.006)**	(.006)**	(.008)**	(.009)**	(.008)**	(.009)**	(.008)**	(.009)**
Total days off	…	…	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	.000	.000
	…	…	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	.000	.000
Constant	(2.206)	(2.283)	(2.192)	(2.268)	(2.473)	(2.567)	(2.408)	(2.504)	2.393	2.492
	(.085)**	(.087)**	(.086)**	(.088)**	(.157)**	(.160)**	(.158)**	(.161)**	(.159)**	(.161)**
Observations	32,015	32,015	32,015	32,015	14,134	14,134	14,134	14,134	14,134	14,134
R-squared	.05	0.05	0.05	0.05	0.05	0.06	0.05	0.06	0.05	0.06
Panel B– 90th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	–0.028	–0.037	–0.028	–0.037	0.079	0.050	0.094	0.067	0.096	0.073
	(.023)	(.026)	(.023)	(.026)	(.060)	(.069)	(.060)	(.070)	(.060)	(.070)
Life insurance	–0.011	–.063	–.012	–.065	–.006	–.089	–.069	–.126	–.064	–.109
	(.177)	(.194)	(.177)	(.193)	(.344)	(.375)	(.344)	(.375)	(.345)	(.375)
Pension	.035	.033	.035	.033	.038	.020	.035	.020	.036	.020
	(.016)*	(.017)	(.016)*	(.017)	(.033)	(.035)	(.033)	(.035)	(.033)	(.035)
Total days off	…	…	–.000	–.000	–.000	–.000	–.000	–.000	.000	–.000
	…	…	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	.000	(.000)
Constant	3.058	3.130	3.063	3.133	.888	1.124	.591	.827	.078	1.040
	(.478)**	(.490)**	(.479)**	(.490)**	(1.267)	(1.297)	(1.270)	(1.303)	(1.277)	(1.311)
Observations	3,213	3,213	3,213	3,213	1,414	1,414	1,414	1,414	1,414	1,414
R-squared	.05	.06	.05	.06	.04	.05	.05	.06	.05	.06
Panel C– 10th percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Health insurance	–0.007	–0.024	–0.007	–0.026	0.012	–0.012	0.013	–0.011	0.009	–0.013
	(.024)	(.025)	(.024)	(.025)	(.035)	(.035)	(.035)	(.035)	(.035)	(.035)
Life insurance	–.696	–.835	–.675	–.810	–2.699	–2.287	–2.679	–2.272	–2.674	–2.265
	(.617)	(.624)	(.618)	(.625)	(.847)**	(.842)**	(.850)**	(.845)**	(.854)**	(.849)**
Pension	.207	.197	.206	.197	.039	–.041	.040	–.040	.043	–.037
	(.046)**	(.048)**	(.046)**	(.048)**	(.070)	(.074)	(.070)	(.074)	(.070)	(.075)
Total days off	…	…	.000	.000	.000	.000	.000	.000	.000	.000
	…	…	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Constant	1.124	1.314	1.094	1.277	1.539	1.239	1.545	1.239	1.534	1.232
	(.239)**	(.263)**	(.242)**	(.266)**	(.461)**	(.499)*	(.462)**	(.500)*	(.464)**	(.501)*
Observations	3,203	3,203	3,203	3,203	1,440	1,440	1,440	1,440	1440	1440
R-squared	.02	.04	.02	.04	.06	.16	.06	.16	.06	.16
Industry/occupation dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Additional covariates	…	…	…	…	Union(0,1)	Union(0,1)	Union(0,1) Full-time(0,1)	Union(0,1) Full-time(0,1)	Union(0,1) Full-time(0,1) Firm size	Union(0,1) Full-time(0,1) Firm size

Standard set of covariates include age; number of children under 6; and indicator variables for black, hispanic, high school graduate, some college, college graduate, married, previously married, urban residence and actual experience (weeks). Data set: NLSY; years: 1990–98; Dependent variable: log (hourly wage). Standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent.

distribution take jobs that have benefits in addition to their regular wages. In this section, quantile regressions are performed on the 10th and 90th percentiles in order to fit median regression analysis on the two parts of the distribution. (The results from the quantile regression model are available upon request to the author.)

The coefficients from the quantile regressions are uniformly positive and generally statistically significant. In the CPS, workers at the 10th percentile respond more strongly to health insurance in terms of wages than those at the 90th percentile.[50] The coefficient estimates for workers below the 10th percentile range between 0.20 and 0.29 while the estimates for the 90th percentile are consistently smaller, between 0.09 and 0.11. The coefficients on pensions demonstrate the differences between the two parts of the distribution: For workers below the 10th percentile, the estimate on pensions range between 0.08 and 0.14, larger than the estimates for the 90th percentile, which are between 0.07 and 0.12. Thus, workers at the bottom part of the distribution respond more strongly to wage changes than do those at the top of the distribution. In the NLSY, the estimates on health insurance are uniformly larger for the 10th percentile than the 90th percentile, mirroring the results from the CPS data set. For pensions however, the estimates for the 10th percentile are larger than their 90th percentile counterparts about half of the time. This suggests that workers at the low end of the distribution respond more strongly in terms of wages for health insurance but that workers at the high end of the distribution may respond more strongly with respect to pensions.

In summary, the results from the quantile regressions generally confirm the ordinary least squares estimates. The effect of health insurance is larger for the bottom tenth of the distribution in both data sets but the effect of pension benefits is somewhat mixed. The evidence in this section continues to be indicative of the "good jobs, bad jobs" story, although the use of the quantile regression method was useful to gain insight into the different relationships between different parts of the wage distribution.

As previously noted, results for additional subgroups including men and women; those above the minimum wage plus 25 cents and plus 50 cents; and several "at-risk" groups, are omitted from the current analysis. The issues specific to these groups are important for two reasons. First, different groups may respond differently to work incentives. For example, do women have less (more) benefits and do they pay more (less) implicitly because of how pay packages work? Second, low-wage groups may reveal an interesting relationship between wages and benefits because their wages are bounded below by the minimum wage. In results not reported, an instrumental variable model was estimated by using State-level minimum wages.[51] Separating the worker population into subgroups provides an opportunity to explore different instruments and to track different behaviors and is a viable area for future work.

Conclusion

Compensating wage differentials are an important aspect of the labor market, yet empirical estimation of these differentials lags behind theory. To date, no researcher has convincingly estimated a hedonic model although many have produced mixed results. The estimates in this article point more to correlations between wages and benefits than to tradeoffs. The analysis did accomplish two main tasks. First, the data employed in the analysis are seldom used by researchers due to confidentiality restrictions. And second, the estimates suggest compensating differentials for subgroups of the population, namely the (weakly) 10th and 90th percentiles but tell a stronger story about the positive correlation between wages and benefits and how they differ at various points in the wage distribution.

The estimates imply that individuals at the top of the wage distribution sometimes earn more than three times as much in certain benefits than those at the bottom of the distribution. Workers above the 90th percentile take slightly more of their compensation in the form of wages than do people below the 10th percentile. Workers at the top of the distribution also take (proportionately) less in insurance and required benefits but roughly the same in pension benefits than do workers at the lower tail of the wage distribution.

In all, the implications of this study are threefold. First, the detailed data from the ECI did generate improvements of point estimates in that statistical significance was achieved in most models. Second, the two different household-level data sets, for the most part, confirmed one another; another check for consistency. And third, point estimates were generally larger for workers at the low end of the wage distribution although estimates for workers above the 90th percentile suggest a compensating differential for health insurance and wages.

The implications of the distribution of benefits on the (gross) wage distribution are important for policy and labor market considerations. Distributional issues as they relate to nonwage forms of compensation are recently receiving more attention and should continue to be explored as better benefit data and better access to such data become available. □

Notes

ACKNOWLEDGEMENT: This article was written while the author was a graduate student at Syracuse University. The author would especially like to thank Brooks Pierce and John Ruser for help using the Bureau of Labor Statistics ECI data. He also thanks William Horrace, Thomas Kniesner, Jeff Kubik, Solomon Polachek, and Timothy Smeeding for helpful comments and suggestions. This research was conducted by having access to restricted BLS data on-site at BLS. The views expressed here are those of the author and do not necessarily reflect the views of BLS.

[1] Sherwin Rosen, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, January 1974, pp. 34–55.

[2] Gary W. Loveman, and Chris Tilly, "Good Jobs or Bad Jobs: What Does the Evidence Say?" *New England Economic Review*, (January/February, 1988), pp. 46–65.

[3] William Wiatrowski, "The National Compensation Survey: Compensation Statistics for the 21st Century," *Compensation and Working Conditions* (Bureau of Labor Statistics, Winter 2000), pp. 5–14.

[4] Employee Benefit Research Institute, "Compensation Costs in Private Industry March 1987 to March 2001," *Facts from EBRI*, August 2001, on the Internet at: **http://www.ebri.org/facts/0801fact.htm**.

[5] Masanori Hashimoto, "Fringe Benefits and Employment," in W. T. Alpert and S. A. Woodbury, eds., *Employee Benefits and labor markets in Canada and the United States* (W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan, 2000), pp. 229–62.

[6] B. K. Atrostic, "Comment," in J. E. Triplett, ed., *The Measurement of Labor Cost,* NBER Studies in Income and Wealth, vol. 48 (Chicago, University of Chicago Press, 1981), pp. 389–94; Arleen Leibowitz, "Fringe Benefits in Employee Compensation," in J.E. Triplett, ed., *The Measurement of Labor Cost.* NBER Studies in Income and Wealth, vol. 48 (Chicago, University of Chicago Press, 1981), pp. 371–89; and Timothy Smeeding, "The Size Distribution of Wage and Nonwage Compensation: Employer Cost versus Employee Value," in J. E. Triplett, ed., *The Measurement of Labor Cost*, NBER Studies in Income and Wealth, vol. 48 (Chicago, University of Chicago Press, 1981), pp. 237–86.

[7] For example, see Stephen Woodbury, "Substitution between Wage and Nonwage Benefits," *The American Economic Review,* March 1983, pp. 166–82 and Edward Montgomery, Kathryn Shaw, and Mary Ellen Benedict, "Pensions and Wages: An Hedonic Price Theory Approach," *International Economic Review,* vol. 33 no. 1, 1992, pp. 111–28.

[8] Charles Brown, "Equalizing Differences in the Labor Market," *The Quarterly Journal of Economics,* February 1980, pp. 113–34.

[9] Jonathan Gruber, and Alan B. Krueger, "The Incidence Mandated Employer-Provided Insurance: Lessons from Workers Compensation Insurance," in D. Bradford, ed., *Tax Policy and the Economy*, vol. 5 (Cambridge, MA, MIT Press, 1991), pp. 111–44.

[10] Jonathan Gruber, "The Incidence of Mandated Maternity Benefits," *American Economic Review,* vol. 84 no. 3, 1994, pp. 622–41.

[11] Joseph G. Altonji, and Christina H. Paxson, "Labor Supply Preferences, Hours Constraints, and Hours-wage Trade-offs," *Journal of Labor Economics,* April 1988, pp. 254–76.

[12] Brooks Pierce, "Compensation Inequality," *Quarterly Journal of Economics,* November 2001, pp. 1493–1525; Pierce Brooks, "Compensation Inequality," *Bureau of Labor Statistics Working Paper 323* (Bureau of Labor Statistics, June 1999); and Craig A. Olson, "Do Workers Accept Lower Wages in Exchange for Health Benefits?" *Journal of Labor Economics,* vol. 20, no. 2 (part 2), 2002, pp. S91–S114.

[13] Pierce "Compensation Inequality," 2001; Pierce, "Compensation Inequality," 1999.

[14] Olson, "Do Workers Accept Lower Wages in Exchange for Health Benefits?" 2002. Olson also estimates the same models using data from the April 1993 Consumer Price Index Fringe Benefit Supplement and again finds negative, though statistically insignificant, coefficients in the range of –0.290 to –0.030.

[15] See also, Thomas C. Buchmueller, John DiNardo, and Robert G. Valletta, "Union Effects on Health Insurance Provision and Coverage in the United States," *Industrial and Labor Relations Review,* July 2002, pp. 610–27.

[16] The earlier studies are: Arleen Leibowitz, "Fringe Benefits in Employee Compensation, 1981;" B. K. Atrostic, "Alternative pay measures and labor market differentials," *BLS Working Paper 127* (Bureau of Labor Statistics, 1983); Woodbury, "Substitution between Wage and Nonwage Benefits, 1983;" and Brown, "Equalizing Differences in the Labor Market, 1980." The later studies are: Gruber, "The Incidence of Mandated Maternity Benefits," 1994; Gruber, and Krueger, "The Incidence Mandated Employer-Provided Insurance, 1991; and Altonji, and Paxson, "Labor Supply Preferences," 1988.

[17] Pierce, "Compensation Inequality," 2001; Pierce, "Compensation Inequality," 1999; and William J. Carrington, Kristin McCue, and Brooks Pierce, "Nondiscrimination Rules and the Distribution of Fringe Benefits," *Journal of Labor Economics,* 2002, vol. 20 no. 2, part 2, pp. S5–S33.

[18] The BLS staff pointed out that the distinction between defined benefit and defined contribution plans in the ECI survey is sometimes vague and thus calculating average values by categories may be erroneous. Thus, for the regressions in this study, the pension variable refers to all pension plans including defined benefit and defined contribution plans.

[19] For more on the details of the ECI data set, see John W. Ruser, "The Employment Cost Index: What is it?" *Monthly Labor Review,* September 2001, pp. 3–20.

[20] Bureau of Labor Statistics. "National Employment, Hours, and Earnings from the Current Employment Statistics Survey," Standard Industrial Classification code based, on the Internet at: **http://www.bls.gov/ces** (extracted July 2002). Additionally, the BLS uses certain imputation procedures to address sample attrition over time. A set of these imputed values was dropped from the sample used in this study. The author would especially like to thanks Brooks Pierce for calling attention to both of these issues.

The construction of the consistent weights modifies the industry weights in the ECI:

$$\hat{w}_{yi} = \frac{w_{qi} \times Employment_{yi}}{\sum_q \sum_i w_{qi}}$$

where w = industry weight in the ECI; q = quarter; i = industry; y = year; and $Employment_{yi}$ = external employment count in industry i in year y.

[21] Due to the confidentiality of the ECI data set, industry and occupation cell sizes were restricted based on the number of observations in each. In addition, because some agriculture and government employees are not included in the ECI, they were eliminated from the NLSY and CPS samples. This leaves nine occupation categories (executive, administrative and managerial; sales; administrative support; precision production, craft and repair; machine operators, assemblers and inspectors; transportation and material moving; handlers, equipment cleaners, helpers, and laborers; service except private household; and professional, technical and specialty) and seven industry categories (mining; construction; manufacturing; transportation; wholesale and retail trade; finances, insurance and real estate; and services and public administration).

[22] Bureau of Labor Statistics "Employer Costs for Employee Compensation: Historical Listing, 1986–2001," on the Internet at: **ftp://ftp.bls.gov/pub/special.requests/ocwc/ect/ECECHIST.PDF** (September, 2001).

[23] Smeeding, "The Size Distribution of Wage and Nonwage Compensation," 1981.

[24] The NLSY does not ask whether a respondent has leave or not; instead it asks how many days are provided to the respondent. In this study, the author chose to use the total number of days off from the NLSY (vacation plus sick) rather than use the imputed dollar values from the ECI because actual days off is probably a better measure. The total number of days off variable was top-coded at 260 because that is the greatest number of days an individual working five days a week, 52 weeks a year could use. The difference between using the NLSY days-off variable and the ECI dollar variable was not very significant.

[25] Daniel S. Hammermesh, "LEEping into the future of labor economics: the research potential of linking employer and employee data," *Labour Economics*, March 1999, pp. 25–42.

[26] John M. Abowd, and Francis Kramarz, "Econometric analyses of linked employer-employee data," *Labour Economics*, March, 1999, pp. 53–76.

[27] Robert F. Elliott and Robert Sandy, "Adam Smith may have been right after all: A new approach to the analysis of compensating differentials," *Economics Letters*, April 1998, pp. 127–31.

[28] Greg J. Duncan and Daniel H. Hill, "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data," *Journal of Labor Economics*, October 1985, pp. 508–32; and Wesley Mellow and Hal Sider, "Accuracy of Response in Labor Market Surveys: Evidence and Implications," *Journal of Labor Economics*, October 1983, pp. 331–44.

For an early look at some of these issues, see also R. Smith, and R. Ehrenberg, "Estimating Wage-Fringe Trade-Offs: Some Data Problems," in J.E. Triplett, ed., *The Measurement of Labor Cost*, vol. 48 (Chicago, University of Chicago Press, NBER Studies in Income and Wealth, 1981), pp. 347–70.

[29] For a comprehensive review of this literature, see John Bound, Charles Brown, and Nancy Mathiowetz, "Measurement Error in Survey Data," in James J. Heckman and Edward Leamer eds., *Handbook of Econometrics, Volume 5* (Amsterdam, North-Holland, 2001), pp. 3707–45.

[30] For an extended discussion, see Smeeding, "The Size Distribution of Wage and Nonwage Compensation,"1981.

[31] The health insurance question in the CPS is slightly different than the question in the NLSY; it asks whether the respondent was covered by a health plan provided by their employer. This implies that individuals answer having already accepted (or rejected) health insurance offers. The questions for pensions however, are similar in the two household samples.

[32] The estimates use the hourly wage reported in the NLSY and CPS, not the ECI. This approach is preferred because the focus is on the tradeoffs made by workers, not employers.

[33] Actual experience is only included in the NLSY sample and is calculated by summing the number of weeks worked for each year of the survey from 1979 to the year of interest. If, in any given year, a respondent was a student and the data for the number of weeks is missing, weeks worked are then set equal to zero in that year. Additional missing data are then dropped from the sample. Values for 1995 (and 1997) are set equal to the average of 1994 and 1996 (and 1996 and 1998).

[34] Estimates reported use the NLSY and CPS sample weights; using the ECI sample weights makes little difference on the results in most models.

[35] Moshe Buchnisky, "Recent Advances in Quantile RegressionModels: A Practical Guideline for Empirical Research," *The Journal of Human Resources*, Winter 1998, pp. 88–126; and Roger Koenker, and Kevin F. Hallock, "Quantile Regression," *Journal of Economic Perspectives*, Fall 2001, pp. 143–56. The author thanks the BLS reviewer for making this analogy.

[36] Olson, "Do Workers Accept Lower Wages?" 2002.

[37] Pierce, "Compensation Inequality," 1999.

[38] See the dissertation by Jonathan A. Schwabish, "Three Essays in Inequality," unpublished doctoral dissertation, Syracuse University, Syracuse, New York, 2003 for an explanation of the results for these groups.

[39] As discussed in the literature section, the studies by Gruber "The Incidence of Mandated Maternity Benefits," 1994 and Gruber and Krueger, "The Incidence Mandated Employer-Provided Insurance," 1991 use variation in State mandated benefit coverage as instruments to correct for endogeneity introduced by their benefit variables. The current study performed similar exercises with the CPS data using variation in State minimum wages as instruments for the wage. The instruments proved to be weak however, and thus are not reported.

[40] Craig Copeland, "Pension Plan Participation Continued to Rise in 2000—What Next?" *EBRI Notes* (Employee Benefit Research Institute, March 2002), pp. 4–7.

[41] Rachel Christensen, "Value of Benefits Constant in a Changing Word: Findings from the 2001 EBRI/MGA Value of Benefits Survey," *EBRI Notes* (Employee Benefit Research Institute, March 2002), pp. 1–3.

[42] Note that the mean age in the NLSY will differ by approximately 1 year in each survey year because of the panel nature. The CPS, by contrast, will basically have the same mean age in each survey. Thus, the comparison in 1998 is the closest in terms of ages.

[43] As mentioned earlier in the endnote section, the estimates use the hourly wage reported in the NLSY and CPS, not the ECI. This approach is preferred because the focus is on the tradeoffs made by workers, not employers.

[44] In regressions where the benefit variables enter individually, coefficients are markedly larger in magnitude—they are not reported in this study because the more comprehensive models are more informative and perform just as well.

[45] Christensen, "Value of Benefits," 2002.

[46] Loveman and Tilly, "Good Jobs or Bad Jobs," 1988, pp. 46–65.

[47] Note that the percentile points in the remaining tables were calculated in the household surveys, not from the ECI data.

[48] Recall that the NLSY has a compressed age distribution, whereas the CPS includes persons 18 to 64 years of age. This may account for some of these differences.

[49] A fixed effects model solely on the ECI data was also considered. However, sample attrition issues within the ECI as well as confidentiality issues led this to be too difficult a task.

[50] Tables containing the results for the 10th and 90th percentiles for both data sets are available upon request to the author and can be found in Schwabish, "Three Essays in Inequality," 2003.

[51] The results from the instrumental variables model are not a main focus of this study for several reasons, including mixed sign and lack of statistical significance on the instrumental variable (the State minimum wage); poor overall fit; and mixed results in the second stage.