APPENDIX C

How Robust Are Hedonic Wage Estimates of the Price of Risk: The Final Report

by Dan A. Black, Jose Galdo, and Liqun Liu

How Robust Are Hedonic Wage Estimates of the Price of Risk?

The Final Report

by

Dan A. Black
Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
danblack@maxwell.syr.edu

Jose Galdo
Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
jcgaldo@maxwell.syr.edu

Liqun Liu
Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
lliu04@maxwell.syr.edu

June 2003

[R 82943001]

Project Officer
Nathalie Simon, Economist
Research and Program Support Division
National Center for Environmental Economics
Office of Policy Economics and Innovation
Ariel Rios Building, Mail Code 1809T
1200 Pennsylvania Avenue, NW
Washington, DC 20460

We thank Bill Evans, Tom Kniesner, Nathalie Simon, and Laura Taylor for helpful comments. We also thank Seng Eun Choi and Rebecca Walker for research assistance.

Table of Contents:

I. Introduction

At least since Adam Smith's *Wealth of Nations*, economists have recognized that workers require compensation to accept the risk of death or dismemberment on the job. While this wage premium provides employers with incentives to reduce the risk on the job, the calculus of the marketplace allows workers and employers to trade the costs of reducing workplace risk against the benefits associated with the reduction.

This calculus, when applied to large numbers of workers, allows a researcher to calculate the value of a statistical life, or the wage reduction associated with reducing the expected number of deaths by one worker. As this value represents the amount of wages that workers are willing to forgo to reduce risk, the value of a statistical life appears to be a useful tool for evaluating individuals' willingness to pay for reductions in risk in other areas. Indeed, it is a measure of the price of risk. The Environmental Protection Agency (EPA) often considers regulations that both impose costs on industry and reduce the deaths from environmental contamination. While the costs may often be calculated with a great deal of accuracy, the problem for policymakers is to value the corresponding benefits. The price of risk appears to be a useful tool for such evaluations.

When basing policy on estimates of the price of risk, the precision and accuracy of the estimates become of utmost importance. Yet, Viscusi (1993), in his review of labor market studies of the value of life, reports that the majority of the estimates are in the $3 to $7 million range [in December 1990 dollars, p. 1930], and this range excludes studies that Viscusi felt were flawed. While this represents over a 133 percent variation, Viscusi correctly notes that much of the variation should be expected, as the studies used different methodologies and different samples. Workers may differ in their attitudes toward risk, and the mixes of workers in these various studies differ substantially. His review, however, leaves unanswered how much of this variation results from differences in the sample of workers, measures of job risk, and the specification of the estimating equation.

In this report, we use three data sets to estimate the price of risk: the Outgoing Rotation Groups of the Current Population Survey, the March Annual Demographic Supplement of the Current Population Survey, and the National Longitudinal Survey of Youth (1979). Labor economists frequently use these three data sets to estimate wage equations. We match these data to two sources of job risk data: the Bureau of Labor Statistics estimates from their Survey of Working Conditions and the National Institute of Occupational Safety and Health estimates from their National Traumatic Occupational Fatality survey. We then use these data to estimate the price of risk. Among our major findings are:

- First, and foremost, the estimates are quite unstable. Small changes in the specification of covariates or the risk measured used result in large variations in the estimated price of risk. Many of the estimates indicate that the price of risk is negative, which is contrary to the theoretical framework used;
- The instability of the estimates does not appear to be the result of the misspecification of the functional form of the regression function. More flexible functional forms of the regression function provide similar estimates to similarly specified OLS equations, and the more flexible functional forms also produced unstable estimates when changing the covariates or risk measures. In our view, therefore, the instability is not the result of equation misspecification;
- We find overwhelming evidence that the job risk measures contain much measurement error. Moreover, the measurement error is nonclassical, as it is correlated with covariates that are usually put into earnings of wage equations. Estimates that do not account for such measurement error may be highly attenuated, which would cause an understatement of the value of risk reduction.

There may, of course, be other biases that offset the attenuation that usually occurs with severe measurement error;

- We find some evidence that job risk is correlated with characteristics not commonly available in labor economic data sets. Using data from the NLSY, we find that job risk varies inversely with Armed Forces Qualification Test scores and varies positively with illegal drug use. This suggests that job risk may be correlated with regression error, rendering OLS estimates inconsistent.

Collectively, these findings lead us to have severe doubts about the usefulness of existing estimates to guide public policy. These estimates are so highly sensitive to the risk measure used and the specification of the wage equation that the selection of any particular value of the price of risk seems arbitrary.

The rest of this report is structured as follows. The next section contains all of the estimates. The first subsection presents the basic estimation results, while the second subsection examines the sensitivity of estimates to assumptions about functional form. The third subsection examines the extent and impact of measurement error; the fourth subsection looks at possible correlations between the job risk measures and the regression error. In the final section, we offer some brief concluding comments.

II. Estimating the Price of Risk

A. Baseline Estimates

The basic notion of hedonic models of risk is to ask the question: "All else the same, how much must I compensate a worker to accept an increase to the risk in the worker's job?" The model is conceptually very simple. Consider a worker who faces a risk of death on-the-job, denoted r^*, and is paid a wage w. Given a von Neumann-Morgenstern expected utility function, we may characterize the expected utility function of the worker as:

$$E(U) = (1 - r^*)U(w \mid X) + r^* D$$

where D is the disutility of death, $U(w \mid X)$, is the utility of earning a wage w, and X is a vector of all other factors that affect the worker's utility. Von Neumann-Morgenstern are unique up to an affine transformation so that we may add or subtract a constant of $-D$.[1] This allows us to rewrite the above equation as:

$$E(U) = (1 - r^*)u(w \mid X) \ ,$$

where $u(w \mid X) = U(w \mid X) - D$.

If we hold constant the expected level of utility, we may ask "How much must we compensate the worker in order for the worker to accept an increase in job risk?" The answer to that question in differential form is simply:

$$dE(u) = 0 = -u(w \mid X) + \frac{\partial u(w \mid X)}{\partial w} \frac{\partial w}{\partial r^*},$$

or

$$\frac{\partial w}{\partial r^*} = \frac{u(w \mid X)}{\partial u(w \mid X) / \partial w} = \varphi(w, X) > 0 \ .$$

In general, the slope of the wage-risk locus depends on the base level of utility, the vector of covariates X, and the level of wages. In general, theory offers no guidance as how to specify the $\varphi(w, X)$ function. A particularly convenient form is $\varphi(w, X) = \gamma w$ so that we have:

$$\frac{\partial w / \partial r^*}{w} = \gamma \ .$$

This form is particularly convenient because it arises from the well known Mincerian equation

$$ln(w) = X\beta + r^*\gamma + \varepsilon_i$$

[1] The variable y is an affine transformation of x if $y = a + bx$, where $b > 0$ which is required to maintain the preference ordering.

While this is the starting point for virtually all hedonic labor market studies, it is worth noting that it is based on a very strong assumption that $\varphi(w, X) = \gamma w$.

Thus, the starting point for our analysis is a wage equation of the form:

$$ln(w_i) = X_i\beta + r_i^*\gamma + \varepsilon_i \tag{1}$$

where $ln(w_i)$ is the natural logarithm of the *ith* worker's wage, r_i^* is the measure of risk (potentially a vector), X_i is a vector covariate, (β, γ) are coefficients to be estimated, and ε_i is the error term of the regression. We assume that $Cov(X_i, \varepsilon_i) = 0$ and $Cov(r_i^*, \varepsilon_i) = 0$ (so that the risk measures and other covariates are exogenous). This form of the wage equation is what Viscusi (1993) calls the "basic approach in the literature" and admits a natural interpretation for γ as the price of risk.

We begin our analysis by estimating equation (1) with Ordinary Least Squares (OLS). We report estimates with three different samples of workers: the Annual Demographic Survey (or March) Current Population Survey (March CPS), the Outgoing Rotation Groups of the Current Population Survey (ORG CPS), and the 1979 panel of the National Longitudinal Survey of Youths. The March CPS and the ORG CPS are nationally representative samples of workers, while the NLSY is a rich panel data set of individuals aged 14 to 21 in 1979. We describe the data sets in more detail in the Appendix B.

There are two major sources of government-reported job risk: the Bureau of Labor Statistics (BLS) estimates from their Survey of Working Conditions and the National Institute of Occupational Safety and Health (NIOSH) estimates from their National Traumatic Occupational Fatality survey. The NIOSH data provide one-digit occupation or industry mortality rates by state. While the BLS data contain counts of deaths by three-digit occupation or industry codes, they do not provide any regional variation. The risk measures have their own distinct costs and benefits. The BLS data, available from 1995 to 2000, contain very detailed measures of the annual number of deaths, but the data suppression procedure requires at least 5 deaths in a cell before the number of deaths is reported.[2] Thus, there are a substantial number of missing values for these data. The use of annual data may be subject to a great deal of sampling error associated with the annual fluctuation in the number of deaths. Moreover, these data only provide the counts of the number of deaths in each industry or occupation. To make these a rate, it is necessary for researchers to estimate the number of workers in an industry or occupation. To estimate these numbers, we use both the March CPS data and the ORG CPS data, but undoubtedly this estimation generates measurement error in our risk measures. Finally, the BLS data assume that job risk is a constant across the country, which clearly is not the case.

The NIOSH data provide fatality rates by one-digit industry or occupation codes by state. It reports 5-year averages from 1981 to 1985, 1986 to 1990, and from 1991 to 1995. Thus, it does not require the researcher to estimate the number of workers in an industry or occupation cell, allows the job risks measure to vary by state, and smoothes much of the sampling variation by using the 5-year average. The use of the 5-year average and the coarser one-digit industry or occupation codes by state reduces, but does not eliminate, the problem of missing values because of data suppression. On the other hand, these data treat police officers and dental assistants as having the same job risk as both are in the one-digit "service worker" occupation. The use of 5-year averages, while smoothing the sampling variation, may miss important time series variation.

Combining the two data sources, we estimate our models for the ORG CPS and March CPS using the BLS data for the years 1995 to 2000.[3] For the NIOSH data, we use the ORG CPS and March CPS for the years 1985 to 1995 and the NLSY data for the years 1986 to 1993. For both the BLS data and the NIOSH data, we estimate separate equations for each sex and use both the occupation and industry risk measure. We limit the sample to workers who are aged 25 to

[2] The BLS suppresses reports with fewer than five deaths.
[3] We choose not to estimate the NLSY data with the BLS data because of the relatively small sample sizes of the NLSY and the irregular sampling of the NLSY in the late 1990s.

60 inclusive. Because theory provides little guidance as to the exact specification of the X_i we provide several different specification of the vector. For the CPS data sets, our basic specification includes a quartic in the worker's age, a vector of dummy variables that control for education level, union status, marital status, and a vector of variables of controls for the worker's race and ethnicity. We also add additional controls for the worker's firm size, which is not consistently available for the ORG CPS data so it is not used for this data set. In subsequent specifications, we include controls for state of residence, and then one-digit industry and occupation. For the NLSY, the basic specification also includes experience, tenure, and worker's Armed Forces Qualification Test (AFQT) scores.

While all of these covariates have been used in countless wage equations, our use of the state and one-digit industry and occupation variables warrants some discussion. Unfortunately, for the BLS data our measures of job risk are ultimately assigned by the worker's three-digit industry or occupation. There is a long history in labor economics of adding industry and occupation variables as covariates in wage equations. Indeed, in the 1980s and early 1990s there was a very visible strand of literature (e.g., Krueger and Summers 1988) that examined whether the payment of above equilibrium wages could be detected from these controls. This literature has documented very large earnings differentials across industries and occupations that were not explained by the type of covariates we include in our analysis. While a portion of these wage differentials may reflect differences in job risk, undoubtedly a substantial portion of these wage differentials reflect other unobserved features of jobs. Using industry and occupation covariates, therefore, sweeps out a good deal of unobserved heterogeneity. The costs of such controls, however, are that they remove much of the variation in our job risk measures. Of course, a similar problem arises in our use of the NIOSH data when we include both state and one-digit industry and occupation variables. In our view, the inclusions of such controls is crucial to control for such unobserved heterogeneity.

While we use both industry-based and occupation-based risk measures, we would be remiss if we did not comment on the relative merits of the two risk measures. At first glance, the use of the industry measure seems inappropriate. After all, this measure assigns the same job risk to a secretary in the coal mining industry as to the coal miner, clearly overstating the secretary's level of job risk and understating the coal miner's job risk. In contrast, the use of occupational risk would combine the job risk of a secretary in the coal mining industry with a secretary in the insurance industry, presumably a pair with a much more homogeneous job risk. Yet, this line of argument is deceiving. Mellow and Sider (1983) document that industry is measured more accurately than occupation, a point to which we return below.

Finally, for Table 1, we report four specifications. The first specification contains a set of basic controls that are found in most wage equations. While we try to make the various data sets have similar specifications, these basic controls differ across data sets because of difference in the covariates that are available (e.g., the ORG CPS data do not consistently contain a firm size measure and the NLSY data contain the AFQT score of recipients). The second specification contains a vector of dummy variables that control for the worker's state of residence. This allows us to control for some differences in the cost of living, state taxes, worker's compensation programs, and other state programs that may confound the estimates. In the third specification, we add either industry or occupation controls depending on whether we are using the industry or occupation risk measure. When using the industry risk measure, we add one-digit occupation controls, and when using the occupation risk measure, we add one-digit industry controls. (We seek to identify the source of parameter instability with this specification.) The last specification contains both industry and occupation controls.

With this rather long introduction, in Panels 1 through 10 of Table 1A, we present our estimates of equation (1) for men; t-statistics are reported in parentheses. Panel 1 presents the estimates using the March CPS matched to the NIOSH industry risk measure. Each of the estimates provides a positive estimate of the impact of job risk on wages, although the estimates are highly variable. The inclusion of state of residence controls more than doubles the estimated price of risk, the inclusion of the one-digit occupation controls reduces the estimate by more than 40 percent, and the inclusion of the industry controls reduces the coefficient by more than 60 percent.

We may use the estimates to construct a value of a statistical life, although we report the values only when they are positive as the theory predicts. Assume that a worker has wage and salary earnings of $35,000, about the mean in 1994 for men ($34,137). We may take the estimates in Panel 1 of Table 1A, divide them by 100,000 (to normalize them back to deaths per 100,000 workers so that we obtain an estimate of the price of risk, γ), and then calculate value of a statistical life using the formula $\$35,000 \times (e^{\gamma} - 1) \times 100,000$ where γ again is the estimated price of risk. In Panel 1, the estimated value of a statistical life varies from about $3.7 million to $16.4 million. Half of the estimated values of a statistical life are within Viscusi's range , which is $3.3 to $7.8 million when inflation adjusted to 1994, but the estimates vary by over 440 percent.

The range of estimates from Panel 1, of course, gives us pause about the quality of the estimates. Unfortunately, there is strong reason to suspect that these estimates may be capturing variations in working conditions other than simple job safety. Literally hundreds of wage studies have used industry dummies as covariates in earnings or wage equations. While these dummies undoubtedly capture variation in job risk, they also capture variation in other working conditions that are not measured in most commonly used data sets. To the extent that other working conditions covary with job risk, the estimates in Panel 1 are biased. We interpret the instability of the parameter estimates in Panel 1 as evidence of the measure of job risk being correlated with the regression error, or $cov(r_i^*, \varepsilon_i \mid X) \neq 0$. From a theoretical perspective, this is hardly surprising. Wealthier workers tend to buy safer jobs. One suspects that they also purchase cleaner jobs, with better hours and nicer offices. Given that we observe only a modest fraction of the characteristics that affect the worker's productivity and only one measure of job-related amenities (job risk), it is hardly surprisingly that we have biased estimates.

Determining the nature of the bias, however, is more difficult. To see why, suppose that the true model of wage determination is:

$$ln(w_i) = X_i\beta + r_i^*\gamma + Z_i b + \varepsilon_i' \qquad (1')$$

where Z_i is a vector of job disamenities and $b > 0$ by assumption. Because wealthier workers prefer jobs with fewer disamenities, or job amenities are a normal good, we expect that $\text{cov}(r_i^*, \varepsilon_i') < 0$ and $\text{cov}(Z_i, \varepsilon_i') < 0$ so that more productive workers have jobs with fewer disamenities. If we mistakenly estimate equation (1) rather than equation (1'), it is straightforward to show that:

$$\hat{\gamma}^{OLS} = \frac{\gamma\,\text{var}(r_i^* \mid X_i) + \text{cov}(r_i^*, Z_i \mid X_i)b + \text{cov}(r_i^*, \varepsilon_i' \mid X_i)}{\text{var}(r_i^* \mid X_i)}.$$

As we expect that job disamenities covary positively ($\text{cov}(r_i^*, Z_i \mid X_i) > 0$) and as $b > 0$, the second term in the numerator causes us to overestimate the price of risk, but as $\text{cov}(r_i^*, \varepsilon_i' \mid X_i) < 0$, the third term in the numerator tends to causes us to underestimate the price of risk. Which bias dominates is, of course, not known. Obviously, the inclusion of one-digit industry and occupation controls imperfectly controls for some of the missing unobservables.

In Panel 2 of Table 1A, we repeat the exercise for the NIOSH occupation measure of job risk. Contrary to the prediction of the theory, three of the four estimates are negative and significant! Again, the magnitudes of the estimates are quite unstable, again suggesting that job risk covaries with the unobservables (ε_i) in the wage equation. This is quite a disconcerting result. The NIOSH occupation job-risk measure is constructed from the same data sources as the

NIOSH industry job-risk measure. Moreover, one expects that the occupation measure of job risk might be a more accurate measure of job risk than industry-based measures.

In Panels 3 and 4, we estimate equation (1) using the BLS measures of job risk and the March CPS from 1995 to 2000. While the BLS job risk measure uses a much more disaggregated measure of industry or occupation (3-digit as opposed to one-digit) to assign job risk, the BLS measure ignores spatial variation in job risks. The BLS estimates are also based on annual fatalities, which tend to be volatile, especially for relatively small industries or occupations. The BLS-based estimates are quite unstable. For instance, using the industry-based measures, the estimates range from -188 to 239, and the estimates are quite precisely estimated.

Economists have long been concerned about biases in estimates of the return to schooling that results from the sorting of higher ability workers into higher levels of schooling. This "ability bias" could cause economists to severely over-estimate the returns to schooling; see Card (1998, 2001) for excellent reviews. We worried that our estimates suffered from similar ability bias. If low-ability workers were being sorted into dangerous jobs (so that $\operatorname{cov}(r_i^*, \varepsilon_i^{'} \mid X_i) < 0$), our estimates would be biased downward. In an attempt to assess how severe such a problem might be, we took a page from the returns-to-schooling literature and found a data set that contains test scores for respondents. While test scores tend to be an imperfect measure of ability, our hope was that the use of test scores as a covariate might remove much of the ability bias.

Toward that end, we used the National Longitudinal Survey of Youth 1979 data. Because respondents took the Armed Forces Qualification Test (AFQT), we have measures of the respondents' aptitude. The AFQT is a set of 10 tests. We first demeaned test scores, conditioning on the age of the respondents at the time they took the test. We then took the first two principal components of the demeaned test scores as our measures of aptitude and added the test score measures to our set of basic controls. In Panels 5 and 6, we estimate equation (1) with our NLSY sample using data from 1986 to 1993 and the NIOSH job risk measures. To our disappointment, the results are reasonably similar to the results from the corresponding two panels of Table 1A. While there are some modest differences, it is important to note that others have found variation arising from the use of different survey instruments and respondents in different locations. Moreover, the estimates using the NIOSH occupation-based job risk measures remain negative and significant. The three positive estimated values of a statistical life are outside of Viscusi's range of $3.3 to $7.8 million, and two are over $20 million.

Finally, for men we also estimate equation (1) using the Outgoing Rotation Group (ORG) data from the CPS and both the NIOSH and BLS job risk measures. Unfortunately, the ORG data does not contain a consistent measure of firm size, but the data do afford larger samples and have a measure of wages that is superior to the March data. In particular, the March data requires the researcher to impute the value of the wage by dividing earnings last year by usual weekly hours and weeks worked. As a result, the earnings may be from a variety of different jobs last year, not just in the job measured. Moreover, the March data requires workers to accurately recall last years earnings, the usual hours of work, and the number of weeks worked last year. One suspects that each of these questions is subject to considerable measurement error. In contrast, the ORG ask about the wages on the major job last week, which appears to provide much better wage estimates. When looking at hourly wage, there are mass points at the minimum wage, whole dollar amounts, and numbers divisible by five and ten as one might expect from hourly wages.

The results are reported in Panels 7 through 10 of Table 1A. Unfortunately, the results are just as disappointing as the results from the other data sets. Of the 16 different estimates, nine are negative and statistically significant. Of the remaining seven estimates, four are outside of Viscusi's range of $3.3 to $7.8 million.

Historically, the value of a statistical life literature has focused on the job risk of men. While we point out below that there are some sensible reasons for concentrating on men, we did not want to ignore the wage hedonic for female workers. In Table 1B, we replicate our estimates for women. Generally, the qualitative findings are quite similar for men and women. The estimates of the price of risk for women are quite volatile, just as they are for men. We will not test our reader's patience with a complete description of the results in Table 1B, but we will use estimates from Panel 1 to construct value of life estimates for women, ignoring the negative and statistically insignificant result in column 5. Using these estimates, the range of estimates in Panel 1 imply value of life estimates that range from $9.7 to $17.6 million, assuming that we evaluate the women's value of life at an earnings of $35,000 as well.[4] As in the estimates we have for men, there is an extremely wide variation in the price of risk.

While we have demonstrated the large variations in the estimated coefficients for the job risk measure, we have not demonstrated how unusual this variability is in the estimation of wage equations. To illustrate how stable coefficients usually are, in Table 2 we report the estimated returns to a bachelor's degree relative to a high school degree using the same specification as appears in Panel 8 of Tables 1A and 1B. There is a long history in labor economics of estimating such returns to education; see Card (1998) for an excellent review. It is generally recognized that when estimating the returns to schooling, one does not want to condition on the worker's occupation because a better occupation is a part of the way in which schooling raises earnings, but to keep our results comparable with Panel 8, we also provide results that condition on occupation. The estimated coefficients for both men and women are remarkably consistent. Unless one conditions on occupation, the results never differ by more than 10 percent. Even when conditioning on occupation, the coefficients fall by less than 40 percent.

Finally, while we have reported a large number of specifications, we wanted to provide some guidance as to our preferred specification. To us, there is no doubt that the final

[4] In these data, women earn about 60 percent of what men earn. One could, therefore, reduce these estimates by 40 percent to evaluate the value of life at the mean earnings of women.

specification that contains the basic controls, the controls for state of residence, and the industry and occupation controls is superior to the other specifications. For reasons that we explained above, we think the ORG CPS data provide a better wage measure than the March CPS data. Similarly, while we think the richness of the NLSY data is valuable, the limited sample sizes and the limited age range, we again favor the ORG data. Finally, while the BLS data's three-digit disaggregation is very appealing, the use of a single national number looses much variation that is extremely useful for identification so we favor the use of the NIOSH data, although we remain agnostic as to whether one should use the occupation or industry risk measure.

In summary, the evidence from the estimates presented in Table 1 lead us to draw five conclusions:

- First, and foremost, estimation of equation (1) produces quite unstable estimates of the price of risk. Small changes in the specification of covariates result in large variations in the estimated price of risk;
- Second, estimates from the men's and women's sample appear to provide reasonably similar estimates of the price of risk. As a result, we will concentrate our discussion on the results for men, but we will continue to present the results for women so that the interested reader may compare differences in estimates between the sexes;
- Third, many of our estimates are within generally accepted bounds for estimates of the value of life. For instance, for men 6 of our estimates are within Viscusi's range of $3.3 to $7.8 million. Other estimates are quite similar to other reports in the literature. For instance, our estimate using the NIOSH occupation risk data with the March CPS data with controls for both industry and occupation ($0.6 million) is virtually the same as Kniesner and Leeth's (1991) estimate of $0.7 million using CPS data and the industry NIOSH data;

- Fourth, both industry and occupation controls substantially affect the estimated coefficients, even when using the "opposite" measure of risk. Henceforth, we report only the specifications that have both industry and occupations;
- Fifth, despite the more disaggregated industry and occupation categories that the BLS uses in the construction of their job risk measures, we find no evidence that the BLS data are superior to the NIOSH data. As a result, we will focus our discussion on the NIOSH data because these data afford us a longer time horizon. Again, we will continue to report estimates using the BLS data so that the interested reader may compare the NIOSH-based estimates to the BLS-based estimates.

Thus, we now attempt to find reasons for the instability of these estimates.

Generally, we think there are three reasons to explain the instability of the estimates found in Table 1. First, we may be suffering from a bias resulting from the correlation of the job risk measure and the regression error, or $\mathrm{cov}(r_i^*, \varepsilon_i^{'} \mid X_i) \neq 0$. This would result in standard endogeneity bias. Second, our functional form in equation (1) could be incorrect. The results of Heckman et al. (1998) and Heckman, Ichimura, and Todd (1997, 1998) can be interpreted as finding that the misspecification of the functional form of the conditional mean functions accounted for a great deal of the heterogeneity in estimates of the impact of job training programs. Third, because the measures of job risk are imperfect, the resulting measurement error might result in significant biases.

The concern regarding $\mathrm{cov}(r_i^*, \varepsilon_i^{'} \mid X_i) \neq 0$ proves to be a difficult nut to crack. Unlike most recent advances in applied microeconomics, our estimates of the price of risk have focused on the equilibrium relationship between risk and wages. Many of the recent advances in applied microeconomics have focused on finding "natural experiments" that induce exogenous variation in the variable of interest (job risk in our case) and examine how agents react to that random variation. For instance, Angrist (1990) uses the draft lottery during the Vietnam Era to access the impact of military service on earnings. Because of exogenous variation in military service that the draft lottery induces, Angrist is able to obtain a consistent estimate of the impact of military service on earnings. Similarly, Angrist and Evans (1998) document that variation in the gender composition of a woman's first two children affects the likelihood that the woman will have additional children, with women whose first two children are the same gender being more likely to have an additional child. Because the gender composition of a woman's offspring is uncorrelated with her productivity in the labor market, Angrist and Evans are able to obtain estimates of the impact of children on labor supply. See Angrist and Krueger (1999) for an excellent discussion of the recent advances in applied microeconomics.

Unfortunately, such truly exogenous variation appears very difficult to obtain in labor market studies of the price of risk. One might hope that government policies may be used for such variation. For instance, Evans, Farrelly, and Montgomery (1999) use the imposition of workplace bans on smoking, which implicitly increase the cost of smoking, to examine the impact of such mandates on smoking. Government safety regulations might appear to provide similar types of natural experiments. However, while the Occupational Safety and Health Administration (OSHA) has implemented several policies to reduce workplace risk, Viscusi (1981) notes that in the first six years of OSHA's existence workplace accident rates declined less than 16 percent, and most of this decline was the result of the changing industry structure. Indeed, Kniesner and Leeth's (1995) review of the literature suggests that OHSA has had no measurable impact on worker safety. Thus, it appears unlikely that government safety regulations will provide sufficient variation to measurably affect the price of risk.

Similarly, one might hope that technological advances would provide natural experiments. Indeed, there are technological advances that have had large impacts on workplace safety (e.g., the introduction of long-wall coal mining), but these technological advances also affect the demand for labor and the skill mix of labor in the industry or occupation. This makes it extremely difficult to distinguish the impact of demand changes on wages from the impact of reduction in risk on wages. Hence, technological advances do not appear to be legitimate natural experiments. In the near future, we intend to explore an idea that Bill Evans suggested to us: the use of regional variation in the improvement of auto travel safety resulting from the introduction of air bags and improved enforcement of drunk driving laws. Unfortunately, results from this project are months away and so we must rely on our use of equilibrium variation in job risk. We can, however, address the issues of the proper functional form and the role of measurement error. In the next two subsections, we address these two issues in some detail.

B. The Role of Functional Form and Support

Again, consider the standard hedonic wage equation:

$$ln(w_i) = X_i\beta + r_i^*\gamma + \varepsilon_i \,. \tag{1}$$

The equation makes three strong assumptions that may do violence to the data. First, it assumes that the researcher knows the appropriate vector of covariates (X_i, r_i^*). Second, it assumes the coefficients (β, γ) are constants, rather than functions or random vectors. Thus, the impact of risk on log wages is the same for a 45-year-old black male accountant as for a 27-year-old white male high school graduate working in the oil fields of Texas. Third, it assumes a log-linear relationship between the wage and the covariates. As Angrist and Krueger (1999) emphasize, the use of OLS estimation may provide very misleading estimates if these assumptions are incorrect.

For the ORG CPS and March CPS data, our sample size is sufficiently large and our covariates X_i are of sufficiently low dimension that we may employ the cell-matching or frequency estimator. We may for the *kth* cell consider the estimation of the equation:

$$ln(w_{ik}) = \alpha_k + r_i^* \gamma + \varepsilon_{ik} \tag{2}$$

where w_{ik} is the *ith* worker's wage, α_k is a constant to be estimated, and ε_{ik} is the regression error. For the NLSY, the data are a bit sparser so we recoded the test scores variables into deciles, transformed the experience and tenure variables, originally measured in months of work, into years, and combined some of the post-high school variables.

The estimation of equation (2) allows for a complete set of interactions for every variable included in X_i and imposes no linearity restriction. Hence, we can interpret equation (2) as being estimates of :

$$ln(w_i) = g(X_i) + r_i^* \gamma + \varepsilon_i . \tag{2'}$$

In this model, the function $g(X_i)$ is a nuisance parameter. We continue to make the assumption that the job risk parameter, r_i^*, enters the equation linearly and that it remains additively separable, in the logarithmic form, from the other covariates, X_i. Finally, if equation (1) represents the true form of the model, the estimation of equation (2) will provide consistent estimates of the price of risk, γ.

In Panels 1 through 8 of Table 3A, we report the results of the estimation of equation (2) for men and in Panels 1 through 8 of Table 3B we report the results for women as well. Taken as a whole, the volatility of the results in Table 3 are remarkably consistent with corresponding volatility of the results in Table 1. The estimates continue to be highly volatile when changing the specification of the vector of covariates, X_i. There is perhaps some evidence that the estimates are somewhat larger when using the flexible functional form. In our view, this suggests that instability of the estimates is not a result of the log-linear specification of the vector of covariates, X_i.

Of course, even this specification assumes that the relationship between the logarithm of wages and job risk is linear. Viscusi (1981) reports substantial variation in estimates of the value of life by quartiles of the distribution of job risk. For instance, the implied value of life for workers in the first quartile of fatality risk is \$5 million while the implied value for workers in the fourth quartile is only \$2.8 million.

To examine whether there are substantial nonlinearities in the wage-risk locus, we next divide jobs into risk deciles and then estimate the equation:

$$\ln(w_i) = X_i\beta + r_{di}\gamma + \varepsilon_i \tag{3}$$

where r_{di} is the risk decile of the ith worker and γ is now a vector of coefficients to be estimated. The use of these discrete cells allows us to trace out any nonlinearity. In Panels 1 through 10 of Table 4A, we report the results of the estimation of equation (3) for men and in Panel 1 through 10 of Table 4B for women.

The results show a highly nonlinear relationship between the wages and risk levels. For instance, focusing on the last column of Panel 8 of Table 4A, initially there is an increase in wages as risk levels with the coefficient on the second to the fourth deciles remaining positive. The estimated coefficient then remain approximately zero as job risk increases. Nor are these results unusual. A quick review of all the panels in Tables 3A and 3B shows a consistent lack of a monotonic relationship between job risk and wages. Of course, interpreting this nonmonotonicity is difficult. It may be the result of the misspecification of equation (1); perhaps significant interactions between the job risk measures and the covariates, X_i, have not been modeled. Alternatively, this nonmonotonicity may be indicative of a covariance between the job risk measures and the regression error, or $\operatorname{cov}(r_i^*, \varepsilon_i \mid X_i) \neq 0$.

If the nonmonotonicity is the result of the exclusion of relevant interaction terms, we may rely on nonparametric regressions to produce estimates of the relationship between job risk and wages without making any functional form assumptions. Exploiting recent advances in applied microeconometrics, we next estimate the wage-risk locus nonparametrically. We use the propensity score matching estimator of Rosenbaum and Rubin (1983); see Heckman, Ichimura, and Todd (1997, 1998) and Smith and Todd (2000) for a discussion of propensity score estimates and examples of their use. Until recently, propensity score matching has been limited to cases in which the variable of interest was binary. For case job risk, this would require the division of jobs into risky and safe classifications, a much too restrictive formulation in our view. Fortunately, Imbens (1999) and Lechner (2000) have shown that propensity score matching extends to finite numbers of alternatives.

We divide jobs into K risk categories (for quintiles, $K = 5$). Let the wage of the *ith* worker in the *jth* risk category, Y_{ij}, be given by:

$$Y_{ij} = g_j(X_i) + \varepsilon_{ij} \qquad j = 1, 2, \ldots K \tag{4}$$

where X_i is a vector of characteristics that determines earnings and ε_{ij} is again the error term. The function $g_j(\cdot)$ is an unknown function that determines wages. We may define the price of risk:

$$p_{ijk}(X_i) = Y_{ij} - Y_{ik}, \tag{5}$$

which is the cost per hour of moving the worker from the *jth* risk class to the *kth* risk class.

The fundamental problem is that we observe either Y_{ij} or Y_{ik} but never observe both. We estimate the "missing" wage using nonparametric methods, a nearest neighbor estimator. After estimating the missing wages, the price of risk is:

$$\hat{p}_{ijk}(X_i) = Y_{ij} - \hat{Y}_{ik} \tag{6}$$

where $\hat{Y}_{ik}$ is the estimated missing wage. Given these individual prices of risk, the average price of risk for moving from the *jth* to the *kth* risk category may be calculated as

$$\hat{p}_{jk} = \frac{\sum_{i=1}^{N_j} \hat{p}_{ijk}(X_i)}{N_j}. \tag{7}$$

The nonparametric estimation of the price of risk avoids making any assumptions about the functional form of the $g_j(\cdot)$ and allows the price of risk to vary across individuals. Thus, moving from the first to the second decile of risk may have a different price than moving from fourth to the fifth decile.

As Heckman et al. (1998) emphasize, there is an added benefit to nonparametric estimation: it forces researchers to confront the "support problem," which is most easily seen when considering the cell-matching estimator but similarly exists for the propensity score estimator. For instance, suppose researchers wish to know the price of moving a 55-year-old white male with a bachelor's degree from the first decile of risk to the tenth decile of risk. The researchers may well find there are no 55-year-old white males with a bachelor's degree in the tenth decile of risk. The data simply will not allow researchers to calculate that price of risk because nonparametric estimation relies on matching workers across the various categories of risk. Of course, with parametric regression such as those in equation (1), we can extrapolate outside the range of the data. Extrapolation outside the range of the data, however, is simply identification by functional form assumption.

While "propensity score" matching may appear quite abstract, the intuition behind the estimator is really quite simple. We wish to compare someone in a "risky job" with someone in a "safe job," but clearly we want to make sure the individuals are comparable. One could try to match workers exactly on the X vector, but for many workers there may not be anyone with an exact match. The basic idea of nonparametric regression is to find someone "similar" without requiring an exact match. The genius of the Rosenbaum and Rubin's (1983) result is that, if the appropriate assumptions hold, we may match workers only on the estimated probability that they are in a risky job. This greatly simplifies the matching and can lead to faster rates of convergence. See Smith (2000) for an excellent non-technical introduction.

In Panels 1 through 10 of Tables 5A and 5B, we provide nearest neighbor estimates for our nonparametric approach. Nearest neighbor estimates simply match people in the *jth* group to the person who is "closest" to them in the *kth* group. Thus, as we are matching on the person's propensity score, for person i in the *jth* group we observe their propensity score, s_{ij}^{0}, find the person in the *kth* group whose propensity score is the closest to s_{ij}^{0}, and use this person's wage as the missing counterfactual. To guarantee that the match is of reasonable quality, we apply a caliper of 0.01, so if the difference between the treatment group observation and nearest neighbor from the comparison group exceeds 0.01, the observation is discarded. We match on the most exhaustive set of covariates for each of the three data sets that we use: the CPS March Demographic Supplement, the CPS ORG data, and the NLSY. The fifth quantile is the group with the highest risk jobs and the first quantile is the group (with the lowest risk jobs).

Several features of the estimates are of interest. First and foremost, the estimates vary as much as the OLS regression estimates vary. Many of the estimates of compensation necessary to take on added risk are negative. Even when the estimates are positive, the estimates do not monotonically increase in risk as we move to higher risk quintiles. Looking across data sets, we see large movements in the estimated comparison group wage. For instance, focusing on the first column of Table 5A, the wages of the first quantile (the comparison group) vary from \$10.99 to \$8.19. This difference shows the differences in who gets placed in the quantile and who gets matched to an observation in the treatment group (the higher risk group) across the differing data sets.

Collectively, the results in Tables 3 through 5 provide some evidence that the variability of the parametric estimates do not appear to be the result of any restrictive assumptions imposed by the parametric representation of equation. The problem of the parameter variability, therefore, would seem to rest elsewhere. Of course, this does not mean that the log-linear form of the wage equation traditionally used in this literature is correct. Rather, it simply documents that other problems with the estimation that appear in parametric and nonparametric estimates. In the next section, we explore the role of measurement error in the estimation of the price of risk and find evidence that this measurement error may well be the source of at least some of the volatility.

C. The Role of Measurement Error

The quality of estimates is necessarily limited by the quality of measurement. No matter how sophisticated the theoretical and econometric models, data of poor quality may still provide estimates of poor quality. In the next section, we suggest why the data from the BLS and NIOSH, while providing extremely accurate measures of the aggregate job risk in the United States, may not provide accurate estimates of the job risk of those workers in our sample.

There are essentially three problems in our measurement of job risk. First, because we divide workers into industries or occupations–some of which are quite small–we may have considerable sampling variation within these industry and occupation cells. Both the BLS and NIOSH data recognize this problem and suppress data when the number of fatalities is too low, but this inherent sampling variation creates measurement error. Second, within occupations, there may be a great deal of heterogeneity in the actual job risk and the assignment of that job risk may be extremely nonrandom. For instance, employers may assign male and older clerks at convenience stores evening and late night hours when the risk of holdup–and injury during the robbery–are particularly high and assign female and younger clerks daytime hours. Because we only measure the aggregate job risk of convenience stores clerks, however, this would result in our overestimating the job risk of young and female clerks and underestimating the job risk of older and male clerks. Finally, because we need to assign workers to an industry or occupation, the quality of our measurement is limited to the quality of the data on industry and occupation assignment. The best available evidence (e.g,, Mellow and Sider, 1983) is that industry and occupation–especially at the three-digit level–are not measured accurately.

1. Documenting the Magnitude of the Measurement Error

If the researcher could measure (X_i, r_i^*) perfectly, Ordinary Least Squares (OLS) estimation of equation (1) would provide consistent and efficient estimates of the parameters (β,γ), assuming the functional form of the conditional mean function was properly specified and the covariates (X_i, r_i^*) are orthogonal to the error term. Unfortunately, there are numerous reasons to suggest that the measure of job risk (r_i^*) is mismeasured and perhaps mismeasured badly. First, government fatality reports are inherently an estimate of job risk: they are realizations of a random variable. For instance, suppose there are N_k workers in the *kth* industry (or occupation) category, and each of these workers is subjected to a risk, r_k^*. Unfortunately for the researcher, the government's tally of deaths in the *kth* category is not exactly equal to the expected number of deaths, $r_k^* N_k$. Rather, the government's tally is equal to the random variable D_k. Using the random variable D_k, the researcher constructs an estimate of r_k^* as $r_k = D_k / N_k$. While $E(r_k) = r_k^*$, it is almost certain that $r_k \neq r_k^*$. Thus, let $r_k = r_k^* + \eta_k$, where η_k is the measurement error associated with the variable r_k.

A simple example illustrates the problem. Suppose there are 400,000 workers in a particular industry or occupation (a relatively large 3-digit industry or occupation, or a large one-digit industry or occupation at the state level) and every worker faces a 5 in 100,000 chance of an on-the-job death, a little larger than the national average in 1995, which was 4.3 per 100,000 workers. If we randomly simulate the number of deaths for five years, we obtain 23, 24, 26, 24, and 11 deaths for a mean death rate of 5.4 per 100,000, very close to the true mean of 5. The standard deviation, however, 1.4, or about 30 percent of the true mean. If we consider an industry or occupation of only 100,000 workers, we get deaths of 3, 3, 9, 8, and 6 for a mean of 5.8, which again is not too far off the true mean of 5. The standard deviation of this sample, however, is 3.4 which over 60 percent of the true mean.

This argument can be formalized. If we assume that the risk of death in an industry or occupation, r_k^*, is the same for all individuals, the number of deaths in an industry or occupation is distributed binomially, with mean $n_k r_k^*$ (where n_k is the number of workers in the industry or occupation) and variance $n_k\, r_k^*(1-r_k^*)$. This implies that our estimate of the death rate has a mean of r_k^* (so that the estimates are unbiased) and has variance $r_k^*(1-r_k^*)/n_k$. A commonly used measure of the precision of the estimate is the coefficient of variation, which is simply the ratio of the standard deviation to the mean. In our case, the coefficient of variation is for the death rate in an industry or occupation is simply $(1-r_k^*)^{1/2} / (r_k^*)^{1/2}(n_k)^{1/2}$. The smaller r_k^* the larger the coefficient of variation, and, of course, the probability of an on-the-job fatality is very small even for extremely dangerous occupations. Thus, there is intrinsically a lot of sampling variation when trying to estimate rare events such as on-the-job fatalities.

Past studies have indicated that job risk differs by firm size, region, and worker characteristics. Thus, when we make the further substitution for the *ith* worker's risk (who is in the *kth* industry/occupation class) that $r_i^* = r_k$, we are undoubtedly introducing measurement error. Thus, let:

$$r_k = r_i^* + \nu_{ik} \tag{8}$$

where ν_{ik} represents the measurement error associated with using r_k as a proxy for r_i^*.

The measurement error undoubtedly attenuates the estimates of the coefficient, γ. Indeed Hausman (2001) terms this the "iron law of econometrics." From an empirical standpoint the relevant question is "How severe is attenuation bias that results from the measurement error v_{ik}?" Fortunately, we have up to four reports on the level of job risk that we may use to determine the extent of the measurement error.

To see why, consider two measures of job risk:

$$r_{1i} = r_i^* + v_{1i} \tag{9}$$

$$r_{2i} = r_i^* + v_{2i} \tag{10}$$

where r_i^* is the true measure job risk, v_{ji} is the measurement error associated with the jth measure of job risk, and r_{ji} is the jth observed measure of job risk. The covariance of the two measures is simply:

$$Cov(r_{1i}, r_{2i}) = Var(r_i^*) + Cov(v_{1i}, r_i^*) + Cov(v_{2i}, r_i^*) + Cov(v_{1i}, v_{2i}) \tag{11}$$

and the variances of the two measure are:

$$Var(r_{1i}) = Var(r_i^*) + 2\,Cov(v_{1i}, r_i^*) + Var(v_{1i}) \tag{12}$$

$$Var(r_{2i}) = Var(r_i^*) + 2\,Cov(v_{2i}, r_i^*) + Var(v_{2i}) \tag{13}$$

which provides us with six unknown parameters and three equations. It also demonstrates why it is essentially impossible to make much progress on the problem in this form: the system is underidentified.

Fortunately, we may follow Griliches (1986) and assume that our measurement error is classical. That is, we may assume $Cov(v_{1i}, r_i^*) = Cov(v_{2i}, r_i^*) = Cov(v_{1i}, v_{2i}) = 0$, which reduces our three-equation system to:

$$Cov(r_{1i}, r_{2i}) = Var(r_i^*) \tag{14}$$

$$Var(r_{1i}) = Var(r_i^*) + Var(v_{1i}) \tag{15}$$

$$Var(r_{2i}) = Var(r_i^*) + Var(v_{2i}) . \tag{16}$$

Of course, with additional covariates, we need to make the additional assumptions that $Cov(v_{1i}, X_i) = 0$ and $Cov(v_{2i}, X_i) = 0$ so that the measurement errors are uncorrelated with covariates. Because we have up to four measures of job risk, the classic errors-in-variables model has empirical content: the covariance of any two measures of risk should have precisely the same covariance as any other two measures.

At this point, it is useful to present a convenient decomposition for OLS regressions. Yule (1907) has shown that the estimation of equation (1) with OLS is equivalent to the following. Estimate:

$$\ln(w_i) = X_i\, b + \varepsilon_i' \tag{17}$$

and recover the residual from the equation, which we denote $\ln(w_i)'$. Then, estimate:

$$r_i = X_i\ \delta + u_i' \tag{18}$$

and recover the residual from the equation, which we denote r_i'. We may then estimate the equation:

$$\ln(w_i)' = r_i'\gamma + \varepsilon_i'' \tag{19}$$

The estimation of equation (19) will yield precisely the same estimate of γ as the OLS of γ from equation (1).

Exploiting Yule's decomposition, our three equations system of covariances would simply become:

$$Cov(r_{1i}, r_{2i} \mid X_i) = Var(r_i^* \mid X_i) \tag{20}$$

$$Var(r_{1i} \mid X_i) = Var(r_i^* \mid X) + Var(v_{1i} \mid X) = Var(r_i^* \mid X) + Var(v_{1i}) \tag{21}$$

$$Var(r_{2i} \mid X_i) = Var(r_i^* \mid X_i) + Var(v_{2i} \mid X_i) = Var(r_i^* \mid X_i) + Var(v_{2i}) \tag{22}$$

where $Var(v_{1i} \mid X) = Var(v_{1i})$ and $Var(v_{2i} \mid X_i) = Var(v_{2i})$ by the assumptions that $Cov(v_{1i}, X_i) = 0$ and $Cov(v_{2i}, X_i) = 0$. As $Var(r_i^*) \geq Var(r_i^* \mid X)$, the addition of covariates must always reduce the signal-to-noise ratio or $Var(r_i^* \mid X) / \left(Var(r_i^* \mid X) + Var(v_{ji})\right) < Var(r_i^*) / \left(Var(r_i^*) + Var(v_{ji})\right)$. Because job risk varies with the observable characteristics, X, conditioning on the observable characteristics removes the variation in the risk measures that are correlated with X. If the measurement error is uncorrelated with the measurement error, as in the case of classical measurement error model, then the inclusion of covariates leaves the variance of the measurement error unaffected and reduces the variation in actual job risk, r_i^*. Thus, the addition of covariates increases the attenuation bias associated with the measurement error: the inclusion of covariates, while necessary to control for the heterogeneity in workers and jobs, removes information

In Table 6, we present the correlation and Yulized residual correlations for the various job risk measures. We use data from the 1995 CPS, including both the ORG and March Supplement. The results are quite depressing. The raw correlation before conditioning on any covariates is modest at best, ranging from 0.53 to 0.30. As the correlation differs in magnitude, we have at least some evidence that the measurement error is not classical. When we condition on the full set of covariates, the correlations range from 0.41 to 0.02! The inclusion of both the state controls and the industry and occupation controls in particular reduces the correlation among the various measures. (Notice that in absence of measurement error the correlations should be 1.)

In Table 7, we produce the full range of Yulized residual covariance, which in turn may be used to construct any estimate desired. The OLS estimates of the price of risk are simply the ratio of the risk measure covariance with the wage measure, divided by the variance of the risk measure. Similarly, we may form any IV estimate desired by dividing the covariance of a risk measure and the wage measure by the covariance of the two risk measures. Thus, the ratio of the variance of the risk measure to its covariance with the other risk measure provides a measure of the magnitude of the attenuation bias resulting from measurement error in the job risk measures. The ratios of the variance-to-covariance are large, particularly for the BLS occupation measure, suggesting that OLS estimates in Table 1 are substantially attenuated. Of course, the *negative* measures of job risk are substantially attenuated as well. Moreover, notice that the covariances of the logarithm of wages and the various job risk measures are quite different and often of the opposite sign, which forces us to conclude that the measurement error is nonclassical.

A second manner in which we might find evidence of measurement error is to compare the estimates of fatality rates for subgroups of the population that we derive using our data with the fatality rates for these subgroups that NIOSH compiles. Because of the aggregation bias that exists in our data, we may see substantial differences between our estimates and the NIOSH Census of occupational deaths. Our data assign the mean risk rate to everyone within an occupation regardless of age, race, and sex, but that may be incorrect. Even if we limit ourselves to a homogeneous occupation in the same industry, there may be substantial differences in risk, say, between a police officer in Washington, DC, and one in Larned, Kansas (population, 4,236). Yet, we suspect that the policeman in Larned is more likely to be white than the officer in Washington. Similarly, even within the city of Washington, we suspect that younger and male police officers may be given somewhat more dangerous assignments than more senior and female officers. Such aggregation bias may add substantially to the measurement error.

The actual situation is much more complicated. Industry and occupation are very poorly measured, even in carefully collected data sets such as the CPS. For instance, using a CPS supplement that interviewed both employees and their employers, Mellow and Sider (1983) document that employer and employee agree on three-digit industry codes only 84.1 percent of the time. Even for the broader one-digit industry codes, the rate of agreement is only 92.3

percent. The situation for occupation codes is even worse. Employee and employer agree only 57.6 percent of the time about the three-digit code and only 81.0 percent of the time for one-digit codes. Thus, there is a substantial degree of measurement error in the industry and occupation measures. Mellow and Sider document that for the sample in which both firm and worker agree on the three-digit industry code, the estimated price of risk for non-fatal accidents is 50 percent higher than for the sample as a whole.

In Table 8, we depict aggregate fatality rates from the NIOSH census for workers by race, sex, age categories, industry, and occupation. These compilations provide the level of job risk by each of these categories. We then match our various measures of job risk to the CPS and attempt to replicate the NIOSH census by aggregating the CPS data over the observed characteristics. The results indicate a substantial amount of error in our measures of job risk. The job risk that black Americans face is substantially underestimated in our measure and the job risk that white Americans face is overestimated. Similarly, we substantially overestimate the job risk that women face and underestimate significantly the job risk that men face. Generally, our estimates of the job risk of blacks are understated, although the BLS occupation job risk overstates the fatality rate for almost every other group. This can result because the job risk measures are suppressed if an occupation or industry cell has too few deaths to be disclosed. We overestimate the risk to younger workers and underestimate the risk to older workers. Finally, there are some large discrepancies across industry and occupation divisions as well.

In Table 9, we exploit the time series of data available from the NIOSH. We compare the aggregate fatality rate that NIOSH calculates each year to the implied fatality rates when we match the NIOSH industry and occupation fatalities rates to the ORG CPS. Two features of the data are very distinctive. First, the matched data always understate the level of job risk. Second, because the NIOSH data are updated every five years, the matched data are too volatile in the years that the new estimates are released and do not show a sufficient time series variation to match the aggregate time series trend.

The form of this measurement error is particularly troublesome. We find the measurement error is correlated with various covariates that are typically included in wage or earnings equations. Given the state of the measurement error literature, this correlation makes it impossible to recover unbiased parameter estimates of the price of risk. Moreover, given that we find convincing evidence that the measurement error is correlated with observable factors that affect wages (the covariates), we expect that the measurement error is probably correlated with unobservable factors that affect wages (the regression error). Such complex correlations among the job risk, the covariates, and the regression error make it impossible to obtain consistent estimates of the price of risk. We simply need better data.

2. Attempts to Correct or Mitigate the Measurement Error

In this section we explore three different methods to help mitigate the impact of the measurement error. We pursue these corrections with trepidation. Given the form of the measurement error that we described in the previous section, the estimates we present in this section are inconsistent.

We begin by considering a simple means of increasing the signal-to-noise ratio in data with multiple reports. Consider:

$$\overline{r_i} = \frac{\sum_{j=1}^{4} r_{ji}}{4} \tag{23}$$

or the simple average of the four job risk measures. This does not reduce the mean bias of the measurement error because we continue to use all of the available measures of job risk. Thus, if we systematically understate the job risk of African Americans in each of our four risk measures, the average of these risk measures will continue to understate the job risk of African Americans.

It will in general reduce the variance associated with the measurement error, unless error terms are perfectly correlated. In Tables 10A and C, we estimate our equation using data from both the ORG and March data for men, while in Tables 10B and D we repeat the exercise for women. The estimates are again highly sensitive to the specification of the equation and are often negative.

In a recent paper, Lubotsky and Wittenberg (2001) considered how to construct an index from the multiple reports. If we let b_j for $j \in \{1,2,3,4\}$ be the OLS estimates from entering all the noise measures of job risk in the same, we simply construct:

$$b^* = \sum_{i=1}^{4} \frac{\text{cov}(r_i, \ln(w))}{\text{cov}(r_1, \ln(w))} b_i \tag{24}$$

where $\text{cov}(r_i, \ln(w))$ is the covariance between the ith risk measure and the natural logarithm of wages. Lubotsky and Wittenberg show that b^* is a lower bound on the price of risk when the measurement error is uncorrelated with the regression error ($\text{cov}(\nu_i, \varepsilon) = 0$) and the measurement error is uncorrelated with the true measure of job risk ($\text{cov}(r^*, \nu_i) = 0$). Importantly, the Lubotsky and Wittenberg result allows the various measurement errors to be correlated across job risk measures, or $\text{cov}(\nu_j, \nu_i) \neq 0$. While we have used the first job risk measure to normalize the estimate, this is arbitrary. Lubotsky and Wittenberg term the estimator in equation (24) the "post hoc" estimator. The post hoc estimator is essentially a weighted average of the individual coefficients where the weights are determined by the relative strength of the covariance between the individual risk measure and the dependent variable.

In Table 11A and 11C, we present the "post hoc" estimator for men using 1995 for both the ORG and March data; Tables 11B and 11D repeats the exercise for women. In each case, the "post hoc" estimator provides negative estimates of the price of risk. Thus, the "post hoc" estimator does nothing to correct the fundamental problem with these estimates.

Finally, we consider the classic solution to a measurement error problem: Instrumental Variables estimation. As we have multiple reports, we may easily find an instrument if we are willing to assume, in addition to the assumption of Lubotsky and Wittenberg, that $\text{cov}(\nu_j, \nu_i) = 0$. In Table 12, we produce our IV estimates of the price of risk using the NIOSH job risk data and the March and ORG data for men. Our selection of these data is not random. We picked a set of data where the covariances between the Yulized residual of wage and the Yulized residual of the job risk measures are positive for both measure of job risk. We use the most extensive set of controls. The resulting IV estimates are quite variable. The IV estimates range from being the same magnitude as the OLS estimates to larger by a factor of 10.

Thus, the IV estimates illustrate the potential attenuation that may plague the OLS estimates. As Black, Berger, and Scott (2000) and Kane, Rouse, and Staiger (1999) emphasize, however, these estimates may be biased away from zero if there is a negative covariance between the measurement error and the true value of job risk. Moreover, Shogren and Stamland (2002) document that the failure to account heterogeneity in the skill to avoid accidents may cause us to overestimate the price of risk. Thus, while the presence of measurement error that we have documented and Hausman et al.'s (1991) "iron law of econometrics" suggest that current estimates of the price of risk are severely attenuated, other biases may cause us to overestimate the price of risk.

Thus, we conclude with the same caveat that we began with: We believe that the estimates reported in this section are inconsistent. Thus, one should not use these estimates for setting policy. It is important, however, to note that existing estimates may suffer from substantial attenuation bias to the extent that they have not controlled for measurement error in job risk measures.

D. The Role of Unobservables

One common concern in the use of observational data is the possibility that unobservable factors might confound estimates. In the context of equation (1), the concern is that the risk measure might be correlated with other factors that affect earnings. In this section, we exploit the richness of the NLSY data to see if such concerns are justified.

Our approach is to use the 1984 survey of the NLSY to examine whether job risk is correlated with other factors observed in the NLSY data that are not commonly observed in other data sources. In Table 13, we depict the correlation between job risk and education, illegal drug use, and AFQT score. Education is, of course, generally observed in most data sets that labor economists use and it is highly correlated with our measures of drug use. Few data sets, however, contain information on respondents' illegal drug use, but one suspects that illegal drug use may be correlated with unobservables that affect wages. We find that illegal drug users do take more risk than those who do not (admit) take illegal drugs. Similarly, we find that those who score higher in the Armed Forces Qualification Test have safer jobs.

Our findings are similar to Hersch and Viscusi (2001). They document that cigarette smokers assume more job risk than nonsmokers and receive less of a compensating differential for their injuries. Similarly, we find that individuals willing to undertake the risk associated with taking illegal drugs or who are less able to score highly on standardized tests also take on higher job risk. In the case of the AFQT scores, there is little doubt that the correlation between job risk and the omitted variable biases the estimate of the price of risk in models without standardized test scores. Numerous models have documented that standardized test scores are highly correlated with wages. Hence, the failure of the data sets such as the CPS to have these measures and the correlation we document suggest that job risk is an endogenous variable. While the interpretation of the correlation between drug use and job risk is a bit more problematic, it suggests that there are unobservable characteristics that affect the propensity to take drugs and assume job risk, which affect wages.

With panel data, there is another approach that may be used to attempt to control for unobservable factors that might be confounding our estimates of the price of risk. If one is willing to assume that the form of the wage equation is:

$$ln(w_{it}) = X_{ii}\beta + r_{it}^{*}\gamma + \alpha_i + \varepsilon_{it}, \tag{25}$$

we may allow for a correlation between the individual fixed-effect, α_i, and the measure of job risk, r_{it}^{*}. Thus, we are essentially assuming that any unobservable characteristics that affect both job risk and wages are time invariant. While this is clearly a restrictive assumption, if it is true then the use of a fixed-effect model would eliminate the bias from these unobservable factors.

In Tables 14A and 14B, we provide estimates of the price of risk using equation (25) and data from the NLSY for men and women. Unfortunately, if not surprisingly, the results are disappointing. For men, the estimates are generally negative and not significant. For women, the estimates are generally positive (as the theory requires), but the estimates are generally not statistically significant. Collectively, the fixed-effects estimates do not appear to provide credible estimates of the price of risk. In our view, this is hardly surprising. It is generally recognized that fixed-effects models exacerbate the problems of measurement error (see Griliches 1986), and these data contain a great deal of measurement error.

III. Conclusion

The existing estimates of the price are largely based on methodologies similar to those we have used in this study: the application of OLS to log linear wage equations. Several of our estimates appear to be consistent with that literature. Unfortunately, these estimates are not robust. Changes in the specification of the equation or changes in the job risk measure can result in very large changes in the estimated price of risk.

In our attempts to explain why the estimates are so unstable, we find evidence that the functional form of the regression equation has little impact on the estimates. Using more flexible functional forms, we found the estimates were similar to the corresponding log linear regression traditionally used in the literature. In addition, the more flexible methods also produced highly unstable estimates when we changed the covariates or altered the risk measure.

We did find, however, compelling evidence that there was a great deal of measurement error in the various measures of job risk. Because we have multiple measures of job risk, we may look at the correlation among the various measures of job risk. The correlation is seldom above 0.5 and the inclusion of richer sets of covariates lowers those correlations. In addition, there appears to be a systematic bias that is correlated with many of the covariates that labor economists often include in wage equations. The existing measurement error literature provides little guidance in how to correct for such nonclassical measurement error. This finding argues strongly for better efforts in data gathering.

Finally, we find some evidence that the assignment of job risk is correlated with the regression error. Because many standard data sets such as the CPS contain only modest sets of covariates, evidence from the NLSY suggests that job risk is correlated with standardized test scores and illegal drug use, data elements that are not typically available to researchers. Thus, it seems quite likely that job risk measures may be endogenous in many wage equations.

V. Appendix

Appendix A

We initially intended to use the Panel Study of Income Dynamics (PSID) as our panel data for this project. Ultimately, we decided that the National Longitudinal Study of Youth (NLSY) was a better data set because it contained standardized test scores for respondents. The initial results with the PSID, however, were quite similar to the results from the two CPS data sets and the NLSY. In Table A1, we provide estimates for men using specifications similar to Table 1 of the report. Again, the estimates are very sensitive to specification and are often negative.

Table A1. Estimated Price of Risk for Male Workers, 1985-1992 PSID Data and NIOSH Risk Data

A. Industrial Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-200	-273	467
	(-3.31)	(-4.19)	(3.88)
B. Occupation Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-958	-1152	433
	(-17.46)	(-19.19)	(5.63)

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a dummy variable indicating whether the worker is black/white, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 25,971 observations in the men's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses.
Source: Authors' calculations.

Appendix B

1. March CPS Data

The Annual Demographic Survey or March Supplement Survey, sponsored jointly by the U.S Bureau of Labor Statistics and the U.S Census Bureau, is one of the primary sources of annual income. Using a selected random sample of some 50,000 households from all 50 states and the District of Columbia, the Annual Demographic Survey gives a detailed analysis of geographical mobility, education attainment, work experience, annual income and poverty status of persons 15 years old and over.

In our study we constrained the data to individuals between 25 and 60 years old. The dependent variable is the log of hourly wages which is based on labor income from non-farm workers and non-self-employees. We divided the annual labor income by the total number of weeks times the hours per week worked during the year. We did not used allocated or inputting information for earnings, hours or weeks. In addition, we did not give any special treatment to top-coding observations.

We use the standard industrial and occupational classification to build 1-digit occupation (11 major groups) and 1-digit industry (10 major groups) categories. Based on these major categories we merge the March CPS data with the NIOSH risk data for each worker in the sample. In the case of BLS risk data that is given at 3-digit industry and occupational level, we merge it with the March CPS data using the 3-digit level for both industry and occupational categories.

The remaining independent variables used in the regressions are:

- Age: We use a quartic in age;
- Race/Ethnicity: 5 dummy variables for white, blacks, Asian, Hispanic and others;

- Education: 9 dummy variables for the highest completed years of education: less than junior high school, junior high school, some high school, high school, some college, college degree, master degree, professional degree, and Ph.D. degree;
- Marital Status: 4 dummy variables for married, single, widowed, divorced; the excluded
- Firm Size: 6 dummy variables for firms with less than 10 workers, 10-25 workers, 26-99 workers, 100-499 workers, 500-1000 workers, and more than 1000 workers,;
- State fixed effects: 51 dummy variables for each state in U.S, including the District of Columbia. ;
- Industry/Occupation: Dummy variables for 10 industrial categories and 11 occupational ones;

The mean values and standard deviations for the variables used in the regressions are in the Appendix tables.

2. NLYS Data

The NLSY is a nationally representative sample of 12,686 young men and women who were 14-21 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994 and are currently being interviewed on a biennial basis.

In our regression model, the panel dataset is extracted from 1986 to 1993, and the observations are aged from 25 to 60. Then the panel dataset is merged with NIOSH risk data. In the basic control, the independent variable is the log hourly wage, and the independent variables include age, union status, experience, tenure status, marital status, education background, firm size, and test score. The data are constructed as follows:

- *Lwage*: log of average real wage (adjusted by each year's price index) on all jobs held during the year. There are two variables used to construct wages: total income from wages and salary in the past calendar year and number of hours worked in the past calendar year. The wage variable is equal to total wage income divided by total hours worked;

- *Union*: indicates whether any job held during the year was covered by a collective bargaining agreement. The union indicator is equal to one if wages are covered by collective bargaining on any of the five jobs;
- *Experience*: total months the respondent has been employed since age 16;
- *Tenure*: total months the respondent has worked for the current employer;
- *Education*: highest grade or year of education completed;
- *AA (BA) degree:* dummy variable indicating the respondent has a 2-year (4-year) college degree as of each year interview;
- *Marital Status*: indicates by three dummy variables, which are married with spouse present, married without spouse present, never married;
- *Industry and Occupation*: information on industry of primary job (CPS) and all other jobs are available in every year in the sample. We use the standard industrial and occupational classification based on the Census data of the corresponding years to build 1-digit occupation (11 major groups) and 1-digit industry (10 major groups) categories;
- *AFQT*: AFQT (Armed Forces Qualification Test) is a test used by the military to judge whether an applicant is suitable for military service. The test was given to NLSY respondents;
- *Firm Size*: a dummy variable indicating whether the employees of a firm are greater than 1,000 or not.

3. ORG Data

The Outgoing Rotation Groups are of the Current Population Survey (CPS) which is a monthly survey of about 50,000 households. Each household entering the CPS is in the survey for 4 consecutive months, out for 8, and then returned for another four months before leaving the sample permanently. Since 1979 only households in months 4 and 8 have been asked their usual weekly earnings and usual weekly hours. These are the outgoing rotations groups that are put together into a single Outgoing Rotation Group file. Hence, an individual appears only once in any file year, but may reappear in the following year.

In our study we constrained the data to individuals between 25 and 60 years old. The dependent variable is the log of hourly wages which is based on labor income from non-farm workers and non-self-employees. We divided the weekly labor income by the number of hours worked per week. We did not consider allocated or inputting information for earnings or hours. Again, we did not give any special treatment to top-coding observations.

We use the standard industrial and occupational classification to build 1-digit occupation (11 major groups) and 1-digit industry (10 major groups) categories. Based on these major categories we merge the ORG data with the NIOSH risk data for each worker in the sample. In the case of BLS risk data that is given at 3-digit industry and occupational level, we merge it with the March CPS data using the 3-digit level for both industry and occupational categories.
The rest of independent variables used in the regressions are constructed in a similar manner to those from the March Supplement. The mean values and standard deviations for the variables used in the regressions are in the appendix tables.

4. NIOSH Risk Data

NIOSH risk data is constructed by the National Institute for Occupational Safety and Health (NIOSH), the Federal agency responsible for conducting research and making recommendations for the prevention of work-related disease and injury. It collects the death certificates from all 50 states and the District of Columbia in answer to the need for a comprehensive enumeration of workers who sustain a fatal work-related injury. The fatality rates were calculated as deaths per 100,000 workers. Rates were not calculated for categories with less than three fatalities or less than 20,000 employees.

The NIOSH risk containing both industry and occupation one-digit level job risks, which varies across state was merged to March CPS, NLSY and ORG data by state and one-digit industry/occupation categories according to the following scheme: data from 1983-1985 was merged to NIOSH 1985, data from 1986-1990 was merged to NIOSH 1990, and data 1991-2000 was merged to NIOSH 1995.

5. BLS Risk Data

The BLS Risk data is based on the Census of Fatal Occupational Injuries which has been conducted in all 50 States and the District of Columbia by the Bureau of Labor Statistics, as part of the BLS occupational safety and health statistics program. Information about each fatality is obtained by cross-referencing source documents (death certificates, workers' compensation records and reports to federal and state agencies). The Census gives the number of fatalities at 3-digit Industry and Occupational level for each cell with at least 5 deaths per year.

We build the BLS industry and occupational fatality risk rate as the ratio between the number of fatalities in each 3-digit industry and occupational category and the weighted number of workers in each cell. The weighted number of workers in each 3-digit category was extracted from the March CPS Data and ORG Data. Finally the risk rate was expressed as:

$$Risk_i = \frac{\#Fatalities_{ij}}{\#Workers_{ij}} * 100,000$$

for i= industry, occupation.

j=3-digit cells.

Appendix B Table 1. Mean Values of 1993 NLSY Data

Variables	Male	Female
Hourly Real Wage	9.54 (6.39)	7.87 (4.26)
Marital Status:		
Never Married	0.43 (0.50)	0.44 (0.50)
Married Spouse Present	0.44 (0.50)	0.44 (0.50)
Other	0.13 (0.34)	0.12 (0.32)
Job Risk		
Industry Risk	4.76 (5.05)	3.07 (3.70)
Occupation Risk	6.50 (7.72)	4.95 (7.70)
Age	32 (2.21)	32 (2.23)
Background		
Experience	11.30 (5.50)	10.52 (3.15)
Union	0.28 (0.45)	0.25 (0.43)
Tenure	7.46 (20.73)	6.37 (18.78)
Firm Size	0.61 (0.49)	0.59 (0.49)
Education		
AFQT	10.64 (2.66)	9.93 (2.29)
Less high school	0.02 (0.14)	0.01 (0.07)
Some high school	0.09 (0.28)	0.04 (0.20)
High school		
Some college	0.29 (0.45)	0.35 (0.48)
Associate Degree	0.09 (0.29)	0.13 (0.34)
Bachelor Degree	0.16 (0.37)	0.17 (0.37)
Master Degree	0.06 (0.23)	0.09 (0.29)
PhD Degree	0.02 (0.13)	0.01 (0.08)
Professional Degree	0.02 (0.14)	0.02 (0.15)
Race		
White	0.52 (0.50)	0.50 (0.50)
Black	0.28 (0.45)	0.31 (0.46)
Hispanic	0.20 (0.40)	0.19 (0.39)
Number of Observations	1,210	1,166

Appendix B Table 2. Mean Values of 1995 March CPS and 1995 ORG Data

	1995 March CPS Data		1995 ORG Data	
Variables	**Male**	**Female**	**Male**	**Female**
Hourly real wage	10.42	7.84	10.09	7.80
	(8.39)	(12.3)	(6.05)	(5.02)
Age	40.05	40.09	39.91	40.11
	(9.43)	(9.37)	(9.36)	(9.26)
Marital Status				
Married	0.713	0.655	0.710	0.638
	(0.45)	(0.47)	(0.45)	(0.48)
Widowed	0.004	0.023	0.004	0.024
	(0.06)	(0.15)	(0.06)	(0.15)
Divorced	0.109	0.175	0.108	0.181
	(0.31)	(0..38)	(0.31)	(0..38)
Never married	0.172	0.145	0.176	0.155
	(0.37)	(0.35)	(0.38)	(0.36)
Education				
Less junior high school	0.030	0.018	0.019	0.011
	(0.17)	(0.13)	(0.13)	(0.10)
Some high school	0.076	0.062	0.065	0.053
	(0.26)	(0.24)	(0.24)	(0.22)
High school	0.322	0.343	0.328	0.338
	(0.46)	(0.47)	(0.46)	(0.47)
Some college	0.181	0.198	0.184	0.199
	(0.38)	(0.39)	(0.38)	(0.39)
College degree	0.189	0.181	0.196	0.193
	(0.39)	(0.38)	(0.39)	(0.39)
Master	0.067	0.064	0.071	0.071
	(0.25)	(0.24)	(0.25)	(0.25)
Ph.D.	0.015	0.010	0.016	0.007
	(0.12)	(0.10)	(0.12)	(0.08)
Race/Ethnicity				
White	0.855	0.833	0.852	0.829
	(0.34)	(0.37)	(0.35)	(0.37)
Black	0.076	0.100	0.083	0.112
	(0.26)	(0.30)	(0.27)	(0.31)
Asian	0.033	0.034	0.035	0.034
	(0.17)	(0.18)	(0.18)	(0.18)
Hispanic	0.128	0.108	0.076	0.061
	(0.33)	(0.31)	(0.26)	(0.24)
Firm Size				
< 10 workers	0.140	0.139	----------	----------
	(0.34)	(0.34)		
[10-25] workers	0.092	0.08	----------	----------
	(0.29)	(0.27)		
[25-99]workers	0.143	0.127	-----------	-----------
	(0.35)	(0.33)		
[100-499] workers	0.153	0.158	-----------	-----------
	(0.36)	(0.36)		
[500-1000] workers	0.058	0.069	----------	----------
	(0.23)	(0.25)		
> 1000 workers	0.407	0.423	----------	----------
	(0.49)	(0.49)		
No. of observations	25,681	24,814	65,365	63,063

VI. References

Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administration Records." *American Economic Review* 80(3) (June): 313-36.

Angrist, Joshua D., and William N. Evans. 1998. "Children and their Parent's Labor Supply: Evidence from an Exogenous Variation in Family Size." *American Economic Review* 88(3) (June): 450-77.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics,* Volume 3A. Amsterdam: North Holland.

Black, Dan A., Mark C. Berger, and Frank A. Scott. 2000. "Bounding Parameter Estimates with Nonclassical Measurement Error." *Journal of the American Statistical Association* 95(451)(September): 739-48.

Card, David. 1998. "The Casual Effect of Education on Earnings." in Orely C. Ashenfelter and David Card's *Handbook of Labor Economics* 3A Amsterdam: Elsevier Science, 1801-63.

________. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69(5): 1127-60.

Evans, William N., Matthew C. Farrelly, and Edward Montgomery. 1999. "Do Workplace Smoking Bans Reduce Smoking?" *American Economic Review* 89(4)(September): 728-47.

Griliches, Zvi. 1986. "Economic Data Issues." in Zvi Griliches and Michael D. Intriligator (eds.), *Handbook of Econometrics* Volume 3. Amsterdam: North Holland.

Hausman, Jerry A., Whitney K. Newey, Hidehiko Ichimura, and James L. Powell. 1991. "Identification and Measurement Polynomial Errors-in-Variables Models." *Journal of Econometrics* 50(3) (December): 273-95.

Heckman, James J., Hidehiko Ichimura, Jeffrey A. Smith, and Petra Todd. 1998. "Characterizing Substitution Bias Using Experimental Data." *Econometrica* 66(5) (September): 1017-98.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1998. "Matching as an Econometric Estimator." *Review of Economic Studies* 65(2) (April): 261-94.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching As An Econometric Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies* 64(4): 487-36.

Hersch, Joni, and W. Kip Viscusi. 2001. "Cigarette Smokers as Job Risk Takers." *Review of Economics and Statistics* xx(2) (May): 263-82.

Imbens, Guido W. 1999. "The Role of the Propensity Score in Estimating Dose-Response Functions." NBER Technical Working Paper # 237. Cambridge, MA: National Bureau of Economic Research.

Kane, Thomas J., Cecilia E. Rouse, and Douglas Staiger. 1999. "Estimating Returns to Schooling When Schooling is Misreported." NBER Working Paper #7235. Cambridge, MA: National Bureau of Economic Research.

Kniesner, Thomas J., and John D. Leeth. 1995. *Simulating Workplace Safety.* Boston: Kluwer Academic Publisher.

Krueger, Alan, and Lawrence Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica* 56(2) March:

Lechner, Michael. 2000. "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." Unpublished paper. St. Gallen, Switzerland: University of St. Gallen. August.

Lubotsky, Darren, and Martin Wittenberg. 2001. "Interpretation of Regressions with Multiple Proxies." Unpublished paper, Princeton University.

Mellow, Wesley, and Hal Sider. 1983. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics* 1(4): 331-44.

Rosenbaum, Paul, and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1) (April): 41-55.

Shogren, Jason F., and Tommy Stamland. 2002. "Skill and the Value of Life." *Journal of Political Economy* 110(5) (October): 1168-1173.

Smith, Jeffrey. 2000. "A Critical Survey of Empirical Methods for Evaluating Active Labour Market Policies." *Schweizerische Zeitschrift für Volkswirtschaft und Statistik (Swiss Journal of Economics and Statistics)*, 136(3): 247-268.

Smith, Jeffrey, and Petra Todd. 2000. "Does Matching Overcome LaLonde's Critique of Non-experimental Estimators." Unpublished paper. London, Ontario: University of Western Ontario. November.

Viscusi, W. Kip. 1981. "Occupational Safety and Health Regulation: Its Impact and Policy Alternatives." In J. Crecine (ed.), *Research in Public Policy Analysis and Management,* Volume 2. Greenwich, CT: JAI Press, pp. 281-99.

Viscusi, W. Kip. 1993. "The Value of Risks to Life and Health." *Journal of Economic Literature* 31(5) (December): 1912-46.

Yule, G. Udny, 1907. "On the Theory of Correlation for any Number of Variables, treated by a New System of Notation." *Proceedings of the Royal Society of London.* Series A, containing Papers of Mathematical and Physical Character, 79(529) (May 14): 182-93.

Table 1A. Estimated Price of Risk for Male Workers

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	210	467	280	106
	(12.83)	(26.33)	(15.3)	(4.78)
Value of Statistical Life (in millions of dollars)	7.4	16.4	9.8	3.7

*There are 266,534 observations in the regressions.

Panel 2. March CPS and NIOSH Occupation Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Industry	no	no	yes	yes
Occupation	no	no	no	yes
Risk/100,000	-262	-188	-90.5	19.47
	(-19.28)	(-13.01)	(-4.56)	(1.18)
Value of Statistical Life (in millions of dollars)	----	----	----	0.7

*There are 266,534 observations in the regressions.

Panel 3. March CPS and BLS Industry Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-84.5	7.73	187	-188
	(-3.76)	(0.35)	(7.77)	(-6.79)
Value of Statistical Life (in millions of dollars)	----	0.3	6.6	----

*There are 102,411 observations in the regressions.

Panel 4. March CPS and BLS Occupation Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-12.3	-5.17	-25.2	18.1
	(-2.15)	(-0.91)	(-4.46)	(3.24)
Value of Statistical Life (in millions of dollars)	----	----	----	0.6

*There are 102,411 observations in the regressions.

Table 1A cont. Estimated Price of Risk for Male Workers

Panel 5. NLSY and NIOSH Industry Risk: 1986-1993*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	281	673	754	-290
	(3.28)	(6.82)	(7.34)	(-5.59)
Value of Statistical Life (in millions of dollars)	9.8	23.6	26.5	----

*There are 20,338 observations in the regressions.

Panel 6. NLSY and NIOSH Occupation Risk: 1986-1993*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-253	-228	-290	-145
	(-5.19)	(-4.33)	(-5.59)	(-2.14)
Value of Statistical Life (in millions of dollars)	----	----	----	----

*There are 20,338 observations in the regressions.

Panel 7. ORG and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	160	359	464	84.2
	(20.39)	(42.75)	(55.72)	(8.01)
Value of Statistical Life (in millions of dollars)	5.6	12.6	16.3	2.9

*There are 550,119 observations in the regressions.

Panel 8. ORG and NIOSH Occupation Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-400	-338	-365	105
	(-54.27)	(-43.13)	(-46.06)	(10.58)
Value of Statistical Life (in millions of dollars)	----	----	----	3.7

*There are 550,119 observations in the regressions.

Table 1A cont. Estimated Price of Risk for Male Workers

Panel 9. ORG and BLS Industry Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-126	-22.7	170	-121
	(-9.36)	(-3.51)	(12.16)	(-6.90)
Value of Statistical Life (in millions of dollars)	----	----	6.0	----

*There are 242,109 observations in the regressions.

Panel 10. ORG and BLS Occupation Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-117	-83.8	-126	35.1
	(-16.06)	(-10.78)	(-17.82)	(4.39)
Value of Statistical Life in millions of dollars)	----	----	----	1.2

*There are 242,109 observations in the regressions.

Table 1B. Estimated Price of Risk for Female Workers

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	276	501	340	-31.9
	(10.77)	(18.91)	(12.4)	(-0.99)
Value of Statistical Life (in millions of dollars)	9.7	17.6	11.9	----

*There are 250,354 observations in the regressions.

Panel 2. March CPS and NIOSH Occupational Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-359	-262	-128	-59.1
	(-14.61)	(-10.29)	(-7.56)	(-2.07)
Value of Statistical Life (in millions of dollars)	----	----	----	----

*There are 250,354 observations in the regressions.

Panel 3. March CPS and BLS Industry Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	33.8	87.0	19.7	-292
	(0.87)	(2.27)	(4.81)	(-5.78)
Value of Statistical Life (in millions of dollars)	1.2	3.0	0.7	----

*There are 102,411 observations in the regressions.

Panel 4. March CPS and BLS Occupation Risk: 1995-2000*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-5.06	-4.92	-2.60	94.8
	(-0.22)	(-0.21)	(-0.11)	(4.10)
Value of Statistical Life (in millions of dollars)	----	----	----	3.3

*There are 102,411 observations in the regressions.

Table 1B cont. Estimated Price of Risk for Female Workers

Panel 5. NLSY and NIOSH Industry Risk: 1986-1993*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	129	716	700	-334
	(1.15)	(5.67)	(5.34)	(-1.79)
Value of Statistical Life (in millions of dollars)	4.5	25.1	24.6	----

*There are 19,272 observations in the regressions.

Panel 6. NLSY and NIOSH Occupation Risk: 1986-1993*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-547	-290	-250	48.20
	(-5.28)	(-2.61)	(-2.25)	(0.25)
Value of Statistical Life (in millions of dollars)	----	----	----	1.7

*There are 19,272 observations in the regressions.

Panel 7. ORG and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	293	438	415	-2.63
	(25.54)	(37.08)	(36.67)	(-0.19)
Value of Statistical Life (in millions of dollars)	10.3	15.4	14.6	----

*There are 556,532 observations in the regressions.

Panel 8. ORG and NIOSH Occupation Risk: 1985-1995*

Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-291	-231	-133	33.6
	(-22.71)	(-17.41)	(-10.17)	(2.16)
Value of Statistical Life (in millions of dollars)	----	----	----	1.2

*There are 556,532 observations in the regressions.

Table 1B cont. Estimated Price of Risk for Female Workers

Panel 9. ORG and BLS Industry Risk: 1995-2000*				
Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-55.10	-24.3	97.6	---
	(-2.30)	(-1.03)	(4.14)	(-11.01)
Value of Statistical Life (in millions of dollars)	----	----	3.4	----

*There are 246,904 observations in the regressions.

Panel 10. ORG and BLS Occupation Risk: 1995-2000*				
Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Occupation	no	no	yes	yes
Industry	no	no	no	yes
Risk/100,000	-130	-109	-68.1	56.7
	(-5.55)	(-4.82))	(-3.18)	(2.52)
Value of Statistical Life (in millions of dollars)	----	----	----	2.0

*There are 246,904 observations in the regressions.

Note: 1. The dependent variable is the natural log of the worker's real wage. For the basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, and a dummy variable indicating whether the worker is under a union contract or not, and dummy variables for the worker's marital status. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The pooled data sets (1985-1995, 1995-2000, 1986-1993) are used to estimate the time-specific fixed-effect (within-group estimate).
2. In NLSY data, the independent variables for the basic regression include a quartic in the workers' age, education level, union, working experience, tenure status, AFQT scores, race/ethnicity categories, and dummy variables for the worker's marital status. .
Source: Authors' calculations

Table 2. Estimated Returns to a Bachelor's Degree for ORG Data, 1995

Panel 1. Male workers-ORG 1995*				
Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Industry	no	no	yes	yes
Occupation	no	no	no	yes
BA Degree	0.397	0.386	0.385	0.240
(relative to high school degree)	(62.56	(47.53)	(47.63)	(27.86)
*There are 51,659 observations in the regressions.				
Panel 2. Female workers-ORG 1995*				
Basic Controls	yes	yes	yes	yes
State	no	yes	yes	yes
Industry	no	no	yes	yes
Occupation	no	no	no	yes
BA Degree	0.479	0.459	0.444	0.277
(relative to high school degree)	(76.68)	(57.72)	(56.24)	(33.68)
*There are 53,291 observations in the regressions.				

Note: The dependent variable is the natural log of the worker's real wage. For the basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the worker is under a union contract or not, and dummy variables for the worker's marital status. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses.

Table 3A. Semi-parametric Estimated Price of Risk for Male Workers

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	493	212	47.6
	(18.84)	(12.13)	(0.14)
Value of Statistical Life (in millions of dollars)	17.3	7.4	1.7

*There are 266,535 observations in the regression.

Panel 2. March CPS and NIOSH Occupation Risk: 1985-1995*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-235	-276	-377
	(-10.44)	(-19.03)	(-13.94)
Value of Statistical Life (in millions of dollars)	----	----	----

*There are 250,354 observations in the regression.

Panel 3. March CPS and BLS Industry Risk: 1995-2000*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-70.8	-43.8	-218
	(-2.88)	(-1.08)	(-1.14)
Value of Statistical Life (in millions of dollars)	----	----	----

*There are 112,416 observations in the regression

Panel 4. March CPS and BLS Occupation Risk: 1995-2000*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-11.91	-7.13	376
	(-1.86)	(-0.53)	(2.97)
Value of Statistical Life (in millions of dollars)	----	----	13.2

*There are 102,411 observations in the regression.

Table 3A cont. Semi-parametric Estimated Price of Risk for Male Workers

Panel 5. ORG and NIOSH Industry Risk: 1985-1995*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	151	323	451
	(19.23)	(35.65)	(7.36)
Value of Statistical Life (in millions of dollars)	5.3	11.3	15.8
*There are 546,210 observations in the regressions.			
Panel 6. ORG and NIOSH Occupation Risk: 1985-1995*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-403	-399	450
	(-54.45)	(-46.42)	(7.31)
Value of Statistical Life (in millions of dollars)	---	---	15.8
*There are 546,210 observations in the regressions.			
Panel 7. ORG and BLS Industry Risk: 1995-2000*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-142	-94.4	-339
	(-10.44)	(-6.25)	(-7.21)
Value of Statistical Life (in millions of dollars)	---	---	---
*There are 242,109 observations in the regressions.			
Panel 8. ORG and BLS Occupation Risk: 1995-2000*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-123	-120	13.60
	(-17.20)	(-14.44)	(0.66)
Value of Statistical Life (in millions of dollars)	---	---	0.5
*There are 242,109 observations in the regressions.			

Note: There is a fixed-effect in each regression for each combination of the independent variables. The basic regression includes age, education level, and race/ethnicity category. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses.
Source: Authors' calculations.

Table 3B. Semi-parametric Estimated Price of Risk for Female Workers

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	689	291	330
	(16.19)	(10.57)	(0.44)
Value of Statistical Life (in millions of dollars)	24.2	10.2	11.6

*There are 250,354 observations in the regression.

Panel 2. March CPS and NIOSH Occupation Risk: 1985-1995*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-337	-377	-73.7
	(-7.14)	(-13.94)	(-0.27)
Value of Statistical Life (in millions of dollars)	----	----	----

*There are 243,958 observations in the regression.

Panel 3. March CPS and BLS Industry Risk: 1995-2000*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	78.7	95.4	-526
	(1.81)	(1.28)	(-1.47)
Value of Statistical Life (in millions of dollars)	2.8	3.3	----

*There are 108,324 observations in the regression.

Panel 4. March CPS and BLS Occupation Risk: 1995-2000*

Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	7.64	48.1	-116
	(0.32)	(1.41)	(-0.24)
Value of Statistical Life (in millions of dollars)	0.3	1.7	----

*There are 84,534 observations in the regression.

Table 3B cont. Semi-parametric Estimated Price of Risk for Female Workers

Panel 5. ORG and NIOSH Industry Risk: 1985-1995*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	272	386	187
	(23.68)	(30.37)	(1.61)
Value of Statistical Life (in millions of dollars)	9.5	13.5	6.6
*There are 556,532 observations in the regressions.			

Panel 6. ORG and NIOSH Occupation Risk: 1985-1995*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-297	-251	-90.5
	(-23.07)	(-0.76)	(-0.76)
Value of Statistical Life (in millions of dollars)	----	----	----
*There are 556,532 observations in the regressions.			

Panel 7. ORG and BLS Industry Risk: 1995-2000*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-77.60	-22.9	-432
	(-3.21)	(-0.85)	(-5.60)
Value of Statistical Life (in millions of dollars)	----	----	----
*There are 242,109 observations in the regressions.			

Panel 8. ORG and BLS Occupation Risk: 1995-2000*			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-222	-191	180
	(-7.64)	(-5.89)	(2.64)
Value of Statistical Life (in millions of dollars)	----	----	6.3
*There are 242, 109 observations in the regressions.			

Note: There is a fixed-effect in each regression for each combination of the independent variables. The basic regression includes age, education level, race/ethnicity category, and marital status dummy variables. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses.
Source: Authors' calculations.

Table 4A. Estimated Price of Risk for Males by Decile of Risk

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995*

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.05	-0.02	0.05
[0.9, 1.4]	(-7.54)	(-2.89)	(0.81)
Third	-0.08	-0.02	-0.01
[1.4, 1.7]	(-10.40)	(-4.48)	(-1.50)
Fourth	-0.02	0.01	0.01
[1.7, 2.1]	(-4.35)	(2.08)	(1.59)
Fifth	-0.08	-0.00	-0.00
[2.1, 2.6]	(-11.66)	(-1.24)	(-0.13)
Sixth	0.03	0.07	0.02
[2.6, 3.2]	(4.62)	(10.48)	(2.86)
Seventh	-0.02	0.08	0.02
[3.2, 4]	(-4.09)	(11.88)	(2.64)
Eighth	0.01	0.10	0.02
[4, 6.4]	(1.88)	(15.88)	(2.56)
Ninth	0.06	0.15	0.03
[6.4, 12.5]	(11.44)	(23.36)	(3.81)
Tenth	0.01	0.16	0.05
[> 12.5]	(1.98)	(24.76)	(4.43)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 2. March CPS and NIOSH Occupation Risk: 1985-1995			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.12	0.05	0.01
[0.6, 0.9]	(15.14)	(6.07)	(1.41)
Third	0.11	0.07	0.00
[0.9, 1.4]	(14.80)	(8.70)	(0.11)
Fourth	0.08	0.04	0.03
[1.4, 1.9]	(11.13)	(5.95)	(3.52)
Fifth	0.08	0.06	0.00
[1.9, 2.5]	(10.99)	(7.41)	(1.04)
Sixth	0.02	0.01	0.01
[2.5, 3.2]	(3.10)	(1.62)	(1.12)
Seventh	0.00	0.03	0.01
[3.2, 4.1]	(0.14)	(4.10)	(1.15)
Eighth	0.07	0.09	0.01
[4.1, 6.8]	(10.31)	(11.15)	(1.20)
Ninth	0.07	0.07	0.01
[6.8, 11.9]	(10.18)	(9.03)	(1.40)
Tenth	-0.05	-0.03	0.01
[> = 11.9]	(-7.79)	(-3.96)	(1.40)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 3. March CPS and BLS Industry Risk: 1995-2000			
Basic Control	yes	yes	yes
State	yes	yes	yes
Industry/Occupation	no	no	yes
Second	-0.01	-0.01	0.01
[0.52, 0.67]	(-1.53)	(-1.07)	(1.08)
Third	0.05	0.05	0.01
[0.67, 0.90]	(5.22)	(5.33)	(1.42)
Fourth	0.07	0.07	0.00
[0.90, 1.30]	(7.71)	(7.33)	(0.17)
Fifth	0.05	0.04	0.00
[1.30, 1.98]	(5.53)	(1.64)	(0.63)
Sixth	-0.01	-0.01	-0.04
[1.98, 2.83]	(-1.83)	(-2.08)	(-5.15)
Seventh	-0.05	-0.04	-0.06
[2.83, 4.03]	(-5.51)	(-5.23)	(-7.21)
Eighth	-0.00	-0.00	-0.04
[4.03, 6.51]	(-0.77)	(-0.19)	(-4.28)
Ninth	0.06	0.07	-0.04
[6.51, 13.05]	(7.11)	(8.02)	(-4.65)
Tenth	0.02	0.03	-0.08
[> 13.05]	(2.74)	(3.89)	(8.35)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 4. March CPS and BLS Occupation Risk: 1995-2000

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.26	-0.26	-.017
[0.0, 0.37]	(-8.28)	(-8.19)	(-5.59)
Third	-0.00	-0.04	-0.00
[0.37, 0.86]	(-0.25)	(-0.47)	(-0.33)
Fourth	0.02	0.02	0.05
[0.86, 1.26]	(2.84)	(2.81)	(6.15)
Fifth	0.02	0.02	0.01
[1.26, 2.03]	(2.91)	(2.52)	(2.28)
Sixth	0.19	0.19	0.09
[2.03, 2.61]	(25.09)	(24.99)	(11.74)
Seventh	0.01	0.01	0.01
[2.61, 3.58]	(1.92)	(1.69)	(1.59)
Eighth	0.01	0.01	0.02
[3.58, 5.51]	(1.53)	(1.74)	(3.26)
Ninth	0.01	0.01	0.07
[5.51, 15.19]	(1.35)	(2.04)	(9.52)
Tenth	-0.05	-0.03	0.09
[> 15.19]	(-7.04)	(-5.37)	(10.69)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 5. NLSY and NIOSH Industry Risk: 1986-1993

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.06	-0.03	0.02
[1, 1.4]	(-2.48)	(-0.95)	(0.79)
Third	-0.12	-0.11	-0.02
[1.4, 2]	(-4.02)	(-3.78)	(-0.68)
Fourth	-0.04	-0.03	0.01
[2, 2.1]	(-1.40)	(-0.88)	(0.37)
Fifth	-0.09	0.01	0.05
[2.1, 2.7]	(-3.09)	(0.50)	(1.62)
Sixth	0.02	0.08	0.04
[2.7, 3.4]	(1.21)	(2.98)	(1.25)
Seventh	-0.06	0.06	0.07
[3.4, 4.6]	(-2.62)	(2.29)	(2.00)
Eighth	0.03	0.09	0.01
[4.6, 7.1]	(1.18)	(3.41)	(0.27)
Ninth	0.05	0.13	0.00
[7.1, 12.3]	(2.23)	(4.92)	(-0.05)
Tenth	0.02	0.14	-0.02
[> 12.3]	(0.85)	(4.89)	(-0.42)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 6. NLSY and NIOSH Occupation Risk: 1986-1993

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.09	0.02	0.00
[0.6, 1]	(3.03)	(0.62)	(-0.16)
Third	0.08	0.06	0.03
[1, 1.5]	(2.64)	(1.92)	(0.86)
Fourth	0.07	0.03	-0.01
[1.5, 2.1]	(2.41)	(0.92)	(-0.44)
Fifth	0.04	0.05	0.01
[2.1, 2.7]	(1.46)	(1.68)	(0.21)
Sixth	0.06	0.05	0.04
[2.7, 3.7]	(1.97)	(1.71)	(0.96)
Seventh	0.05	0.06	0.02
[3.7, 4.6]	(1.77)	(2.04)	(0.57)
Eighth	0.16	0.14	0.01
[4.6, 7.9]	(5.86)	(4.99)	(0.28)
Ninth	0.07	0.09	0.01
[7.9, 16.2]	(2.54)	(3.13)	(0.12)
Tenth	-0.04	-0.04	-0.03
[> 16.2]	(-1.49)	(1.39)	(-0.50)

Table 4A. Estimated Price of Risk for Males by Decile of Risk

Panel 7. ORG and NIOSH Industry Risk: 1985-1995

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.06	-0.04	-0.02
[0.9, 1.4]	(-22.24)	(-14.02)	(-6.96)
Third	-0.08	-0.03	-0.01
[1.4, 1.7]	(-22.90)	(-9.23)	(-4.97)
Fourth	-0.03	-0.01	-0.01
[1.7, 2]	(-11.96)	(-6.12)	(-3.35)
Fifth	-0.07	-0.01	-0.02
[2, 2.5]	(-24.74)	(-4.70)	(-6.95)
Sixth	0.01	0.04	-0.01
[2.5, 3.2]	(3.97)	(15.61)	(-3.07)
Seventh	-0.01	0.06	-0.01
[3.2, 4]	(-6.11)	(22.32)	(-3.97)
Eighth	0.01	0.09	-0.01
[4, 6]	(5.63)	(30.19)	(-5.09)
Ninth	0.03	0.11	-0.01
[6, 11.4]	(10.81)	(37.21)	(-4.33)
Tenth	-0.00	0.10	-0.01
[> 11.4]	(-0.90)	(36.03)	(-3.66)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 8. ORG and NIOSH Occupation Risk: 1985-1995			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.13	0.07	0.01
[0.6, 0.9]	(38.56)	(18.43)	(3.86)
Third	0.13	0.09	0.00
[0.9, 1.3]	(36.21)	(22.95)	(2.48)
Fourth	0.06	0.03	0.01
[1.3, 1.8]	(17.91)	(8.95)	(3.14)
Fifth	0.04	0.03	-0.00
[1.8, 2.4]	(11.66)	(8.76)	(-0.20)
Sixth	0.06	0.03	0.00
[2.4, 2.9]	(17.55)	(10.28)	(0.30)
Seventh	0.00	0.00	-0.00
[2.9, 3.9]	(2.07)	(1.77)	(-0.29)
Eighth	0.04	0.04	-0.00
[3.9, 6.5]	(13.83)	(11.19)	(1.44)
Ninth	0.05	0.04	-0.01
[6.5, 10.9]	(17.12)	(12.07)	(-2.83)
Tenth	-0.06	-0.07	-0.00
[> 10.9]	(-20.02)	(-20.40)	(-0.03)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 9. ORG and BLS Industry Risk: 1995-2000

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.04	0.03	0.03
[0.47, 0.6]	(7.33)	(4.83)	(4.54)
Third	0.08	0.06	0.02
[0.6, 0.8]	(13.36)	(10.05)	(4.33)
Fourth	0.06	0.05	0.00
[0.8, 1.1]	(11.88)	(8.31)	(-0.56)
Fifth	0.00	-0.02	-0.03
[1.1, 1.7]	(0.02)	(-4.44)	(-5.59)
Sixth	-0.01	-0.02	-0.05
[1.7, 2.3]	(-1.87)	(-4.51)	(-9.93)
Seventh	-0.01	-0.02	-0.04
[2.3, 3.5]	(-1.55)	(-3.56)	(-7.07)
Eighth	0.02	0.01	-0.03
[3.5, 5.7]	(5.25)	(3.15)	(-6.30)
Ninth	0.05	0.05	-0.03
[5.7, 11.8]	(10.41)	(8.37)	(-5.35)
Tenth	0.05	0.04	-0.04
[> 11.8]	(10.32)	(7.74)	(-9.73)

Table 4A cont. Estimated Price of Risk for Males by Decile of Risk

Panel 10. ORG and BLS Occupation Risk: 1995-2000			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.05	0.05	0.05
[0, 0.3]	(4.43)	(3.89)	(4.25)
Third	-0.04	-0.05	0.01
[0.3, 0.7]	(-7.58)	(-8.71)	(2.44)
Fourth	-0.06	-0.06	0.00
[0.7, 1]	(-12.22)	(-11.31)	(-1.21)
Fifth	0.02	0.01	0.00
[1, 1.7]	(4.50)	(3.02)	(0.48)
Sixth	0.01	0.00	0.00
[1.7, 2.5]	(3.25)	(0.94)	(0.02)
Seventh	0.08	0.08	0.01
[2.5, 3.4]	(19.32)	(17.45)	(3.28)
Eighth	0.02	0.01	0.00
[3.4, 5.3]	(6.23)	(2.74)	(1.94)
Ninth	-0.06	-0.05	0.01
[5.3, 13.9]	(-14.17)	(-11.92)	(3.80)
Tenth	-0.04	-0.04	-0.01
[> 13.9]	(-12.84)	(-10.95)	(-2.86)

Table 4B. Estimated Price of Risk for Females by Decile of Risk

Panel 1. March CPS and NIOSH Industry Risk: 1985-1995

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.08	-0.03	-0.01
[0.9, 1.4]	(-15.64)	(-6.52)	(-1.89)
Third	-0.09	-0.04	-0.01
[1.4, 1.7]	(-15.66)	(-6.56)	(-1.44)
Fourth	-0.10	-0.05	-0.02
[1.7, 2.1]	(-20.82)	(-8.66)	(-3.16)
Fifth	-0.20	-0.08	-0.04
[2.1, 2.6]	(-32.61)	(-12.48)	(-5.53)
Sixth	-0.08	-0.03	-0.00
[2.6, 3.2]	(-12.21)	(-5.18)	(-0.05)
Seventh	-13	-0.03	-0.00
[3.2, 4]	(-22.40)	(-4.53)	(-0.05)
Eighth	-0.03	0.04	-0.01
[4, 6.4]	(-5.19)	(5.82)	(-1.82)
Ninth	0.02	0.12	-0.07
[6.4, 12.5]	(3.72)	(16.56)	(-0.62)
Tenth	0.00	0.12	-0.01
[> 12.5]	(0.63)	(14.13)	(-0.81)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 2. March CPS and NIOSH Occupation Risk: 1985-1995			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.07	0.02	0.00
[0.6, 0.9]	(13.27)	(3.98)	(1.08)
Third	0.05	0.04	0.00
[0.9,1.4]	(9.47)	(6.48)	(1.08)
Fourth	-0.01	-0.03	-0.00
[1.4, 1.9]	(-1.97)	(-4.85)	(-0.01)
Fifth	-0.03	-0.03	-0.02
[1.9,2.5]	(-6.11)	(-5.43)	(-3.18)
Sixth	-0.10	-0.10	-0.02
[2.5, 3.2]	(-16.14)	(-15.10)	(-2.69)
Seventh	-0.10	-0.07	-0.02
[3.2, 4.1]	(-17.25)	(-10.95)	(-2.58)
Eighth	-0.06	-0.00	-0.03
[4.1, 6.8]	(-9.29)	(-1.16)	(-3.32)
Ninth	-0.02	-0.02	-0.04
[6.8, 11.9]	(-2.51)	(-3.02)	(-3.27)
Tenth	-0.11	-0.10	-0.04
[> 11.9]	(-13.30)	(-11.17)	(-3.20)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 3. March CPS and BLS Industry Risk: 1995-2000

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.08	0.08	0.07
[0.52, 0.67]	(12.07)	(13.01)	(12.31)
Third	0.14	0.14	0.11
[0.67, 0.90]	(21.24)	(20.77)	(17.91)
Fourth	0.15	0.14	0.10
[0.90, 1.30]	(22.13)	(20.71)	(14.68)
Fifth	0.07	0.06	0.07
[1.30, 1.98]	(10.10)	(8.93)	(9.52)
Sixth	0.02	0.02	0.02
[1.98, 2.83]	(3.80)	(2.80)	(3.62)
Seventh	-0.03	-0.04	-0.01
[2.83, 4.03]	(-4.57)	(-5.23)	(-0.17)
Eighth	0.00	0.08	0.00
[4.03, 6.51]	(1.13)	(0.89)	(0.42)
Ninth	0.10	0.09	0.01
[6.51, 13.05]	(10.32)	(10.20)	(1.66)
Tenth	0.13	0.13	-0.01
[> 13.05]	(12.05)	(12.74)	(-0.94)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 4. March CPS and BLS Occupation Risk: 1995-2000

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.02	0.02	0.00
[0.0, 0.37]	(3.26)	(2.76)	(0.70)
Third	0.06	0.05	0.06
[0.37, 0.86]	(8.56)	(8.30)	(9.68)
Fourth	0.07	0.06	0.08
[0.86, 1.26]	(9.31)	(8.88)	(11.26)
Fifth	-0.00	-0.00	0.03
[1.26, 2.03]	(-0.34)	(-1.28)	(5.17)
Sixth	0.14	0.13	0.06
[2.03, 2.61]	(17.02)	(16.00)	(7.55)
Seventh	0.04	0.03	0.04
[2.61, 3.58]	(4.77)	(3.92)	(4.83)
Eighth	-0.05	-0.05	-0.01
[3.58, 5.51]	(-6.12)	(-6.49)	(-1.74)
Ninth	-0.11	-0.11	-0.04
[5.51, 15.19]	(-9.65)	(-10.21)	(-3.48)
Tenth	-0.49	-0.04	0.02
[> 15.19]	(-3.42)	(-2.95)	(1.70)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 5. NLSY and NIOSH Industry Risk: 1986-1993

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.06	0.01	0.00
[1, 1.4]	(-2.90)	(0.41)	(-0.14)
Third	-0.11	-0.08	-0.05
[1.4, 2]	(-4.49)	(-2.94)	(-1.67)
Fourth	-0.09	-0.06	-0.05
[2, 2.1]	(-4.16)	(-2.13)	(-1.74)
Fifth	-0.13	-0.02	-0.02
[2.1, 2.7]	(-5.23)	(-0.63)	(-0.48)
Sixth	0.00	0.03	-0.01
[2.7, 3.4]	(-0.04)	(1.17)	(-0.39)
Seventh	-0.12	-0.02	-0.03
[3.4, 4.6]	(-4.98)	(-0.60)	(-0.89)
Eighth	-0.04	0.03	-0.08
[4.6, 7.1]	(-1.69)	(0.96)	(-1.87)
Ninth	-0.03	0.09	-0.11
[7.1, 12.3]	(-1.11)	(3.02)	(-2.20)
Tenth	-0.03	0.10	-0.15
[> 12.3]	(-1.13)	(2.84)	(-2.54)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 6. NLSY and NIOSH Occupation Risk: 1986-1993			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.03	-0.01	0.00
[0.6, 1]	(1.61)	(-0.53)	(0.18)
Third	-0.06	-0.05	0.00
[1, 1.5]	(-2.82)	(-1.96)	(0.15)
Fourth	0.03	0.00	0.07
[1.5, 2.1]	(1.42)	(0.00)	(2.43)
Fifth	-0.09	-0.07	0.02
[2.1, 2.7]	(-3.85)	(-2.92)	(0.66)
Sixth	-0.08	-0.07	0.10
[2.7, 3.7]	(-3.06)	(-2.68)	(2.53)
Seventh	-0.12	-0.10	0.04
[3.7, 4.6]	(-5.29)	(-3.92)	(1.15)
Eighth	-0.08	-0.06	0.02
[4.6, 7.9]	(-3.03)	(-1.90)	(0.46)
Ninth	-0.17	-0.13	-0.05
[7.9, 16.2]	(-5.00)	(-3.72)	(-0.78)
Tenth	-0.14	-0.11	-0.01
[> 16.2]	(-4.53)	(-3.49)	(-0.19)

.

Table 4B. Estimated Price of Risk for Females by Decile of Risk

Panel 7. ORG and NIOSH Industry Risk: 1985-1995			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	-0.07	-0.03	-0.09
[0.9, 1.4]	(-33.87)	(-14.93)	(-4.04)
Third	-0.08	-0.04	-0.02
[1.4, 1.7]	(-32.41)	(-15.09)	(-0.78)
Fourth	-0.08	-0.05	-0.01
[1.7, 2]	(-35.28)	(-19.39)	(-4.68)
Fifth	-0.14	-0.07	-0.01
[2, 2.5]	(-59.60)	(-26.85)	(-7.10)
Sixth	-0.09	-0.04	-0.00
[2.5, 3.2]	(-36.87)	(-16.77)	(-1.56)
Seventh	-0.11	-0.03	-0.00
[3.2, 4]	(-45.25)	(-13.07)	(0.88)
Eighth	-0.03	0.01	-0.01
[4, 6]	(-11.74	(6.44)	(-2.90)
Ninth	0.01	0.09	-0.00
[6, 11.4]	(4.92)	(29.32)	(-2.06)
Tenth	0.03	0.09	0.00
[> 11.4]	(11.79)	(31.06)	(1.24)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 8. ORG and NIOSH Occupation Risk: 1985-1995			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.07	0.02	0.01
[0.6, 0.9]	(35.64)	(9.15)	(1.23)
Third	0.05	0.03	0.00
[0.9, 1.3]	(22.30)	(12.80)	(0.07)
Fourth	-0.01	-0.03	0.00
[1.3, 1.8]	(-4.78)	(-12.33)	(0.47)
Fifth	-0.05	-0.04	-0.02
[1.8, 2.4]	(-17.97)	(-14.22)	(-7.67)
Sixth	-0.03	-0.05	-0.02
[2.4, 2.9]	(-14.02)	(-19.33)	(-7.74)
Seventh	-0.06	-0.07	-0.02
[2.9, 3.9]	(-28.18)	(-29.44)	(-7.98)
Eighth	-0.06	-0.04	-0.02
[3.9, 6.5]	(-24.23)	(13.76)	(-6.96)
Ninth	0.01	-0.00	-0.03
[6.5, 10.9]	(3.19)	(-2.43)	(-6.82)
Tenth	0.01	-0.03	-0.00
[> 10.9]	(6.33)	(-11.90)	(-0.54)

.

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 9. ORG and BLS Industry Risk: 1995-2000

Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.11	0.11	0.09
[0.4, 0.6]	(29.23)	(25.73)	(23.16)
Third	0.20	0.18	0.14
[0.6, 0.8]	(55.69)	(45.41)	(36.06)
Fourth	0.17	0.15	0.10
[0.8, 1.1]	(46.19)	(35.90)	(26.06)
Fifth	0.06	0.04	0.06
[1.1, 1.7]	(15.79)	(9.90)	(14.25)
Sixth	0.05	0.03	0.03
[1.7, 2.3]	(11.67)	(7.68)	(7.89)
Seventh	0.07	0.05	0.02
[2.3, 3.5]	(15.78)	(11.00)	(5.84)
Eighth	0.08	0.07	0.02
[3.5, 5.7]	(17.79)	(14.20)	(4.36)
Ninth	0.11	0.10	0.01
[5.7, 11.8]	(21.08)	(17.11)	(2.78)
Tenth	0.13	0.11	0.01
[> 11.8]	(40.42)	(30.63)	(4.62)

Table 4B cont. Estimated Price of Risk for Females by Decile of Risk

Panel 10. ORG and BLS Occupation Risk: 1995-2000			
Basic Control	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Second	0.00	0.00	0.01
[0, 0.3]	(1.21)	(0.51)	(2.88)
Third	0.01	0.01	0.06
[0.3, 0.7]	(3.63)	(3.08)	(16.15)
Fourth	0.03	0.03	0.05
[0.7, 1]	(7.97)	(7.37)	(12.62)
Fifth	0.03	0.02	0.02
[1, 1.7]	(7.98)	(5.67)	(5.64)
Sixth	0.02	0.01	0.01
[1.7, 2.5]	(4.96)	(2..37)	(3.04)
Seventh	0.09	0.09	0.03
[2.5, 3.4]	(20.63)	(18.51)	(6.72)
Eighth	0.07	0.05	0.04
[3.4, 5.3]	(15.72)	(10.72)	(8.93)
Ninth	-0.04	-0.04	0.01
[5.3, 13.9]	(-6.82)	(-6.39)	(1.97)
Tenth	-0.01	-0.01	-0.02
[>13.9]	(-3.02)	(-5.38)	(-7.74)

Note: 1. The NIOSH and BLS industry and occupation risk data are divided into 10 deciles. The dependent variable is the natural log of the worker's real wage. For the basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the worker is under a union contract or not, and marital status. Workers are aged 25 to 60 inclusive. T-statistics are given in parenthesis. Cut-off points are given in brackets.
2. In NLSY data, the independent variables for the basic regression include a drop in the workers' age, educational level, union coverage, working experience, tenure status, AFQT test scores and race/ethnicity categories.
Source: Authors' calculation

Table 5A. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Males
1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data

Panel 1. March CPS and NIOSH Industry Risk	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	10.55	10.18	10.86	10.85
Mean real wage of matched comparison group	10.99	11.60	11.03	13.42
Treatment Effect	**-0.44**	**-1.42**	**-0.19**	**-2.57**
	(-0.15)	(-0.71)	(-0.56)	(-0.72)

Panel 2. ORG and NIOSH Industry Risk	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	9.28	9.54	9.86	9.89
Mean real wage of matched comparison group	9.58	10.21	10.82	11.30
Treatment Effect	**-0.29**	**-0.66**	**-0.95**	**-1.41**
	(-1.47)	(-3.03)	(-4.24)	(-6.86)

Panel 3. NLSY and NIOSH Industry Risk	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	9.20	8.86	10.19	8.52
Mean real wage of matched comparison group	10.53	10.79	8.69	10.66
Treatment Effect	**-1.33**	**-1.93**	**1.50**	**-2.14**
	(-0.45)	(-1.30)	(1.46)	(-1.36)

Panel 4. March CPS and NIOSH Occupation Risk	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	8.67	10.53	11.68	13.56
Mean real wage of matched comparison group	8.86	9.42	9.72	11.39
Treatment Effect	**-0.18**	**1.11**	**1.95**	**2.17**
	(-0.30)	(1.94)	(2.83)	(2.56)

*Nearest-neighbor matching with a caliper 0.01

**Table 5A cont. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Males
1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data**

Panel 5. ORG and NIOSH Occupation Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	7.72	9.54	11.10	11.98
Mean real wage of matched comparison group	8.37	9.06	9.65	10.95
Treatment Effect	**-0.64**	**0.47**	**1.37**	**1.03**
Panel 6. NLSY and NIOSH Occupation Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	7.74	9.44	10.03	11.02
Mean real wage of matched comparison group	8.19	9.40	9.24	10.57
Treatment Effect	**-0.45**	**0.04**	**0.79**	**0.45**
	(-0.69)	(0.05)	(1.16)	(0.53)
Panel 7. March CPS and BLS Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	10.99	9.85	10.55	13.23
Mean real wage of matched comparison group	9.77	9.63	10.14	11.30
Treatment Effect	**1.21**	**0.22**	**0.41**	**1.93**
	(2.97)	(0.69)	(1.38)	(3.27)
Panel 8. ORG and BLS Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	8.35	8.83	10.13	11.57
Mean real wage of matched comparison group	8.39	9.18	9.75	10.92
Treatment Effect	**-0.03**	**-0.34**	**0.38**	**0.64**
	(-0.19)	(-2.13)	(2.24)	(3.66)

*Nearest-neighbor matching with a caliper 0.01

Table 5A cont. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Males 1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data

Panel 9. March CPS and BLS Occupation Risk	**Quintile 5 vs. 1**	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	9.76	11.47	12.89	11.35
Mean real wage of matched comparison group	9.71	11.30	11.89	11.73
Treatment Effect	**0.05**	**0.17**	**1.00**	**-0.38**
	(0.90)	(0.26)	(0.85)	(-0.46)

Panel 10. ORG and BLS Occupation Risk	**Quintile 5 vs. 1**	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	8.84	10.77	10.93	10.84
Mean real wage of matched comparison group	9.06	10.29	10.41	10.74
Treatment Effect	**-0.22**	**0.48**	**0.53**	**0.10**
	(-1.65)	(3.25)	(3.37)	(0.55)

*Nearest-neighbor matching with a caliper 0.01
Source: Authors' calculations.

Table 5B. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Females 1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data

	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Panel 1. March CPS and NIOSH Industry Risk				
Mean real wage of matched treatment group	8.35	7.57	7.30	8.07
Mean real wage of matched comparison group	7.82	8.08	7.83	8.31
Treatment Effect	**0.53**	**-0.50**	**-0.52**	**-0.23**
	(2.17)	(-1.75)	(-2.16)	(-0.74)
Panel 2. ORG and NIOSH Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	7.87	6.74	6.96	7.46
Mean real wage of matched comparison group	7.54	7.44	7.44	8.01
Treatment Effect	**0.33**	**-0.69**	**-0.47**	**-0.54**
	(1.54)	(-5.08)	(-3.28)	(-4.33)
Panel 3. NLSY and NIOSH Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	8.35	6.97	7.50	8.71
Mean real wage of matched comparison group	5.22	9.74	8.97	8.04
Treatment Effect	**3.13**	**-2.77**	**-1.47**	**0.67**
	(3.49)	(-1.67)	(-1.46)	(0.64)
Panel 4. March CPS and NIOSH Occupation Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	6.10	7.42	8.75	9.74
Mean real wage of matched comparison group	6.69	7.07	7.62	8.77
Treatment Effect	**-0.58**	**0.34**	**0.72**	**0.96**
	(-2.72)	(1.55)	(2.14)	(2.09)

*Nearest-neighbor matching with a caliper 0.01

Table 5B cont. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Females 1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data

	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Panel 5. ORG and NIOSH Occupation Risk				
Mean real wage of matched treatment group	5.99	6.80	8.42	8.88
Mean real wage of matched comparison group	6.87	7.17	8.15	7.79
Treatment Effect	**-0.88**	**-0.37**	**0.27**	**1.08**
	(-1.58)	(-0.74)	(0.31)	(2.07)
Panel 6. NLSY and NIOSH Occupation Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	5.99	6.80	8.42	8.88
Mean real wage of matched comparison group	6.87	7.17	8.15	7.79
Treatment Effect	**-0.88**	**-0.37**	**0.27**	**1.08**
	(-1.58)	(-0.74)	(0.31)	(2.07)
Panel 7. March CPS and BLS Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	7.88	7.46	8.00	8.04
Mean real wage of matched comparison group	7.10	6.82	7.61	8.06
Treatment Effect	**0.78**	**0.64**	**0.39**	**-0.02**
	(-4.62)	(0.87)	(1.04)	(-0.08)
Panel 8. ORG and BLS Industry Risk				
	Quintile 5 vs. 1	**Quintile 4 vs. 1**	**Quintile 3 vs. 1**	**Quintile 2 vs. 1**
Mean real wage of matched treatment group	7.19	6.23	7.31	7.91
Mean real wage of matched comparison group	7.34	7.28	7.71	7.86
Treatment Effect	**-0.14**	**-1.04**	**-0.39**	**0.04**
	(-0.92)	(-5.53)	(-4.13)	(0.31)

*Nearest-neighbor matching with a caliper 0.01

Table 5B cont. Propensity Score Matching Estimators: Nearest-Neighbor Matching* for Females 1993 CPS Outgoing Rotation Data, 1993 March CPS Data, 1995 NIOSH Risk Data, 1994 BLS Risk Data, 1993 NLSY Data

Panel 9. March CPS and BLS Occupation Risk

	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	8.07	8.07	7.34	8.41
Mean real wage of matched comparison group	7.98	7.12	7.43	8.06
Treatment Effect	**0.09**	**0.94**	**-0.09**	**0.34**
	(0.29)	(3.68)	(-0.24)	(1.17)

Panel 10. ORG and BLS Occupation Risk

	Quintile 5 vs. 1	Quintile 4 vs. 1	Quintile 3 vs. 1	Quintile 2 vs. 1
Mean real wage of matched treatment group	7.27	7.46	7.25	8.01
Mean real wage of matched comparison group	7.08	6.73	6.87	7.25
Treatment Effect	**0.19**	**0.72**	**0.37**	**0.75**
	(2.18)	(6.84)	(3.40)	(-3.58)

Note: *Nearest-neighbor matching with a caliper 0.01. The NIOSH and BLS industry and occupation risk data are divided into 5 quintiles. The logit model was used to calculate the probability in each risk quintile versus risk quintile 1. For NLSY data the independent variables include a quartic in the worker's age, education level, experience, tenure, test scores, union, and dummy variables indicating firm size, marital status, industry and occupation status, and race/ethnicity categories. For CPS Outgoing Rotation data, the independent variables include age, age quartic, educational level, firm size, race/ethnicity, marital status, and state. The probabilities are used as propensity scores to get the nearest-neighbor estimations. Here, 0.01 is the nearest-neighbor matching standard.
Source: Authors' calculations.

Table 6. Correlations and OLS Residual Correlations for Male Workers
1995 CPS Outgoing Rotations Data, CPS March Data, BLS and NIOSH Risk Data

A. March CPS Data				
Basic Controls	no	yes	yes	yes
State	no	no	yes	yes
Industry/Occupation	no	no	no	yes
Correlations				
NIOSH Ind / NIOSH Occ	0.50	0.48	0.39	0.41
NIOSH Ind / BLS Ind	0.47	0.40	0.41	0.02
NIOSH Ind / BLS Occ	0.29	0.24	0.25	0.02
NIOSH Occ / BLS Ind	0.36	0.31	0.34	0.04
NIOSH Occ / BLS Occ	0.38	0.34	0.35	0.04
BLS Ind / BLS Occ	0.44	0.39	0.38	0.22
B. CPS Outgoing Rotation Data				
Basic Controls	no	yes	yes	yes
State	no	no	yes	yes
Industry/Occupation	no	no	no	yes
Correlations				
NIOSH Ind / NIOSH Occ	0.53	0.43	0.32	0.28
NIOSH Ind / BLS Ind	0.48	0.45	0.45	0.06
NIOSH Ind / BLS Occ	0.30	0.27	0.26	0.05
NIOSH Occ / BLS Ind	0.37	0.33	0.31	0.07
NIOSH Occ / BLS Occ	0.40	0.38	0.38	0.09
BLS Ind / BLS Occ	0.43	0.40	0.39	0.22

Note: The residual correlations are based on the OLS regression of the risk variable on a set of independent variables. The basic controls are dummy variables for age, age quartic, education, race, ethnicity, union coverage, and marital status. After estimating the residuals for each regression, we estimated the residual correlations for each set of regressions. The number of observations for the 1995 CPS Outgoing Rotations data and 1995 March CPS data are 51,140 and 25,237, respectively.
Source: Authors' calculations.

Table 7. Covariances and Variances of Residual Estimates for Male Workers
1995 CPS Outgoing Rotation Data, NIOSH Risk Data, and BLS Risk Data

Panel 1. NIOSH Industry / NIOSH Occupation

	Basic Controls	State	Ind/Occ
VAR (Lnwage)			
VAR (NIOSH Ind)	0.24	0.23	0.21
VAR (NIOSH Occ)	39.92	33.91	14.80
COV (NIOSH Ind, NIOSH Occ)	45.71	38.57	17.09
COV (Lnwage, NIOSH Ind)	18.43	11.56	4.51
COV (Lnwage, NIOSH Occ)	0.06	0.15	0.01
R2 Lnwage on X	-0.23	-0.13	0.02
R2 NIOSH Ind on X	0.27	0.29	0.37
R2 NIOSH Occ on X	0.03	0.21	0.64
	0.09	0.24	0.66

Panel 2. NIOSH Industry / BLS Industry

	Basic Controls	State	Ind/Occ
VAR (Lnwage)	0.25	0.24	0.21
VAR (NIOSH Ind)	47.67	39.58	18.20
VAR (BLS Ind)	59.36	58.09	35.65
COV (NIOSH Ind, BLS Ind)	24.45	21.72	1.53
COV (Lnwage, NIOSH Ind)	0.09	0.18	0.02
COV (Lnwage, BLS Ind)			
R2 Lnwage on X	0.26	0.28	0.36
R2 NIOSH Ind on X	0.04	0.21	0.63
R2 BLS Ind on X	0.06	0.08	0.43

Panel 3. NIOSH Industry / BLS Occupation

	Basic Controls	State	Ind/Occ
VAR (Lnwage)	0.24	0.24	0.21
VAR (NIOSH Ind)	39.59	32.60	13.75
VAR (BLS Occ)	154.35	152.86	115.10
COV (NIOSH Ind, BLS Occ)	21.58	18.90	2.02
COV (Lnwage, NIOSH Ind)	0.04	0.13	0.01
COV (Lnwage, BLS Occ)	-0.23	-0.18	0.06
R2 Lnwage on X	0.30	0.32	0.41
R2 NIOSH Ind on X	0.03	0.20	0.66
R2 BLS Occ on X	0.04	0.05	0.29

Table 7 cont. Covariances and Variances of Residual Estimates for Male Workers 1995 CPS Outgoing Rotation Data, NIOSH Risk Data, and BLS Risk Data

Panel 4. NIOSH Occupation / BLS Industry

	Basic Controls	State	Ind/Occ
VAR (Lnwage)	0.25	0.24	0.21
VAR (NIOSH Occ)	50.14	42.08	19.02
VAR (BLS Ind)	59.06	57.83	35.39
COV (NIOSH Occ, BLS Ind)	18.31	15.62	1.99
COV (Lnwage, NIOSH Occ)	-0.23	-0.14	0.02
COV (Lnwage, BLS Ind)	-0.07	-0.03	-0.06
R2 Lnwage on X	0.26	0.28	0.36
R2 NIOSH Occ on X	0.09	0.24	0.65
R2 BLS Ind on X	0.06	0.08	0.43

Panel 5. NIOSH Occupation / BLS Occupation

	Basic Controls	State	Ind/Occ
VAR (Lnwage)	0.25	0.24	0.21
VAR (NIOSH Occ)	59.95	51.18	23.63
VAR (BLS Occ)	175.53	173.78	131.13
COV (NIOSH Occ, BLS Occ)	39.72	36.50	5.11
COV (Lnwage, NIOSH Occ)	-0.23	-0.14	0.02
COV (Lnwage, BLS Occ)	-0.21	0.16	0.05
R2 Lnwage on X	0.26	0.28	0.36
R2 NIOSH Occ on X	0.09	0.24	0.65
R2 BLS Occ on X	0.05	0.06	0.29

Panel 6. BLS Industry / BLS Occupation

	Basic Controls	State	Ind/Occ
VAR (Lnwage)	0.25	0.24	0.21
VAR (BLS Ind)	67.78	66.45	42.19
VAR (BLS Occ)	175.02	173.33	130.92
COV (BLS Ind, BLS Occ)	43.91	42.56	16.43
COV (Lnwage, BLS Ind)	-0.08	-0.04	-0.06
COV (Lnwage, BLS Occ)	-0.21	-0.16	0.05
R2 Lnwage on X	0.26	0.28	0.36
R2 BLS Ind on X	0.06	0.08	0.43
R2 BLS Occ on X	0.05	0.05	0.28

Source: Authors' calculations.

Table 8. Comparison of Job Risk Estimates from Matched Data to NIOSH Rates
1995 CPS Outgoing Data, 1995 NIOSH Data, and 1995 BLS Data

Demographics		NIOSH	Matched Data			
			NIOSH Industry	BLS Industry	NIOSH Occupation	BLS Occupation
Sex						
	Male	7.30	4.95	6.05	5.29	7.53
	Female	0.70	2.86	2.44	2.35	2.06
Race						
	White	4.20	3.96	4.34	3.84	5.16
	Black	4.60	4.00	4.13	4.33	5.08
	Other	3.90	3.67	4.33	3.80	4.71
Age Group						
	25-34	3.70	3.95	4.45	3.99	5.49
	35.44	3.90	4.06	4.44	3.94	5.17
	45.54	4.50	3.86	4.01	3.68	4.67
	55.64	6.10	3.76	4.08	3.89	4.83
Industry Division						
Ag/For/Fishing		15.50	15.78	22.77	12.90	19.39
Mining		25.40	23.94	25.56	10.00	18.23
Construction		13.10	12.72	12.10	7.96	14.43
Manufacturing		3.60	3.63	4.10	4.39	4.49
Trans/Comm/PU		10.30	10.32	10.59	6.48	11.71
Wholesale Trade		3.50	3.62	5.84	5.11	6.00
Retail Trade		2.40	2.87	3.37	3.86	4.08
Finance/Insurance		1.20	1.08	1.37	2.17	2.49
Services		1.50	1.52	1.61	2.49	2.17
Public Admin		5.80	4.51	—	2.39	6.23
Occupation Division						
Exec/Admn/Mgr		2.50	3.55	3.55	2.40	2.52
Prof/Spec		1.40	2.48	1.66	1.40	1.24
Tech/Support		3.60	3.07	2.29	3.50	1.89
Sales		2.70	2.94	3.71	2.97	3.96
Clerical		0.70	3.73	3.25	0.63	0.81
Services		2.50	2.50	2.69	2.57	4.10
Farm/For/Fish		16.60	10.84	18.76	17.01	21.70
Crafts		8.00	6.85	7.23	7.89	9.92
Mach Operators		3.50	3.79	3.99	3.23	3.04
Transport		17.80	7.73	12.39	18.11	19.85
Laborers		10.40	6.05	7.63	10.51	11.93

Source: Authors' calculations.

Table 9. Comparison of Job Risk Estimates from Matched Data to NIOSH Rates
1985-1995 CPS Outgoing Rotation Data and NIOSH Risk Data

year	NIOSH	Matched Data NIOSH Industry	NIOSH Occupation
1985	5.80	5.34	5.23
1986	5.10	4.61	4.51
1987	5.20	4.62	4.50
1988	5.00	4.57	4.51
1989	4.80	4.53	4.48
1990	4.60	4.52	4.45
1991	4.50	4.03	3.99
1992	4.30	3.99	3.96
1993	4.40	4.01	3.95
1994	4.40	3.97	3.93
1995	4.30	3.94	3.89

Source: Authors' calculations.

Table 10A. Estimated Price of Risk for Male Workers
1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

A. Risk Measure I			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-315	-145	40.4
	(-8.53)	(-2.92)	(0.58)
B. Risk Measure II			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-376	-214	39.0
	(-9.86)	(-4.20)	(0.54)

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 52,143 observations in the men's regressions. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The risk measure I is equal to the simple mean of valid risk measures of NIOSH industry risk, NIOSH occupational risk, BLS industry risk, and BLS Occupational risk. The risk measure II is the simple mean of the same risk categories, but adding zeros to the missing values.
Source: Authors' calculations.

Table 10B. Estimated Price of Risk for Female Workers
1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

A. Risk Measure I			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-120	93.40	-60.70
	(-1.97)	(1.30)	(-0.74)
B. Risk Measure II			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-345	82.0	-38.60
	(-4.06)	(0.75)	(-0.25)

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 53,819 observations in the women's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The risk measure I is equal to the simple mean of valid risk measures of NIOSH industry risk, NIOSH occupational risk, BLS industry risk and BLS occupational risk. The risk measure II is the simple mean of the same risk categories, but adding zeros to the missing values.
Source: Authors' calculations.

Table 10C. Estimated Price of Risk for Male Workers
1995 March CPS Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

A. Risk Measure I			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-120	73.0	191
	(-2.40)	(1.41	(2.73)
B. Risk Measure II			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-174	24.70	197
	(-3.32)	(0.46)	(2.66)

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 25,621 observations in the men's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The risk measure I is equal to the simple mean of valid risk measures of NIOSH industry risk, NIOSH occupational risk, BLS industry risk, and BLS occupational risk. The risk measure II is the simple mean of the same risk categories, but adding zeros to the missing values.
Source: Authors' calculations.

Table 10D. Estimated Price of Risk for Female Workers
1995 March CPS Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

A. Risk Measure I			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	31.60	124	-56.10
	(0.35)	(1.36)	(-0.51)
B. Risk Measure II			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-129	91.80	37.30
	(-1.07)	(0.75)	(0.21)

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, and a dummy variable indicating whether the workers are under union contract or not. There are 24,758 observations in the women's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The risk measure I is equal to the simple mean of valid risk measures of NIOSH industry risk, NIOSH occupational risk, BLS industry risk, and BLS occupational risk. The risk measure II is the simple mean of the same risk categories, but adding zeros to the missing values.
Source: Authors' calculations.

Table 11A. "Post Hoc" Risk Indicator in a Multiple Regression for Male Workers 1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

Basic Regression	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
NIOSH Industry Risk/100,000	845	1118	172
	(18.25)	(23.13)	(2.63)
NIOSH Occupational Risk/100,000	-723	-527	75.20
	(-18.37)	(-13.11)	(1.37)
BLS Industry Risk/100,000	-231	-301	-297
	(-6.25)	(-8.20)	(-7.92)
BLS Occupational Risk/100,000	59.60	29.90	62.40
	(3.03)	(1.54)	(3.27)
Estimated "Risk" Effect/100,000	-7652	-5991	-237

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 215,365 observations in the men's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The "post hoc" risk indicator is extracted from the coefficients of the industry and occupational risk used simultaneously in the multiple regression. The occupational risk coefficient is weighted by the ratio of the bivariate covariance between the natural log of real wage and the occupational risk, to the covariance between log of real wage and the industry risk. The units of the "post hoc" estimator are given in terms of industry risk.
Source: Authors' calculations.

Table 11B. "Post Hoc" Risk Indicator in a Multiple Regression for Female Workers 1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

Basic Regression	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
NIOSH Industry Risk/100,000	489	600	7.28
	(7.34)	(9.00)	(0.10)
NIOSH Occupational Risk/100,000	-548	-368	53.70
	(-8.45)	(-5.69)	(0.70)
BLS Industry Risk/100,000	-526	-538	-498
	(-7.09)	(-7.38)	(-6.05)
BLS Occupational Risk/100,000	102	72.70	244
	(1.34)	(0.97)	(3.20)
Estimated "Risk" Effect/100,000	-2287	-1809	-575

Note: The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the workers are under union contract or not, and marital status. There are 182,453 observations in the women's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The "post hoc" risk indicator is extracted from the coefficients of the industry and occupational risk used simultaneously in the multiple regression. The occupational risk coefficient is weighted by the ratio of the bivariate covariance between the natural log of real wage and the occupational risk, to the covariance between log of real wage and the industry risk. The units of the "post hoc" estimator are given in terms of industry risk.
Source: Authors' calculations.

Table 11C. "Post Hoc" Risk Indicator in a Multiple Regression for Male Workers*
1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

Basic Regression	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
NIOSH Industry Risk/100,000	630	863	154
	(18.84)	(24.85)	(3.59)
NIOSH Occupational Risk/100,000	-561	-350	105
	(-17.62)	(-10.66)	(2.41)
BLS Industry Risk/100,000	-270	-322	-329
	(-8.77)	(-10.55)	(-10.06)
BLS Occupational Risk/100,000	42.90	12.70	118
	(2.45)	(0.74)	(6.92)
Estimated "Risk" Effect/100,000	-5192	-3810	-115

Note: *Adding zero value for missing risk categories. The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the worker is under union contract or not, and marital status. There are 353,950 observations in the men's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The "post hoc" risk indicator is extracted from the coefficients of the industry and occupational risk used simultaneously in the multiple regression. The occupational risk coefficient is weighted by the ratio of the bivariate covariance between the natural log of real wage and the occupational risk, to the covariance between log of real wage and the industry risk. The units of the "post hoc" estimator are given in terms of industry risk.
Source: Authors; calculations.

Table 11D. "Post Hoc" Risk Indicator in a Multiple Regression for Female Workers*
1995 CPS Outgoing Rotation Data, 1995 NIOSH Risk Data, and 1995 BLS Risk Data

Basic Regression	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
NIOSH Industry Risk/100,000	425	540	1.23
	(10.66)	(13.41)	(0.03)
NIOSH Occupational Risk/100,000	-506	-247	37.40
	(-10.11)	(-5.45)	(0.63)
BLS Industry Risk/100,000	-423	-428	-359
	(-7.62)	(-7.48)	(-5.95)
BLS Occupational Risk/100,000	259	214	534
	(4.08)	(3.43)	(8.44)
Estimated "Risk" Effect/100,000	-1287	-907	-134

Note: *Adding zero value for missing risk categories. The dependent variable is the natural log of the worker's real wage. For basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the worker is under union contract or not, marital status. There are 337,530 observations in the women's regression. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The "post hoc" risk indicator is extracted from the coefficients of the industry and occupational risk used simultaneously in the multiple regression. The occupational risk coefficient is weighted by the ratio of the bivariate covariance between the natural log of real wage and the occupational risk, to the covariance between log of real wage and the industry risk. The units of the "post hoc" estimator are given in terms of industry risk.
Source: Authors' calculations.

Table 12. IV Estimation for Estimated Price of Risk

Panel 1: March CPS and NIOSH Risk: Males, 1985-1995				
	OLS Estimation	**OLS Estimation**	**IV Estimation**	**IV Estimation**
Basic Controls	yes	yes	yes	yes
State	yes	yes	yes	yes
Industry/Occupation	yes	yes	yes	yes
	Industry	Occupation	Industry	Occupation
Risk/100000	107	20.34	101	285
	(4.72)	(1.72)	(1.22)	(4.71)
Instrument	no	no	yes Occupation Risk	yes Industry Risk

Panel 2: ORG and NIOSH Risk: Males, 1985-1995				
	OLS Estimation	**OLS Estimation**	**IV Estimation**	**IV Estimation**
Basic Controls	yes	yes	yes	yes
State	yes	yes	yes	yes
Industry/Occupation	yes	yes	yes	yes
	Industry	Occupation	Industry	Occupation
Risk/100000	84.0	105	293	201
	(7.75)	(10.58)	(10.50)	(13.80)
Instrument	no	no	yes Occupation Risk	yes Industry Risk

Note: For all regressions, the dependent variable is the natural log of the worker's real wage. For the basic regression, the independent variables include a quartic in the worker's age, a vector of dummy variables that control for the worker's education, a vector of dummy variables indicating whether the worker is Hispanic, Asian, African American, or other race, a dummy variable indicating whether the worker is under a union contract or not, and marital status. Workers are aged 25 to 60 inclusive. T-statistics are given in parentheses. The Instrumental Variable (IV) estimation was performed using for the variable industry risk the instrument occupation risk, and for the variable occupation risk the instrument industry risk.
Source: Authors' calculations.

Table 13. Correlation of NIOSH Job Risk Measures and Air Force Qualification Test (AFQT) Score, Illegal Drug Use, and Education, NLSY Data

Panel A. NIOSH Industry / NIOSH Occupation and Other variables	Education	AFQT Score (Age demeaned)	Illegal drug use
Industry job risk	-0.15	-0.10	0.09
Occupation job risk	-0.26	-0.21	0.05
Panel B. Residual Correlations			
Industry job risk	0.02	-0.01	0.06
Occupation job risk	0.01	-0.03	-0.00

Notes: Illegal drug use data are from 1984 and include heroine, cocaine, and marijuana/hashish. In 1984, respondents were aged 19 to 26 inclusive. AFQT-job risk and education-job risk correlations are from 1990 when respondents were aged 25 to 32 inclusive.
Source: Authors' calculation.

Table 14A. Fix-Effect Estimated Price for Male Workers
1986 to 1993 NLS Data and NIOSH Risk Data

A. Industrial Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-47.00	-19.40	-110
	(-0.41)	(-0.17)	(-0.65)
Mean Risk	6.18	6.18	6.18
B. Occupation Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	-78.90	-74.70	5.46
	(-1.51)	(-1.42)	(-0.08)
Mean Risk	7.66	7.66	7.66

Table 14B. Fix-Effect Estimated Price for Female Workers
1986 to 1993 NLS Data and NIOSH Risk Data

A. Industrial Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	266	288	83.20
	(1.75)	(1.87)	(0.39)
Mean Risk	3.61	3.61	3.61
B. Occupation Risk			
Basic Controls	yes	yes	yes
State	no	yes	yes
Industry/Occupation	no	no	yes
Risk/100,000	123	116	74.70
	(1.11)	(1.05)	(0.39)
Mean Risk	3.64	3.64	3.64

Note: For the basic regression, the independent variables include a quartic in the worker's age, education level, union, experience, tenure status, test scores, race/ethnicity category, and marital status. Workers are aged 25 to 60. T-statistics are given in parentheses. There are 20,338 observations in the male regressions and 19,272 observations in the female regressions. There is a fixed-effect for individual in each specification.
Source: Authors' calculations.