STABILITY OF ESTIMATES OF THE COMPENSATION FOR DANGER*

G. R. ARABSHEIBANI

(Royal Holloway College, University of London)

A. MARIN

(London School of Economics)

ABSTRACT

Estimates of the extra earnings for jobs with higher risks of death can be used in cost-benefit studies involving risk changes. Because of this use, the magnitude and stability of the estimated coefficient are important. Part of the current study closely reproduces a widely quoted 1982 study by Marin & Psacharopoulos to check on the stability. We also examine the robustness of the estimate to the inclusion/exclusion of non-fatal risks and other relevant characteristics. While the magnitude of the co-efficient has increased threefold from the earlier study, the coefficient is robust to other changes in the specification.

There could be selectivity bias in the estimates of the extra return because people can select their occupation on the basis of its riskiness. Our findings suggest that one common technique to deal with selectivity bias in a continuous variable can give unreliable results in practice.

Keywords: Cost-Benefit, Value of life, Labour market, Selectivity bias

JEL Numbers: C24, D61, H43, J17, J28, J31

* We are grateful to Franco Insalaco for exceptionally thorough research assistance.
This study was funded by the ESRC . (Award no. R000221162).
Data was provided by the ESRC Data Archive and OPCS (now ONS), some via the Manchester Computer Centre. We are grateful to them, and to the ED, for provision and help with the data, especially to Hilary Beedham, Frances Drever and John Roberts for particular assistance and advice. We are also grateful to Jonathan Wadsworth and to participants at seminars at LSE and the 1997 EEEG Meetings for helpful comments. The usual disclaimer remains necessary.

1. INTRODUCTION

In the past two decades, starting with the work of Smith (1973, 1976) and Thaler & Rosen (1975) there have been a series of empirical studies examining whether, and how much, workers are paid extra for working in more dangerous jobs. A survey of some of the earlier studies, together with an exploration of reasons for differing results, can be found in Marin and Psacharopoulos (1982). More recent surveys can be found in Jones-Lee (1989), Fisher et al. (1989) and Viscusi (1993).

Part of the reason for the interest in such studies is a wish to test the (Adam) Smithian theory of compensating differentials for undesirable characteristics of jobs. Because it seems plausible to suggest that neo-classical labour market theory is least likely to apply when wages are heavily influenced by collective bargaining, particular attention has been paid by some researchers to the effects of unionisation on the compensating differential for job risk, for example Sandy & Elliot (1996) for the UK.[1]

Another reason for the interest is that the results of such studies can provide the basis for an important input into cost-benefit studies. Many government projects, or government regulations of firms' actions, involve a change in the risks of death or non-fatal injury/illness faced by some individuals. There is now wide acceptance among economists (though with exceptions, e.g. Broome [1978, 1985]), that the best way to evaluate such changes is by the willingness-to-pay (WTP) for changes in the risks by those affected by them.

The extra wages received for extra work-related risks can be viewed as the compensating variation for accepting the risks. The empirical estimates of the wage premia for danger can thus be used in cost-benefit studies to evaluate risk changes - at least where the risks are of the same kind. For example if the risk of death or injury from accidents is the measure used in the labour market study and is also the type of risk arising in the project which is being evaluated.

Because of the use in cost-benefit studies, an important consideration is not only whether there is a statistically significant positive coefficient on the risk variable in the wage equation but also the magnitude of the estimated coefficient. The 1982 study by Marin and Psacharopoulos has been widely quoted in the literature, especially (but not only) for the UK, e.g. Jones-Lee (1989), Fisher et.al. (1989). One aim of the current study is to reproduce as closely as possible the earlier study on more recent data, and to check on the stability of the relevant coefficient.

Most labour market studies of WTP, with a few mainly US exceptions (see the Viscusi [1993] survey, but also Siebert and Wei [1994] for the UK), have only looked at the risk of death. It has been suggested that such estimates are unreliable because they may be picking-up the effects of non-fatal job related illness/injury. In the current study we also examine the robustness of the estimates on the compensation for risks of fatal injury if non-fatal injury rates are also included, as well as the robustness to the inclusion/exclusion of some other relevant characteristics, e.g. region or marital status.

The final aspect we examine are the effects on the estimates of allowing for the endogeneity of some of the relevant labour market characteristics. We look at the choice of union membership, but particularly concentrate on the possibility that because people can select their occupation on the basis of its riskiness, this could lead to selectivity bias in the estimates of the extra return to risky jobs.

2. MORE DETAILED BACKGROUND

A detailed description of the relationship to be expected between riskiness and earnings can be found in several sources.[2] For simplicity, the explanation is phrased here in terms of the risk of premature death, but the same argument applies *mutatis mutandis* to non-fatal risks.

The basic hypothesis is that an individual will be prepared to trade-off extra income for an increased probability of death during the coming period. In general, different workers will have different attitudes towards risk, reflecting both tastes as usually understood, perhaps family status and dexterity, calmness and other aspects of the ability to deal with risky situations. As a result, faced by a market opportunity locus, different people will choose different combinations of risk and earnings. As far as firms are concerned, it seems reasonable to postulate that firms can reduce the risks faced by their employees, at a cost. The actual cost of improved safety will differ between firms, e.g. according to industry or technology. The market opportunity locus is established by the interaction of the supply of workers at different degrees of risk and the demand for them. - i.e. resulting from workers' preference functions and firms' profit maximisation.

Having estimated the slope of the frontier between risk and earnings, it is common in the literature to use the results to give “the value of life” or, slightly less misleadingly “the value of a statistical life”.

One route is via expected utility theory.[3] However, given the doubts over both the predictive and normative validity of expected utility theory, it is perhaps more straightforward to go directly to the notion of a "value" per expected (as in the statistical sense) life lost or saved.

If we consider the meaning of the slope of the wage/risk relationship for the employees, we can interpret the “value of life” from a viewpoint which fits into the standard cost-benefit framework. If we measure the probability of death as, say, so many in a thousand per year, and the slope of the earnings function with respect to risk, is, say, £*x* per year, then the payment of £*x* per year required by a worker to accept an increased risk of death of 1/1000 per year, is the compensating variation for this increased risk. A group of 1000 workers would together have the sum of each of their compensating variations adding to £1000 *x*. However, such a group of 1000 workers, each with an increase of 1/1000 in risk of death per year, would suffer one extra death on average. Thus an extra expected life lost would have a sum of compensating variations equal to £1000 *x*, and it is the sum of compensating variations which is the usual criterion used in project appraisal. Therefore the relationship between earnings and the probability of death provides a figure for the “value of life” that is suitable when each person affected by the project has an increased (or decreased) chance of dying, even though it would not be correct if any individual faced certain death as a result of the project.

The £1000 *x* in the above example is the value of a life, not of a life per year. This stock measure is independent of the time period, provided only that both earnings and probabilities of death are measured for the same unit of time. For example, if weekly wages and weekly risks were used in an earnings function instead of annual, the estimated coefficient would be unaltered, since both the independent and the dependent variable would have been divided by 52.
A similar derivation from the firm's viewpoint would give the same figure as the "cost" of an extra death (ignoring the difference between pre- and post- tax wages).

In the approach sketched out here, the notion of the cost or "value" of an expected death or of an expected life saved is just a way of summarising the cost or compensating variation for a small change in the risk faced by an individual. It does **not** state that the value of any particular life is £1000 *x*. Given that some criticisms of the practical use of such figures are at least partly based on asserting the contrary, [4] and since the application to cost-benefit does not require the concept of the “value of (statistical) life", it might be preferable to stop using it. Its main advantage is that it is an easy short-cut way to compare the implications of different studies for cost-benefit applications. However, for studies using regressions which do not include inter-actions between risk and other variables the concept is otiose, as a direct comparison of the risk coefficients is all that is needed. If we define Y as income and p as the probability of death through job related causes, then even when interactions are included, a direct comparison of $\partial Y/\partial p$ will suffice, evaluated at the means of the interacting variables, if necessary.[5] The only requirement is that results are converted to those that would follow if risk were measured in the same units, e.g. per thousand per year.

As can be seen by looking at the three more recent articles mentioned in the Introduction as including surveys of results, such articles typically list results by the “value of life”. The values cover a fairly wide range; e.g. those listed in Viscusi (1993) are from \$0.7 million to \$16.2 million (1990 US \$). Some of the variability can be explained (e.g. see Marin & Psacharopoulos or Viscusi). However, even allowing for the part that can be explained, because studies vary in so many ways, it is difficult to assess how much of the variability is due to systematic differences, such as in data definitions, and how much to lack of robustness of the estimates to sampling. The present study therefore tries to reproduce the earlier Marin & Psacharopoulos study as closely as possible in terms of the variables included and their precise definitions but using data for the following decade. In this way the temporal stability of the size and significance of the coefficients on risk can be assessed.

In view of our comments above, the comparison is of the coefficients of the risk variables in an earnings equation - in practice, with the log of earnings as the dependent variable. Since our risk measures are in per thousand per year, readers who wish to covert the results into a “value of life” need only multiply the coefficient (for regressions with no interactions) by 1000 times the average income for the year/country in whose units they wish to have the “value of life”. [6]

Less attention has been paid to the “value” of a non-fatal injury/illness. In addition to the usefulness of estimating the compensating differential for job related illness/injury as an input into cost-benefit studies, it has also been suggested that the failure to include non-fatal risk measures can bias the estimates of the effects of fatal risk on earnings. If estimates of the compensating differential for fatal risk are "picking up" the effects of omitted non-fatal riskiness, then at the very least the estimates should only be used in cost-benefit analysis of projects where the fatal/non-fatal relationship is likely to be similar to that in job related risks. The effects of including non-fatal risks are examined in Section 6.

3. ESTIMATION PROCEDURES

Most of the empirical studies investigating whether labour markets reward more dangerous jobs have used a "hedonic wage equation" of the form.

$$\log Y_i = a_1 p_i + a_2 q_i + X_i \beta + \varepsilon_{i\,1} \qquad (1)$$

Where Y_i is earnings of individual i,

p_i , q_i job related fatal and non-fatal risk faced by i (q_i often omitted)

X_i a vector of other relevant individual and job characteristics (plus constant).

X_i usually includes the variables indicated by the human capital approach, such as years of schooling and experience. Other variables differ from study to study. They may include union membership, race, marital status and number of children from the personal characteristics side. On the job side they may include firm size, industry dummies or union coverage. Regional dummies may also be included.

Equation (1) can be viewed in two ways. One way is to view it simply as an earnings equation. In this light, it is a reduced form of the labour supply and demand functions. It has the standard exogenous variables on the right hand side, but with the addition of the risk variable (s). It follows the Mincerian human-capital approach which implies that an equation with years of schooling as an explanatory variable should have the log of earnings as the dependent variable. Later in this Section we consider the potential problems if some of the right-hand side variables in Equation (1) are selected by workers (or employers) rather than being exogenous.

The other way of viewing equation (1) follows from the theory mentioned in the previous Section, where it is the "frontier" constituting the market opportunity locus between wages and risk along which individual workers and firms make their choices. The position of the frontier is determined by a kind of matching process between workers and firms for jobs with varying levels of risk (see

Rosen (1974), (1986) or Kahn & Lang (1988)). It is therefore determined by the variables that enter into firms' demand for labour functions (where the labour input consists not only of the hours worked but the risk accepted by workers) and into workers' supply of labour functions (similarly defined), plus possibly other variables that reflect the matching process itself.

In the absence of a strong basis for the functional forms of the supply and demand functions of types of firms and workers and for the distribution of the types of firms and workers, the choice of functional form for the hedonic wage equation is arbitrary. The log of earnings, as in (1), has generally been chosen as dependent variable. One justification in this framework might be that it is known that incomes tend to have a log-normal distribution, and therefore using the log of earnings reduces the chance of problems arising from the maintained hypothesis of normal ε_i .

A few US studies have tried to estimate structural hedonic systems, distinguishing demand, supply and the hedonic wage frontier (e.g. Khan & Lang (1988), Biddle & Zarkin (1988), Viscusi & Moore (1989)). A major problem is finding identifying restrictions. As indicated in the paragraphs above, the attempts to omit variables from the frontier that appear in the supply (utility) functions seem unconvincing.[7]

Typically regional dummies are included in the hedonic wage frontier to identify the supply of labour function, on the assumption that workers' characteristics or preferences are independent of region. However, we think that this is extremely problematic. For example, among the explanations for the "wage curve" of Blanchflower & Oswald (1994) are those that involve local levels of unemployment affecting workers' individual or collectively bargained wage demands for taking a job.

Partly because of the lack of convincing identifying restrictions, we limit ourselves in this study to reduced form hedonic wage equations. Using reduced form equations also enables us to by-pass the possibly arbitrary choice of representing utility functions, which is a feature of some of the few studies which use a structural approach, e.g. Viscusi & Moore (1989).[8]

Whichever of the ways that equation (1) is viewed, there is still the possibility of bias if some of the RHS variables are themselves selected by workers and employers in the current period.

One such possible source of selectivity bias is the choice of union membership, if this is among the exogenous variables. In this case it is possible to use the standard Heckman-Lee two stage estimates to correct for selectivity bias (Maddala 1983).[9] Given the purposes of the current study, we are interested in whether self-selectivity of union membership can affect the compensation of union and union members for risk, rather than in any effects on the estimate of the return to union membership itself.

Another important possible source of bias is the choice by workers of the degree of riskiness of occupation. In this case, one argument is that the endogeneity of the choice may lead to an under-estimate of the compensating variation for risk. Intuitively, those workers with the least aversion to danger will match with the most dangerous jobs (i.e. where it is most expensive for firms to provide greater safety). Such workers would require less compensation for the extra risk than would the average worker, and thus the estimate of the necessary compensating variation for the labour force as a whole will be biased downwards.[10]

More formally, the problem can be seen as a standard selectivity bias problem.[11] We posit that workers choose the riskiness of the occupations that they will enter by a process that can be summarised as in equations (2) and (3)

$$p_i = Z_i \gamma + \varepsilon_{2i} \tag{2}$$

$$q_i = Z_i \theta + \varepsilon_{3i} \tag{3}$$

The Z_i variables will include the X_i of equation (1), because the choice of occupation, and its associated risks, will depend on the expected income. Z_i will also include non-labour income (wealth) and possibly other personal characteristics that do not directly help to determine income in equation (1).

In general, we should not expect there to be a zero covariance between the ε_{1i} and ε_{2i} or ε_{3i}. The error term ε_{1i} can give rise to an income effect in the choice of job related risk. Unobserved personal characteristics affecting the choice of the degree of risk may also affect earnings within jobs with any degree of risk.

If people differ according to their ability to handle danger, then this will affect their earnings related to danger. The last point could be expressed either by making a_1 and a_2 in equation (1) dependent on *i* or, equivalently, by adding terms to (1) to re-write the correct hedonic wage equation as:

$$\text{Log } Y_i = a_1 p_i + a_2 q_i + X_i \beta + u_{1i} p_i + u_{2i} q_i + \varepsilon_{4i} \qquad (4)$$

The selectivity bias occurs if equation (4) is estimated by OLS, because ε_{4i} , u_{1i} and u_{2i} help to determine p_i and q_i. This causation means that there is a non-zero expected value of the error term in equation (4); i.e. $E[u_{1i} p_i + u_{2i} q_i + \varepsilon_{4i}] \neq 0$. The estimation of (4) by OLS is therefore inconsistent.

The standard Heckman-Lee selectivity correction only applies to binary (or multinominal) discrete choices, not to continuous choice variables, such as risks. However Garen (1984, 1988) has shown that consistent estimates can be obtained by first obtaining $\underline{\triangle}_{2i}$ and $\underline{\triangle}_{3i}$ from OLS estimation of equations (2) and (3). These are then used to estimate the wage function:

$$\text{Log } Y_i = a_1 p_i + a_2 q_i + X_i \beta + c_1 \underline{\triangle}_{2i} + c_2 \underline{\triangle}_{3i} + c_3 \underline{\triangle}_{2i} p_i + c_4 \underline{\triangle}_{2i} q_i + c_5 \underline{\triangle}_{3i} p_i + c_6 \underline{\triangle}_{3i} q_i + \varepsilon_{5i} \qquad (5)$$

As shown by Garen, the estimates of the a_1, a_2, β and the c's in (5) are consistent.

In Section 7, we investigate the effects on the estimates of the compensation for risk of allowing for endogenous selection of risk.

4. DATA

The data sources used for income and personal characteristics were the General Household Survey (GHS), for various years in the 1980s. We restricted the sample to male full-time employees in England and Wales between the ages 20-64. For some purposes we also used the income and other characteristics of other members of the same household or the respondent's parents.

As explained in Marin & Psacharopoulos, the analysis was restricted to males because there are major problems with the occupational classification of females in the mortality data. Those under 20 years were omitted because some of the mortality data is not published for this age.

A major series of our problems with the data stem from the changing set of questions used in the GHS. One such change is in the degree of detail of the occupational groups to which respondents are assigned. For 1981 onwards, the 1980 Classification of Occupations is the base, but before 1985 only a broader 16 group classification (KOS) could be calculated. From 1985 it is possible to derive the 166 three-digit occupation units (161 excluding inadequately classified groups).

This degree of detail turned out to be crucial for the analyses. The occupational mortality data (described later) is available for the more detailed occupational classifications. When we experimented with aggregating it up to the 16 KOS group level and comparing it to the 161 group level, the results were clearly much less well defined in terms of the significance levels of the mortality variable - the standard error of the coefficient was larger while nearly always the size of the coefficient was smaller. In some sub-samples the hypothesis of a positive compensating differential in earnings would have been rejected at conventional significance levels for the broader occupational grouping, but not rejected for the preferred narrower grouping.[12] (In 1983, for which the broad grouping was the only one available, the sign, and not just the significance, of the fatal accident variable was unstable and varied with the sample and the other variables included in the regression.)

The other major advantage of using the GHS for 1985 onwards is that the finer level of occupational classification meant that it was possible to update the Goldthorpe and Hope (1974) classification of occupational desirability, by using data from the 1981 Census on cross-classification by 1970 and 1980 occupational units which was provided to us by the OPCS . This enabled a direct comparison between the results of earlier work and the current study. Without this index for 1980s data, assessing the stability over time of the compensating differentials would have been far less satisfactory and convincing. Even using the dummy variable Social Class classification of Erikson - Goldthorpe would have been impossible without the data in the 1985 GHS.

However, some other job characteristics were not collected by the GHS for 1985 onwards. The ones we most would have liked to have, but which were not on the questionnaires for these years, concerned union membership and coverage. These questions were asked for 1983 but not subsequently. Ideally we had wanted to endogenise the choice of union membership as well as of risk taking. Also, for comparability with the earlier study (and other work such as Sandy and Elliot

[1996]), it was necessary to allow for the effects of unions on the compensating differentials for risk. For the latter problem of allowing for the effects of trade unions on the rewards for risk, we tested two alternatives.

FIRST For 1983 we constructed a cross-classification of the 10 one-digit SIC and the KOS (the broad 16 group occupational classification mentioned above), then calculated the percentage of union members in each of the 160 SIC/KOS groups. The percentage 1983 membership was used as a measure of union strength in the same SIC/KOS group as in 1985. Despite known changes in aggregate union membership over the 1980s, the 1983 strength by SIC/KOS should at least be a reasonable instrument for 1985 strength.

SECOND A probit equation was fitted to the 1983 sample to explain the 1-0 dummy of union membership. The estimated coefficients on the RHS variables in the 1983 probit were then applied to the same variables for people in the 1985 sample, to predict their union membership. This predicted 1-0 union dummy could then be used as an independent variable in the 1985 wage equations. The estimated probit coefficients could also be used in "selectivity bias" corrections for 1985.

The two methods have their own advantages and drawbacks as a way around the problem. The first, as stated, can be seen as an instrument/proxy for union strength in the relevant bargaining units. It is also genuinely exogenous for the individual worker. One disadvantage is that it may be that it is whether the individual is a union-member that is relevant for his earnings, including the compensating differential, rather than general union strength in the bargaining group. The other disadvantage is that selectivity bias cannot be tested. The second method allows for the individual's membership both to be relevant and to be endogenous. The disadvantage is that since the probit approach estimates union membership by a cut-off point, the actual magnitude of the coefficients matters for the prediction. Instability of the size of the coefficients between 1983 and 1985 would lead to mis-categorization of some 1985 individuals between union membership and non-membership.

Insofar as it turned out that the two different methods gave rather similar coefficients on the size of the compensating differential for risk, this gave confidence in the reliability of the results.

The other sets of data concerned mortality and non-fatal injuries. The former came from the OPCS Occupational Mortality Decennial Survey for 1979-83, some from the published tables, some unpublished. The published tables on deaths from accidents at work had several problems for our purposes. One (which we overcame with the help of a special 1981 table provided by OPCS) is that the numbers of fatal injuries were for England and Wales, while the number of people working in each occupation group were for Great Britain. Using the specially provided table for the numbers working in each age/occupation group, we were able to construct consistent fatal injury ratios for England and Wales. All our analyses are therefore for males in England and Wales.

The OPCS also provided us with a Table showing the break-down of fatal injuries at work sub-divided by age-group. Although we did not actually use the age-group specific deaths for each occupation group (only for all groups) in constructing our measure of expected deaths, nevertheless we noticed that the totals in some groups differed slightly from those in the published Tables. The discrepancies could not be explained. We experimented initially with both sets - the figures in the specially provided data gave slightly higher and more significant results. On the basis that if the small discrepancies between the data sources were purely random, then the errors-in-variable effects would reduce the coefficient, we subsequently used the totals from the specially provided data.

For the non-fatal injuries we used questions in the 1987-89 GHS on accidents at place of work and during hours of work, that resulted in a visit to a doctor or a hospital.

We followed the same procedure as in Marin & Psacharopoulos to transform the mortality and the non-fatal injury into measures of risk for each of the 161 occupational groups. The measure is the difference (as a rate per thousand workers per year) between the actual numbers of deaths/injuries and the number that would have been expected given the age structure of the occupational group and the overall death/injury rates for each age. That is: for each occupation group, the measure is the observed number of deaths (or injuries) in the group minus the number of deaths (or injuries respectively) expected given the number of employees of each age in the group, divided by total employment in the group (and the whole then multiplied by one thousand to give a rate per thousand).

5. <u>REPLICATION OF EARLIER STUDY ON MORE RECENT DATA</u>

As stated above, a major aim of the research was to repeat the formulation of the equations reported in Marin and Psacharopoulos (1982), but with an up-dated data set. In particular, because it has been widely quoted in cost-benefit studies, a major aim was to see if there was stability in the size of the coefficient on risk of death. As also discussed above, within the common specification in such studies in which the dependent variable is the log of income (and seen most straightforwardly if there are no interactions between risk and other variables) the "value of life" from different studies can be compared just from their coefficients on risk - provided only that, if necessary, the coefficient is converted to what it would be if risk were measured in the same units, e.g. deaths per thousand per year. We do not need to bother with the unit of measurement of income in order to compare studies.

For ease of comparison the main regressions from Marin & Psacharopoulos are reproduced in Table 1.
In the current study, the wage is weekly (from pay-period earnings in the GHS), and therefore there is no weeks worked variable. As mentioned above, the current study also uses different measures of union effects. As it would be tedious to show all the variants, for simplicity in this Section we mainly use the USK variable, which is union coverage in the SIC/KOS group. Results are extremely similar using the predicted union membership dummy.

As before, and as expected (on the arguments in Marin & Psacharopoulos), the measure of <u>total</u> job related mortality, called GENRISK, does not give very satisfactory results. It is worth mentioning that when an interaction between the union variable and GENRISK is included, there are always opposite signs on the coefficients and the interaction terms, which is something not adequately examined in previous studies. This is probably related to the very high positive covariance between GENRISK and the interaction - the correlation coefficient is over +0.9 for the whole sample and two of the sub-groups and +0.83 for the managerial/professional sub-groups. However, without interaction, the results are weak in terms of conventional significance levels.

Thus we should place little reliance on GENRISK either as the basis for cost-benefit studies or as evidence for the existence of compensating differentials for job related mortality. The results in this study therefore strengthen the previous scepticism which was based on both a priori arguments and the empirical evidence. We, therefore, do not bother to show these results (though they can be provided on request).

The basic results for the risk of accidental death at work variable, ACCRISK, are shown in Table 2 . As contrasted with the original results in Marin & Psacharopoulos, and considering the versions without any union/risk interaction, the coefficient on ACCRISK using 1985 data is approximately 3 - 4 times as great for the whole sample, and also for manual workers and for non-manual workers - for the managerial/professional group the discrepancy is slightly greater. The differences are several standard-errors (e.g. nearly five times the S.E. for ACCRISK in column 1 of Table 2). It is also worth noting that this time the coefficients for the non-manual and managerial/professional sub-groups are clearly significant at conventional levels. (Allowing for heteroscedasticity via the Huber/White correction leads to only trivial changes in the "t" values. Even applying the correction for group averaging effects suggested by Moulton (1986) to the ACCRISK variable did not affect conclusions about significance at conventional levels[13].)

Again, we do not show all the variants here, but the results on the sizes and significance of the ACCRISK coefficients are fairly robust to the inclusion of extra dummies such as job tenure, region etc., to the replacement of the Goldthorpe - Hope index by the social class dummies, or to the use of predicted individual union membership instead of the SIC/KOS group union coverage. Thus, insofar as one is interested in the question of whether UK data show that employees are paid compensating differentials for working in jobs according to the risk of dying from a job-related fatal accident, the current study strongly reinforces the earlier findings that UK labour markets do confirm to Smithian predictions. Furthermore this finding here applies even to managerial/professional and non-manual employees (groups who have been omitted in some other studies of the UK in the 1980s), which was unclear in the Marin & Psacharopoulos study.

However, insofar as the interest is in the precise size of the differential as an input into cost-benefit studies, the increase in the magnitude of the relevant coefficient is less re-assuring. It is possible that the increase reflects higher real income opportunities. Theory would predict that safety is a superior good. The log-linear earnings-risk specification ensures that the compensating differential will increase proportionately with levels of income, but a stronger income effect than allowed for in the estimating specification would show up as a higher coefficient when fitted to a data set with higher real incomes. The three-to-four fold magnitude of the increase makes this implausible as the <u>sole</u> explanation. This aspect of the findings is one which we hope to explore further in the future.

As discussed in the earlier study, in assessing the coefficients in the non-manual and managerial/professional groups it should be noted the higher coefficients reflect groups where safety is relatively cheap to achieve and the higher incomes increase the costs to firms of inducing employees to accept danger. Many of these occupation groups are "corner" solutions in that the actual number of deaths is zero; although ACCRISK < 0 for these groups since a positive number of fatal injuries would be expected from economy wide figures.

There is another aspect where the results confirm the earlier study, but where on reflection we feel that the earlier interpretation of the results was insufficient. This concerns the Goldthorpe - Hope index. As in the earlier study, the reason it was included was to capture other aspects of job desirability which earn compensating differentials. However, the positive coefficient contradicts this interpretation. Since the index gives a higher rating to more attractive jobs, on a compensating differentials interpretation the coefficient should be negative. Thus, we now interpret the index as picking-up attributes and abilities required for different jobs which are not captured just by schooling (or other measures of educational attainment, when these were used instead of schooling). As it happens, the size and significance of the coefficients on the risk variable (and on the other variables in the basic equations) are insensitive to whether one uses the cardinal Goldthorpe - Hope index or the social class dummies. This strengthens our current interpretation of the index or the dummies as proxying for abilities/attributes which segment the labour force to an extent, but which are not the subject of questions in the GHS (or similar surveys).

The new interpretation is strengthened by investigating (this time) what happens when both the Goldthorpe - Hope index and the alternative class dummies are omitted (Table 3). The risk variable remains positive, though with reduced significance levels, in all of the three sub-groups but become insignificantly negative for the more heterogeneous whole sample. There is also somewhat more difference between the union measures in this case, which are therefore both shown in Table 3.

The final noteworthy part of the replication concerns the possible interaction of unionisation and of the risk of fatal injury in determining earnings. The earlier study consistently found a negative interaction - i.e. that higher unionisation reduced the compensation for danger (though the interaction was not significant at conventional levels for the ACCRISK measure of risk). This went against US studies up till that time and suggested an interesting difference in the behaviour of unions between the US and UK[14]. The issue has been interesting enough to be studied further by Sandy & Elliot and Siebert & Wei for the UK.

In the current study, when an interaction term was added to the regressions the results were unstable, not only for the interaction term itself but also for the estimated coefficient on the risk variable. Depending on the sample (whole, or each subgroup) and also to some extent on whether the Goldthorpe - Hope index or the Social Class dummies were used and on which of the union measures was used, in some cases the coefficient on risk became very large while in some others it actually became negative. The standard errors of the coefficient estimates were also volatile, but not in a systematic way; so that sometimes the "t" became much larger, other times it was very low. The same applied to the sign and significance of the interaction term.

The likeliest interpretation, we think, is mainly the fairly high correlation between the risk variable and the interaction, e.g. for the whole sample there was a 0.75 correlation between ACCRISK and the interaction with the predicted union membership dummy; (but only -0.06 correlation for the interaction and the union dummy). For the SIC/KOS union variable, the correlation between ACCRISK and the interaction was 0.9. There is more variability in the risk variable than in the unionisation ones. This explanation is reinforced by the fact that in all four of the cases using the SIC/KOS coverage as the union variable, ACCRISK and the interaction took opposite signs; while the same also held for the manual sample using the 0-1 predicted union membership dummy, (in the other three cases for this union variable, both risk and interaction were positive, but ACCRISK was smaller than without the interaction, and insignificant in two of the cases).

An alternative way of seeing the effects of unions (not used in the earlier study) is by fitting completely separate regressions for those predicted to be union members and those predicted to be non-union (i.e. predicted by the coefficients in the 1983 probit). If this is done with membership treated as exogenous (equivalent to assuming zero selectivity bias), the coefficients on ACCRISK are very close with one exception. E.g. for the whole sample the coefficients on ACCRISK were 1.0753 for the union members and 1.0648 for the non-members. The one exception was for manual workers, where the coefficient on ACCRISK was 0.5428 (S.E.= 0.2415) for union members and 1.4686 (S.E.= 0.3427) for non-members. (For this group, when trying an interaction term between the union membership dummy and risk, it had been significantly negative: the only significant interaction for membership and risk - though the above caveat applies since the correlation between the interaction term and risk for manual workers was also 0.75).

The conclusion would seem to be that in the current data set, there is no reliable evidence for an interaction between risk and union strength and very limited evidence for an interaction between risk and union membership. The one exception found, for manual workers, goes in the same direction as in the earlier study for Great Britain (and thus in the opposite direction to most US studies).

As mentioned, we were limited to some extent in our investigation of the effects of unionisation by the unavailability of a union membership question in the GHS after 1983. To predict union membership for 1985 required fitting a probit equation to the 1983 data, and this is also part of the 2-stage Heckman approach to endogenising the effects of union membership in a way that allows for self-selectivity. (Testing for self-selectivity of union membership is an extension, not just a replication of the earlier work.)

Various personal, educational and job related characteristics were used as explanatory variables in the probit. Within the general specification, particular changes sometimes made slight differences to the significance of particular variables, e.g. use of educational level dummies rather than years of schooling or existence of dependent children rather than their number, but had little effect on the overall goodness-of-fit as measured by the number of correct predictions of union/non-union membership in the sample.

As far as use in a 1985 ordinary least-squares equation is concerned, both the predicted union membership variable and our SIC/KOS measure seemed reasonable in themselves as determinants of earnings. If applied to the basic regression, the pattern is as shown in Tables 2&4. For both the whole sample and manual workers, the coefficients on the union membership dummy give lower coefficients than the union average measure (see, for example, Booth [1995]). For the non-manual and the managerial/professional groups, the results are less clear-cut. The negative coefficients on the latter group (though insignificant for the membership dummy) may reflect the likelihood that managers and professional employees who belong to trade unions may well be in different sorts of job slots than those who do not.[15]

When the union membership dummy was endogenised and allowance made for possible selectivity bias, there was little effect on the ACCRISK variable. As with exogenous union membership, there was also only a small different between its size for members and non-members.

The selectivity bias measure ("Inverse Mills Ratio") was sometimes significant in 1983 for union membership but always insignificant for non-union membership in 1983. In 1985 it was insignificant for both.

Our conclusion is that whether or not endogenous membership selectivity is important for estimating the direct effects of unions on wages, it is relatively unimportant for estimating the compensating differential for risk of fatal accident.

6. EXTENDING THE RISK MEASURE

In this and the next Section we examine the results of two major extensions of the original Marin & Psacharopoulos study. One is to examine the effects of including non-fatal risks, the other is to allow for the decision of how dangerous an occupation to work in to be endogenous.

As stated above, from the GHS we constructed measures for the risk of non-fatal accidents by the 161 occupation classification, using the same approach as for ACCRISK; i.e. for each occupation group the measure is the observed number in that group minus the number expected given the age structure of workers in the group, divided by the number of employees in the group. The GHS asked questions on where the accident occurred and when. We constructed the risk measure both for accidents occurring at the place of work and those during hours of work. The results for the two were extremely close, and we only report the latter; since in principle job related risk should include accidents occurring during work but not in the place of work e.g. lorry drivers' accidents should be included. The variable is called RISKHOUR.

The results are somewhat mixed. Unlike the ACCRISK variable on its own, the sign, size and significance of the RISKHOUR variable is somewhat sensitive to the other included variables - e.g. which union measure or other dummies.

If the RISKHOUR variable is used, but not ACCRISK, then it is insignificantly different from zero for the whole sample; e.g. the coefficient has a "t" of +0.119 using the union membership dummy and -0.366 using the SIC/KOS union variable. For the sub-groups it is negative for manual and managerial/professional workers and (significantly) positive for non-manual.

However, what one is mainly interested in is (i) whether the omission of non-fatal accidents noticeably affects the estimated compensating differential for risk of fatal accidents and thus its use in cost-benefit analysis (ii) whether there is a significant extra compensation for risk of non-fatal accidents.

The answer to Question (i) is negative. Despite doubts that have been expressed over the possible sensitivity of the widely quoted "value of statistical life" to omitting non-fatal effects, the coefficient of the ACCRISK variable (still treated as exogenous) does not seem sensitive to the inclusion or exclusion of RISKHOUR. E.g. see Table 5. Similar results for ACCRISK apply to sub-samples.

The answer to Question (ii) is less robust. When both ACCRISK and RISKHOUR are included the coefficient on the RISKHOUR variable alters according to sample and other variables. It is always negative but often with a low "t". In some combinations, however, it is negative and significantly different from zero at conventional levels, e.g. as in Table 5, whereas using the same union variable but a wider set of dummies (e.g. tenure, region, children etc.) give t = -1.43.

Thus, using ordinary least squares estimates, there does not seem to be a consistent positive compensating differential for the risk of non-fatal accidents, especially if one simultaneously allows for fatal accident risk.[16] This is puzzling, and remains to be investigated with data which distinguishes non-fatal accidents by their degree of severity. However, the estimation of the compensating differential for fatal risks is unaffected by the inclusion or exclusion of non-fatal risks.

7. ENDOGENISING RISK CHOICE[17]

The Heckman procedure, for allowing for the possible selectivity bias arising from the endogenous selection of some variable affecting earnings (e.g. as discussed above in the context of union membership), cannot be used for a continuous variable, such as riskiness of occupation. However, as explained in Section 3, a two-stage approach due to Garen (1984) can be used.

In our case, the first stage uses a least squares regression explaining the level of risk chosen, Equation (2) above. Here, the choice of risk is equivalent to the choice of an occupation with that degree of risk. When non-fatal accidents are included, the first-stage is a pair of choice of risk equations; one for risk of fatal and one for risk of non-fatal accidents, Equations (2) and (3). The second stage includes among the right hand variables the errors from the first stage equation(s) and the interaction of the risk variable(s) with the error(s). (That is, Equation (5) above, with the omission of the terms in q_i and $\underline{\triangle}_{3i}$ when only fatal risks are included.)

It is hypothesised that the level of risk chosen will depend on various personal characteristics and the <u>potential</u> level of full income irrespective of risk, i.e. job related and other forms of income. In general the theory would predict that the higher the level of non-job income and human capital, the lower the level of danger chosen (for any given compensation for extra danger). The variables entering into the earnings equation also have to be included in the first stage.

Various combinations of available variables were tried - some such as tenure were of dubious relevance theoretically and never important, and were then dropped. Of the variables used eventually, the degree of importance (size and significance) varied somewhat according to the precise mix but all the first-stage OLS equations led to very similar errors and interactions for use in the second stage - the simple correlations between errors from the different first-stage equations were typically over +0.9, often reaching +0.99. Typical stage-one equations for accident risks are shown in Table 6, cols 1&2 .

The effects of fatal accident risk on earnings, once that risk choice is endogenised, is shown in Table 7, col. 1. As the Garen procedure should lead to heteroscedasticity, the results are shown for estimation including the White correction (without it the "t" on ACCRISK would have been 1.79, other coefficients show even smaller proportional changes in their standard errors). At first sight, the ratio of 2-3 times of the size of the coefficient as compared to that in Table 2 seems reasonable - though see footnote 10. Furthermore, existing UK studies on more restricted data sets by Siebert & Wei and Sandy & Elliot, as well as the Garen US study (1988) showed similar ratios of increase in the compensation for risk of fatal accidents once the choice of risk was endogenised.

However, the estimation techniques and the interpretation of the results are subject to difficulties that do not seem to have been previously discussed. It can be seen even in the result mentioned that the sum of the coefficients on ACCRISK and on the error term from the first stage is very close to the size of the coefficient on ACCRISK when it was treated as fully exogenous. Furthermore, the standard errors of these two terms in the new equation are almost identical (1.687 v. 1.718 here, or 1.441 v. 1.447 without the heteroscedasticity correction).

The coefficients imply that if somebody's choice of a riskier job is predictable on the basis of personal and wealth characteristics they would get about 2½ times as much as a person whose choice was unpredictable on this basis. Since we cannot observe people's idiosyncratic utility functions, it is not obvious why somebody who is more averse to danger because of externally observable personal characteristics and income effects should get higher compensation for risk than somebody who is simply timid by nature.

As noted above, the sum of the coefficients on the now endogenous ACCRISK and the first-stage error is very close to the coefficient on the original exogenous ACCRISK. In itself this is not necessarily implausible and could be coincidental . However this "coincidence" is suggestive of a deeper problem when combined with some other of our results.

In Garen's study, a footnote mentions that in the first-stage regressions he used as dependent variable the inverse normal cumulative density function of his US risk data. In our case, because our measure of risk is actual minus expected and we find it more convenient to use accidents per thousand, our risk measure is not constrained in principle to be in the zero to one range. However, because ACCRISK does have a very limited range in practice, -0.0521 to +0.3869, we tried using an inverse normal CDF function on 2x(ACCRISK +0.06). In this case, as shown in Table 7 Column 3, the coefficient on ACCRISK fell to 0.5592 and was insignificant. [In this equation, the units of measurement of ACCRISK and the first-stage residual are completely different to each other, so there is no point in the sort of analysis on the coefficients that was made above].

Other results indicating the problem can be seen when combining both fatal and non-fatal risks as endogenous in a wage equation, as in Table 7 Column 2. In this case the coefficient on ACCRISK increases by a further factor of 5, i.e. to ten times that obtained when both ACCRISK and RISKHOUR were treated as exogenous. Here again, it is worth noting that the sum of the coefficients on ACCRISK and its first-stage error term remains very close to that when ACCRISK was treated as exogenous, while that on RISKHOUR and its first-stage error term are close to zero (while with each pair the standard error on the risk and the first-stage residuals are extremely close).

For a further comparison, Table 7 Column 4, shows the results for first-stages with inverse normal CDF transformations for both risk measures.

Our explanation for these results is that with the sort of "success" that is likely to be typical in explaining the choice of risk in the first stage regression(s), there is a very high collinearity between the risk measure and its first-stage regression residual. For example, for the regressions shown in Table 6 Columns 1&2, the simple correlation between ACCRISK and its residual is +0.87, while for RISKHOUR and its residual it is +0.89. (With the CDF transformations they are slightly less, +0.82 and +0.86 respectively; even though the residuals for the two forms of ACCRISK first-stage regressions have a +0.95 correlation and RISKHOUR forms have +0.97).

This collinearity of the regressors in the second-stage equations could explain both their instability, especially on the risk variables, as shown both in the results for the whole sample already mentioned, and in results for the sub-samples, not analysed here to save space.

8. CONCLUSIONS

The main conclusion of this study is that the sign and statistical significance of the compensation for accepting the risk of fatal injury on the job are robust both to the time period studied and to the explanatory variables included. However while the robustness of the *size* of the compensation does hold with respect to the explanatory variables included, it does not hold for varying the time period, even for compensation as a proportion of income. Using data a decade later, the proportionate compensation has approximately tripled.

The implications for the use in cost-benefit analysis are mixed. Our results show that there needs to be *some* positive allowance for reductions in risk. The robustness with respect to the other explanatory variables gives strong support for the use of such valuations. Conversely, the variability over time casts some doubt on the desirability of claiming strong support for any particular precise valuation. To modify the latter conclusion would require further research either to produce an explanation for the changes over time or to show by studies similar to this one that temporal instability does not occur in other countries or between other periods in the UK.

References:

Arabsheibani,G.R., J.M.Emami and A. Marin (1996) The Impact of computer Use on Earnings in the UK (unpublished)

Arabsheibani,G.R. and A. Marin (1997) Self Selectivity Bias with a Continuous Variable: Potential Pitfall in a Common Procedure (unpublished)

Biddle,J.E. and G.A.Zarkin (1988) "Worker Preferences and Market compensation for Job Risk" Review of Economics and Statistics 70.

Blanchflower, D.G. and A.J.Oswald (1994) The Wage Curve (MIT Press)

Booth, A.L. (1995) The Economics of the Trade Union (Cambridge U.P.)

Broome, J. (1978) "Trying to Value a Human Life" Journal of Public Economics 9(1) and discussion in 12(2), 1979, same journal.

Broome, J. (1985) "The Economic Value of Life" Economica 52(207)

Fisher, A., L.G.Chestnut and D.M.Violette (1989) "The Value of Reducing Risks of Death: A Note of New Evidence" Journal of Policy Analysis and Management 8(1).

Garen, J. (1984) "The Returns to Schooling: A Selectivity Bias Approach With a Continuous Choice Variable" Econometrica 52(5)

Garen, J. (1988) "Compensating Wage Differentials and the Endogeneity of Job Riskiness" Review of Economics and Statistics 70.

Jones-Lee, M.W. (ed.) (1982) The Value of Life and Safety (North-Holland)

Jones-Lee, M.W. (1989) The Economics of Safety and Physical Risk (Basil Blackwell)

Kahn, S. and K.Lang (1988) "Efficient Estimation of Structural Hedonic Systems" International Economic Review 29(1)

Kniesner, T.J. and J.D. Leeth (1991) "Compensating Wage Differentials for Fatal Injury Risk in Australia, Japan and the United States" Journal of Risk and Uncertainty 4(1)

Maddala, G.S.(1983) Limited Dependent and Qualitative Variables in Econometrics (Cambridge U.P.)

Marin, A. and G.Psacharopoulos (1982) "The Reward for Risk in the Labour Market: Evidence from the United Kingdom and a Reconciliation with Other Studies" Journal of Political Economy 90(4).

Moulton, B.R. (1986) “Random Group Effects and the Precision of Regression Estimates” Journal of Econometrics 32

Rosen, S. (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" Journal of Political Economy 82

Rosen, S. (1986) "The Theory of Equalising Differences" in O. Ashenfelter and R. Layard (eds.) Handbook of Labor Economics (Elsevier)

Sandy, R. and R.F.Elliott (1996) " Unions and Risk: Their Impact on the Compensation for Fatal Risk " Economica 63

Siebert, W.S. and X.Wei (1994) "Compensating Wage Differentials for Workplace Accidents: Evidence for Union and Nonunion Workers in the UK" Journal of Risk and Uncertainty 9

Smith, R.S. (1973) "Compensating Wage Differentials and Hazardous Work" Technical Analysis Paper No. 5, Office of Evaluation, U.S. Dept. of Labor.

Smith, R.S. (1976) The Occupational Safety and Health Act, Its Goals and Its Achievements (American Enterprise Institute)

Thaler, R. and S.Rosen (1975) "The Value of Saving a Life: Evidence from the Labor Market" in N. Terleckyj (ed.) Household Production and Consumption (Columbia U.P.)

Viscusi, W.K. (1993) "The Value of Risks to Life and Health" Journal of Economic Literature XXXI (4)

Viscusi, W.K. and M.J.Moore (1989) "Rates of Time Preference and Valuations of the Duration of Life" Journal of Public Economics 38.

Variable Name Abbreviations in Tables

ACCRISK	Excess rate (per thousand per year) of fatal accidents
ERR ACC	Errors from ACCRISK choice equation
ERR HR	Errors from RISKHOUR choice equation
EXP	Experience, measured by age minus schooling minus 5 (UK) or minus 6 (overseas)
EXP SQ	Experience squared
GH	Goldthorpe - Hope occupational rating
LOGREAL	Log of weekly earnings divided by earnings index for month in which individual surveyed
Ln Y	Log of 1975 annual earnings
RISKHOUR	Excess rate (per thousand per year) of non-fatal injuries during hours of work.
S	Years of schooling
UNIMAN	Estimated union membership dummy
UNION	Percentage union collective bargaining coverage by industry in 1975
USK	Percentage union membership by KOS\industry group in 1983

TABLE 1 **EARNINGS FUNCTIONS - EARLIER STUDY FOR 1975**

Dependent Variable - ln Y

Independent Variable	Whole		Managers and Professionals		Non-Manual Workers		Manual Workers	
Constant	1.9448 (25.40)	1.9477 (25.44)	2.5173 (4.72)	2.3902 (4.56)	1.4484 (7.39)	1.4222 (7.45)	2.4534 (24.93)	2.4573 (24.97)
S	.0584 (23.94)	.0581 (23.42)	.0649 (10.52)	.0635 10.37)	.0634 (16.94)	.0632 (16.98)	.0324 (5.52)	.0324 (5.51)
EX	.0461 (34.19)	.0462 (34.22)	.0459 (8.98)	.0461 (9.03)	.0528 (18.71)	.0528 (18.72)	.0395 (25.34)	.0396 (25.43)
EX SQ	-.0008 (30.34)	-.0008 (30.39)	-.0007 (7.15)	-.0007 (7.23)	-.0009 (16.19)	-.0009 (16.22)	-.0007 (24.30)	-.0007 (24.39)
ln WEEKS	1.1302 (61.55	1.1306 (61.56)	.8366 (6.47)	.8377 (6.47)	1.2889 (27.11	1.2895 (27.13)	1.1123 (59.69)	1.1122 (59.67)
ACCRISK	.3573 (3.84)	.2235 (4.39)	2.9393 (1.85)	.2813 (.57)	1.3000 (1.30)	.7710 (1.86)	.3844 (3.99)	.2626 (5.23)
UNION	.0021 (8.30)	.0020 (8.23)	-.0015 (0.51)	.0033 (3.64)	.0013 (.70)	.0023 (3.88)	.0018 (6.13)	.0016 (5.96)
GH	.0081 (21.95)	.0081 (21.93)	.0180 (8.82)	.0181 (8.86)	.0034 (4.88)	.0034 (4.87)	.0059 (9.40)	.0059 (9.52)
UNIONxRISK	-.0045 (1.72)	-	-.0966 (1.75)	-	-.0194 (.58)	-	-.0038 (1.49)	-
R^2	.58	.58	.38	.38	.54	.54	.59	.59
N	5464	5464	688	688	1335	1335	3378	3378

Figures in brackets are absolute values of t-ratios

TABLE 2 **BASIC EQUATIONS, ACCRISK**

Dependent Variable - Logreal

Independent Variable	Whole Sample		Manual		Non-Manual		Managerial/Professional	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	3.6027 (74.461)	3.6040 (74.652)	3.6926 (42.005)	3.6727 (41.526)	3.966 (41.623	4.0026 (40.112)	3.1752 (20.374)	3.0493 (18.541)
S	0.0436 (12.621)	0.0437 (12.659)	0.0381 (5.922)	0.0383 (5.968)	0.0330 (5.099)	0.0334 (5.159)	0.0626 (10.646)	0.0620 (10.570)
EXP	0.0367 (18.439)	0.0366 (18.415)	0.0247 (9.229)	0.0247 (9.249)	0.0406 (10.182)	0.0405 (10.149)	0.0509 (11.565)	0.0511 (11.646)
EXP SQ	-0.0006 (15.259)	-0.0006 (15.241)	-0.0004 (8.120)	-0.0004 (8.134)	-0.0007 (8.428)	-0.0007 (8.352)	-0.0008 (8.620)	-0.0008 (8.652)
ACCRISK	0.9464 (6.257)	-0.637 (1.545)	0.7472 (4.505)	1.8334 (3.129)	3.1036 (6.701)	4.9688 (3.129)	1.7928 (2.969)	-3.2016 (1.450)
USK	0.1106 (4.180)	0.1281 (4.788)	0.3233 (8.272)	0.3500 (8.451)	0.132 (0.235)	-0.068 (0.788)	-0.1310 (2.251)	0.0869 (0.795)
GH	0.0119 (21.559)	0.0115 (20.786)	0.0111 (11.073)	0.0110 (11.018)	0.0084 (7.185)	0.0086 (7.287)	0.0140 (6.455)	0.0133 (6.137)
USK x RISK	-	2.3803 (4.130)	-	-1.5690 (1.932)	-	-2.8038 (1.228)	-	6.5074 (2.351)
R^2	.320	.322	.168	.170	.278	.280	.247	.251
N	3608	3608	1802	1802	734	734	974	974

Figures in brackets are absolute values of t-ratios

TABLE 3 **OMITTING GOLDTHORPE INDICES**

Dependent Variable - Logreal

Independent Variable	Whole Sample	Manual	Non-Manual	Managerial/ Professional	Whole Sample	Manual	Non-Manual	Managerial/ Professional
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CONSTANT	3.7091 (72.910)	3.9247 (44.588)	4.1369 (43.418)	3.938 (40.168)	3.5690 (35.858)	3.8682 (35.858)	4.0472 (38.032)	3.7934 (32.506)
S	0.0808 (25.457	0.0523 (8.140)	0.0494 (7.894)	0.0729 (12.662)	0.0829 (22.574)	0.0649 (8.389)	0.0490 (7.086)	0.0741 (11.345)
EXP	0.0437 (21.011)	0.0273 (9.950)	0.0436 (10.637)	0.0489 (11.049)	0.0556 (15.068)	0.0382 (7.284)	0.0572 (8.493)	0.0609 (8.065)
EXP SQ	-0.0007 (16.816)	-0.0004 (8.673)	-0.0008 (8.840)	-0.0007 (8.109)	-0.0011 (10.271)	-0.0007 (5.234)	-0.0012 (6.228)	-0.0011 (8.065)
ACCRISK	-0.2177 (1.451)	0.2205 (1.346)	3.8719 (7.107)	1.2278 (2.017)	-0.0144 (0.085)	0.3877 (1.971)	3.3001 (6.611)	0.6937 (0.944)
USK	-0.0263 (0.972)	0.3207 (7.952)	0.0709 (1.242)	-0.0759 (1.299)	-	-	-	-
UNIMAN	-	-	-	-	0.0622 (3.941)	0.0990 (4.517)	0.1111 (3.635)	0.0229 (0.731)
R^2	.235	.110	.227	.220	.244	.109	.281	.245
N	3646	1812	735	991	2688	1289	594	724

Figures in brackets are absolute values of t-ratios

TABLE 4 **USING ESTIMATED UNION DUMMY**

Dependent Variable - Logreal

Independent Variable	Whole	Manual	Non-Manual	Managerial/ Professional
	(1)	(2)	(3)	(4)
CONSTANT	3.5701 (63.036)	3.6843 (34.614)	3.8840 (35.829)	2.9825 (16.109)
S	0.0472 (11.613)	0.0457 (5.874)	0.0350 (4.833)	0.0641 (9.635)
EXP	0.0487 (13.755)	0.0336 (6.576)	0.0566 (8.587)	0.0653 (8.691)
EXP SQ	-0.0010 (9.727)	-0.0007 (4.824)	-0.0013 (6.333)	-0.0012 (5.569)
ACCRISK	1.0753 (6.210)	0.9412 (4.699)	3.1152 (6.369)	1.3360 (1.833)
UNIMAN	0.0642 (4.276)	0.0174 (5.042)	0.0628 (2.013)	-0.0308 (0.954)
GH	0.0107 (17.099)	0.0117 (9.347)	0.0072 (5.340)	0.0143 (5.633)
R^2	.317	.169	.314	.276
N	2669	1284	594	712

Figures in brackets are absolute values of t - ratios

TABLE 5 (EXOGENOUS) INCLUSION OF RISKHOUR, WHOLE SAMPLE

Dependent Variable - Logreal

Independent Variable	Whole	
	(1)	(2)
CONSTANT	3.6748 (75.014)	3.6246 (73.563)
S	0.0415 (11.823)	0.0424 (12.141)
EXP	0.0372 (18.643)	0.0365 (18.348)
EXP SQ	-0.0006 (15.516)	-0.0006 (15.197)
ACCRISK	-	1.0542 (6.665)
USK	0.1159 (4.325)	0.1177 (4.420)
GH	0.0106 (19.906)	0.0117 (21.100)
RISKHOUR	-0.7E-06 (0.366)	-0.0005 (2.314)
R^2	.313	.322
N	3608	3608

Figures in brackets are absolute values of t - ratios

TABLE 6 **CHOICE OF RISK**

Dependent Variables as shown

Independent Variable	ACCRISK	RISKHOUR	ACCRISK TRANSFORMATION	RISKHOUR TRANSFORMATION
	(1)	(2)	(3)	(4)
CONSTANT	0.0614 (8.204)	65.330 (11.542)	-0.4277 (6.045)	-0.5684 (10.000)
SECOND WAGE INCOME	-0.6E-06 (0.375)	-0.0008 (0.754)	0.6E-05 (0.431)	-0.9E-05 (0.818)
WIFE'S SCHOOLING	-0.5E-05 (0.016)	-0.4489 (1.975)	-0.0020 (0.713)	-0.0042 (1.835)
SINGLE	-0.0096 (2.515)	-8.7411 (3.012)	-0.1293 (3.569)	-0.0963 (3.309)
PENSION RIGHTS CURRENT OR PREVIOUS	-0.0064 (2.956)	-4.9290 (3.024)	-0.0332 (1.608)	-0.0404 (2.433)
RENTING	0.0040 (1.975)	3.8704 (2.498)	0.0243 (1.261)	0.0447 (2.885)
OUTRIGHT HOME OWNER	-0.0051 (1.711)	1.0072 (0.446)	-0.0505 (1.764)	0.0211 (0.917)
INDEX OF FATHER'S WAGE	0.9E-06 (2.369)	-0.0003 (1.063)	0.8E-05 (2.271)	-0.3E-05 (1.134)
HOUSEHOLD NON-LABOUR INCOME	0.5E-07 (0.080)	0.0001 (0.249)	0.2E-06 (0.044)	0.2E-05 (0.475)
RACE DUMMY	-0.0093 (2.193)	0.6065 (0.179)	-0.0831 (1.963)	0.0019 (0.056)
WIFE'S LABOUR INCOME	-0.3E-06 (1.040)	-0.0001 (0.395)	-0.5E-05 (1.542)	-0.2E-05 (0.906)
DEPENDENT CHILD DUMMY	0.0006 (0.306)	1.0366 (0.687)	0.0121 (0.633)	0.0161 (1.048)
S	-0.0012 (2.756)	-2.5476 (7.475)	-0.0228 (5.296)	-0.0270 (7.806)
EXP	0.0008 (1.785)	-0.1561 (0.481)	0.0025 (0.628)	-0.0024 (0.751)
EXP SQ	-0.2E-04 (1.547)	-0.1E-04 (0.002)	-0.8E-04 (0.700)	0.3E-04 (0.319)
GH	-0.0013 (19.638)	-0.6112 (11.903)	-0.0142 (22.080)	-0.0061 (11.809)
USK	0.0075 (2.231)	17.627 (6.929)	0.0082 (0.257)	0.1603 (11.809)
R^2	.240	.216	.315	.235
N	2930	2930	2676	2676

TABLE 7 EARNINGS WITH ENDOGENOUS RISK WHOLE SAMPLE

Dependent Variable - Logreal

Independent Variable	(1)	(2)	(3)	(4)
CONSTANT	3.4347 (28.345)	4.4903 (27.021)	3.5621 (48.034)	3.8917 (34.929)
S	0.0486 (9.571)	-0.0146 (1.602)	0.0457 (10.233)	0.0280 (4.334)
EXP	0.0480 (11.479)	0.0344 (7.848)	0.0506 (13.945)	0.0492 (13.518)
EXP SQ	-0.0009 (8.500)	-0.0007 (6.380)	-0.0010 (10.228)	-0.0010 (10.090)
ACCRISK	2.5812 (1.530)	12.377 (6.074)	0.5592 (0.647)	1.7145 (1.725)
USK	0.0183 (2.389)	0.4648 (8.635)	0.0900 (2.793)	0.1927 (4.925)
GH	0.0132 (5.695)	0.0077 (3.389)	0.0105 (8.702)	0.0072 (4.840)
ERR ACC	-1.6045 (0.934)	-10.931 (5.451)	0.0441 (.527)	-0.0388 (0.422)
ACCRISK x ERR ACC	0.3302 (0.229)	-0.4282 (0.262)	0.1473 (0.539)	-0.1242 (0.375)
ACCRISK x ERR HR		0.0310 (2.575)		1.4838 (1.778)
RISKHOUR		-0.0270 (8.880)		-0.0074 (3.982)
ERR HR		0.0266 (8.728)		0.6636 (3.758)
RISKHOUR x ERR ACC		-0.0309 (2.354)		-0.0012 (1.251
RISKHOUR x ERR HR		-0.1E-05 (0.425)		0.0027 (2.519)
R^2	.316	.342	.316	.321
BREUSCH-PAGAN CHI-SQ TEST	60.85 (8 D of F)	175.97 (13 D of F)	59.53 (8 D of F)	125.78 (13 D of F)
N	2676	2676	2676	2676

1. Figures in Brackets are absolute values of t - ratios using White heteroscedasticity correction
2. Errors in columns (1) and (2) from equations (1) and (2) in Table 6
3. Errors in columns (3) and (4) from equations (3) and (4) in Table 6

ENDNOTES

1. In this paper, we follow the usage in this literature, whereby “risk” refers to job related danger of death or non-fatal illness/injury, rather than the usage elsewhere in economics.

2. An early example is Thaler & Rosen (1975).

3. A thorough account is in Jones-Lee (1989).

4 Part of the disagreement stems from the fact that the verbal discussion above is equivalent to, and based on, some diagrammatic treatments where the individual has indifference <u>curves</u> between income and probability. See, for example, Thaler & Rosen (1975) or Rosen(1986). In many formal treatments using a von Neumann-Morgenstern framework the utilities are assigned to life and death, and the probabilities then used to calculate expected utility. The standard assumptions rule out any preference for probability *per se,* only for the outcomes to which the probability is attached. This approach would imply indifference <u>lines</u> in the space of income and probability. Such straight lines would have to intersect a vertical line at probability equals one, i.e. a finite differential for certain death. However, indifference <u>curves</u> can be asymptotic to such a vertical line and thus need not imply a value of any particular person's life. Broome, in Jones-Lee (1982) and elsewhere, discusses the alternative problems that can arise from the framework used here.

5. Most studies use $\log Y$ as the dependent variable. In this case the coefficient of $\partial \log Y / \partial p$ does not even depend on the units in which Y is measured.

6. This follows from the previous footnote plus that: Install Equation Editor and double-click here to view equation.

7. Biddle & Zarkin attempt to justify this by distinguishing *individual* workers’ choices from the frontier which is a given constraint to the individual worker.

8. Biddle & Zarkin use a translog approximation instead. The Viscusi & Moore approach allows them to estimate the compensating variation for risk in a life-cycle context, including discounting. The more standard approach implicitly assumes that workers can completely change their job each period.

9. For reasons to be discussed in the next Section, with our data we could not use the simultaneous Maximum Likelihood estimates to jointly estimate the union membership choice probit and union/non-union wage equations including risk.

10 This particular aspect of the possibility of self-selection bias may be less relevant to our study than to some others. This is because we consider all occupations, rather than just the most dangerous ones, across the whole male labour force. Compare Rosen (1986) p.661.

11. The explanation here approximately follows that in Garen (1988), which is based on the fuller theoretical development in Garen (1984).

12 The problem of the appropriate level of aggregation in this study is not the same as that investigated for the US by Kniesner and Leeth (1991). Their study was concerned with aggregating over <u>all</u> variables,

including the dependent one (and therefore using average earnings and other characteristics for a group rather than individual earnings functions), not with aggregating the risk level to which individuals are assigned.

The analysis by Moulton (1986) suggests that using the least aggregated occupational grouping would give more reliable estimates of the ‘t’ statistics. See the statement in Section 5, and the next note, on the unimportance of the Moulton correction using our ACCRISK grouping.

13. The only few minor exceptions did not concern cases which matter for our conclusions. These cases were USK in col.7 (t=-1.522), and in the unstable interaction regressions USKxRISK in col.4 (t=-1.632), ACCRISK in col.6 (t=1.839) and USKxRISK in col.8 (t=1.823).

14. A few subsequent US studies found opposite results. The various findings are summarised in Sandy and Elliot (1996).

15 We noticed one interesting result when checking whether it made a noticeable difference to the ACCRISK coefficient if the same wider set of dummies were included in the income equation as had been used in the probit equation to explain union membership. Although the coefficients on ACCRISK were very similar to those in the basic regressions, the coefficients on the union variables (both the membership dummy and the KOS/SIC coverage) became negative even for the whole sample (t = -2.63 for the SIC/KOS group coverage variable, t = -1.81 for the membership dummy), as well as in the sub-groups. This switch to a negative coefficient was caused by the inclusion of the SIC (industry) and workplace size dummies. If both of these sets were omitted, but the rest of the dummies from the probit (e.g. race, region etc.) included, the coefficients on the union variables remained positive (and significant). Going back to 1983 data, inclusion of the SIC and size dummies in a wage equation drastically reduced the size and significance of the effect of unionisation on wages, without actually making it negative.

Among possible explanations is that the size of the workplace, and the industry, are genuinely independent determinants of earnings, as well as being linked to the likelihood of union membership. For 1985 it may be that, given the ongoing changes in industrial relations, labour market conditions and industrial structure, being a member of a trade union or having wages determined in a heavily unionised group was actually a disadvantage. However, regressions which do not allow separately for SIC or for the premium for working in a larger plant/office, may mask the disadvantage of unionisation in 1985. Further detailed investigation with other data sets than the GHS will be necessary to see how far the pattern extends to later years, and if it does, to pin-down an explanation. (We have found similar results for 1990 as well as 1985 when working on Arabsheibani, Emami and Marin [1996]).

16. Similar results have been found in some US studies. See Viscusi (1993) for a summary of US studies and a brief discussion of some of the problems in obtaining reliable estimates of non-fatal injury compensation. Siebert & Wei (1994) did find positive results for the UK, but their sample was for manual workers only and used the broader occupational grouping.

17. The results in this section are discussed in more detail in Arabsheibani and Marin (1997).