

IZA DP No. 4439

A Risk Augmented Mincer Earnings Equation? Taking Stock

Joop Hartog

September 2009

A Risk Augmented Mincer Earnings Equation? Taking Stock

Joop Hartog

*University of Amsterdam
and IZA*

in cooperation with:

Simona Bajdechi Raita, Peter Berkhout, Jose Cabral Vieira, Luis Diaz Serrano, Bas Jacobs, Helena Skyt Nielsen, Hans van Ophem, Erik Plug, Juerg Schweri, Wim Vijverberg, Dinand Webbink, Stefan Wolter

Discussion Paper No. 4439
September 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

A Risk Augmented Mincer Earnings Equation? Taking Stock*

We survey the literature on the Risk Augmented Mincer equation that seeks to estimate the compensation for uncertainty in the future wage to be earned after completing an education. There is wide empirical support for the predicted positive effect of wage variance and the negative effect of wage skew. We discuss robustness of the findings across specifications, potential bias from unobserved heterogeneity and selectivity and consider the core issue of students' information on benefits from education.

JEL Classification: J31, D8

Keywords: human capital, earnings function, risk

Corresponding author:

Joop Hartog
Amsterdam School of Economics
University of Amsterdam
Plantage Muidergracht 12
1018 TV Amsterdam
The Netherlands
E-mail: j.hartog@uva.nl

* I have benefited from discussions in the joint UvA-VU Labour seminar, at CREB (University of Barcelona), the SOLE/IZA Transatlantic Meeting in Buch am See 2007, University of Maastricht, SUNY Binghamton, SUNY Albany and Queen's University Kingston (Ont., Canada). Comments by Jim Rebitzer, Michael Sattinger and Pieter Gautier are also gratefully acknowledged.

1. Basic hypothesis

Suppose, individuals can either go to work or go to school for five years. The post-school log wage is 0.325 higher than the wage from going straight to work. Then, in the basic Mincer framework¹ the rate of return to schooling would be 6.5%. Now suppose, the direct, unschooled wage is fixed, but the post-school wage is uncertain, at a realistically modest standard deviation of 0.25. Then, if an individual would make it to one standard deviation above the mean in the post school distribution, the rate of return would be 11.3%, but if he would not reach above one standard deviation below the mean, his return would not surpass 1.3%. Instead of a single return of 6.5% for everyone, two thirds of all individuals would have a return somewhere in the wide interval between 1.3 and 11.3 %.

Of course the essential question would be which part of this variation is risk and which part is known to the student but hidden as unobserved heterogeneity to the outsider. Yet, there can be no doubt that the decision to engage in an education is a decision under uncertainty. There are at least three dimensions in which information is incomplete. The potential student will generally not fully grasp the requirements of the school curriculum and of the occupations available after graduation, she will usually not even be fully sure of her own abilities and preferences and she will not know what the exact returns to the investment will be. With uncertainties so prominent, one is inevitably led to expect that the returns to education will be shaped not only by compensation for postponing earnings, but also by compensation for risk. In fact however, the basic Mincer earnings equation is

¹ The basic Mincer framework leads to earnings equation $\ln Y_s - \ln Y_0 = rS$. The standard deviation is taken from Hartog, van Ophem and Bajdechi (2007). That paper presents a simulation of the distribution of rates of return as a function of stochastic properties of alternative life cycle wage profiles.

derived under conditions of certainty on future earnings in the alternatives and the estimations similarly ignore the impact of uncertainty. Research on returns to education has focussed for decades on getting an unbiased estimate of the causal effect of schooling on earnings. Main emphasis is on the proper econometric modelling to deal with omitted ability bias, measurement errors in education and the endogeneity of the schooling decision. Admittedly, these are serious and stubborn problems, as illustrated by the fact that the dust has not settled down and that there is as yet no consensus on true, causal, rates of return to education and their variation in response to individual and institutional conditions. Perhaps all we can be confident about is an interval of the return for an additional year of education between a few percent to an upper limit of some 20 percent.

With uncertainty so prominent, one might have thought that there is a large literature to deal with it. In fact, analyses investigating the consequences of uncertain completion of an education, incomplete information on abilities and even preferences and on the poor predictability of the career development that comes with an education are very scarce². This paper can only make a small contribution, by treating just one aspect of the pervasively present uncertainty. We focus on the notion that an individual contemplating an education does not anticipate a given post-school wage rate but rather an entire wage *distribution*, without knowing exactly where in that distribution she will end up. As the individual perceives that the returns to his investment are not certain, he will demand compensation. This paper is about the existence and magnitude of a risk premium in the returns to education. It turns out that a basic idea on an econometric approach to this problem has been

² The classic reference for human capital investment under uncertainty is Levhari and Weiss (1974). Interest in the role of risk is growing; see the *Labour Economics Special Issue on Education and Risk*, 2007

around for a long time³. The literature has started with King (1974), who used aggregate data by occupation to estimate the effect of variance and skew of wages within an occupation on the mean occupational wage. The variance of earnings is taken as a measure of risk that individuals want compensation for. Positive skewness of the earnings distribution points to the opportunity of attaining real high earnings and this is something people want to pay for with reduced expected earnings. With micro data one may proceed in two steps. First, estimate a standard Mincer earnings equation, group the residuals by education-occupation classification and take the within-group distribution of residuals as indicating the uncertainty associated with choosing the particular education-occupation combination. In the second stage, add the variance of these education-occupation residuals to the earnings function and estimate risk compensation as the regression coefficient of earnings on the variances. The two-stage approach was first applied by McGoldrick (1995). Following King (1974), she also included the skew of residuals by occupation in the regression equation.

Compensation for earnings risk has now been estimated in close to twenty studies. Generally, the studies report an earnings premium for risk and a rebate for skew, as theory predicts. In this paper we review these studies and assess the evidence so far. In section 2 we present a formal model underlying the estimations, in section 3 the empirical specifications. In Section 4 we present estimation results. Section 5 addresses the problems of bias from heterogeneity and presents estimates that control for these problems. Section 6 considers self-selection and the information set that potential students have. Section 7 presents

³ After I suggested this approach to Luis Diaz Serrano for his dissertation, he dug up the early references. A referee on one of the papers discussed here noted that the same idea had been repeatedly formulated by students.

supporting evidence and Section 8 discusses alternative explanations for our results. Section 9 wraps up, with conclusions and a research agenda.

2. A simple formal model.

To derive the compensation for earnings postponement and for accepting earnings uncertainty, we will first consider schooling choice under utility maximization rather than earnings maximisation (which is unavoidable if we analyse uncertainty in a utility framework). In the next stage we will examine earnings uncertainty proper.

2.1 *Utility maximization in the absence of uncertainty*

Individuals face two alternatives: go straight to work, and earn an annual non-stochastic income Y_0 for the rest of their working life, or go to school for s years, and then after school earn a non-stochastic income Y_s for the rest of their working life. We assume individuals have an uninhibited choice between alternatives. In equilibrium, lifetime utility should be equal. We seek to derive the premium M_s that accomplishes equilibrium, hence

$$Y_0 = (1 - M_s)Y_s \tag{1}$$

i.e., M_s is the mark-off on Y_s to equate lifetime utility. Discounting at a rate δ and setting utility of zero income at zero, we get

$$\int_0^{\infty} U(Y_0) e^{-\delta t} dt = \int_s^{\infty} U(Y_s) e^{-\delta t} dt \quad (2)$$

where $U(Y)$ is a standard utility function. With time-independent income (and utility) this solves into

$$U(Y_0) = e^{-\delta s} U(Y_s) \quad (3)$$

A simple first-order expansion of $U(Y_0)$ around Y_s generates

$$\begin{aligned} U(Y_0) &= U(Y_s) + (Y_0 - Y_s) U'(Y_s) \\ &= U(Y_s) - M_s Y_s U'(Y_s) \end{aligned} \quad (4)$$

Combining (3) and (4) yields the solution for M_s :

$$M_s = (1 - e^{-\delta s}) \frac{U(Y_s)}{U'(Y_s)} \frac{1}{Y_s} \quad (5)$$

Clearly, this is a generalisation of Mincer's earnings function. The last two terms jointly are the inverse of the income elasticity of utility. Under earnings maximization, $U(Y)=Y$, and (5) reduces to the term in parentheses, implying $\ln Y_s = \ln Y_0 + \delta s$, the standard Mincer equation.

Assuming a constant income elasticity of utility simplifies equation (5). Specifically, with Constant Relative Risk Aversion (CRRA)

$$U(Y) = \frac{1}{1-\rho} Y^{1-\rho} \quad (6)$$

we get

$$M_s = \frac{1 - e^{-\delta s}}{1 - \rho} \quad (7)$$

In an expected utility framework, monetary returns will be depreciated by the declining marginal utility of income.

2.2 *Uncertainty: Risk Aversion and Skewness Affection*

Consider an individual who must choose between two options, one with a fixed income Y^* , and the other with random income Y at an expected income of $E[Y] = \mu$. Define Θ as the generalized absolute risk premium: $\Theta = E[Y] - Y^* = \mu - Y^*$: an individual in the risky situation receives an expected income that exceeds the income of the risk free activity by an amount of Θ .

We seek to establish the equilibrium risk premium Θ , i.e., the gap between expected income in both positions at which a utility maximizing individual is indifferent between the two positions. Utility is defined again as a continuous differentiable function of income $U(Y)$, with $\partial U / \partial Y > 0$. We also assume risk aversion $\partial^2 U / \partial Y^2 < 0$. Indifference requires

$$U(\mu - \Theta) = E[U(Y)] \quad \text{with} \quad E[Y] = \mu \quad (8)$$

To solve for Θ , we follow Pratt (1964, his equations (4) to (7)). For the left-hand side of (8) we write

$$U(\mu - \Theta) = U(\mu) - \Theta U'(\mu) \quad (9)$$

where $U'(\mu)$ is $\partial U / \partial Y$ evaluated at $Y = \mu$. For the right-hand side we retain one more term than Pratt did in his Taylor series expansion:

$$U(Y) = U(\mu) + (Y - \mu)U'(\mu) + \frac{1}{2}(Y - \mu)^2 U''(\mu) + \frac{1}{6}(Y - \mu)^3 U'''(\mu) \quad (10)$$

where $U''(\mu) = \partial^2 U / \partial Y^2$ evaluated at $Y = \mu$, and $U'''(\mu) = \partial^3 U / \partial Y^3$ at $Y = \mu$. Hence

$$E[U(Y)] = U(\mu) + \frac{1}{2}m_2 U''(\mu) + \frac{1}{6}m_3 U'''(\mu) \quad (11)$$

since $E[Y] = \mu$, and with m_2 and m_3 defining the second and third moments of Y (the variance and the skewness) around μ . Equating (9) and (11), we are now able to solve for Θ :

$$\Theta = \frac{1}{2}V_a m_2 - \frac{1}{6}V_a \frac{U'''(\mu)}{U''(\mu)} m_3 \quad (12)$$

where $V_a = -U''(\mu)/U'(\mu)$ is the degree of absolute risk aversion. If, by analogy to risk aversion, we define absolute skewness affection as $F_a = -U'''(\mu)/U''(\mu)$, we may write

$$\Theta = \frac{1}{2}V_a m_2 - \frac{1}{6}V_a F_a m_3 \quad (13)$$

Equation (13) is the standard equation for a risk premium as derived by Pratt (1964) and Arrow (1965), but now expanded with skewness affection. In this paper, we refer to the second moment as risk.

The sensitivity of absolute risk aversion to income (or wealth) Y can be derived as

$$\frac{d}{dY} \left[-\frac{U''}{U'} \right] = \frac{-U'U''' + (U'')^2}{(U')^2} \quad (14)$$

As Arrow (1965) argues, increasing absolute risk aversion is an absurd assumption, as it would imply investing less in risky alternatives if income (wealth) increases. Assuming the same holds for choice in the labour market, we require *decreasing* absolute risk aversion: the expression in equation (14) ought to be negative. For this to hold, a necessary but not sufficient condition is $U''' > 0$, as was first pointed out by Tsiang (1972, p. 359): “Thus, if we regard the phenomenon of increasing absolute risk aversion as absurd, we must acknowledge that a normal risk-averse individual would have a preference for skewness in addition to an aversion to dispersion (variance) of the probability distribution of returns.” Since, for a risk averter, V_a is positive and since F_a is positive if we assume decreasing absolute risk aversion, we conclude from (13) that the absolute risk premium Θ is positive in risk (variance) and negative in skewness. This motivates our terminology of skewness affection; for evidence on skewness affection from choices made in gambling and betting, see Garrett and Sobel (1999) and Golec and Tamarkin (1998). In the literature on lifetime wealth accumulation, skewness affection is called prudence (Gollier, 2001, p 238).

The empirical model derives more naturally when the compensation for uncertainty is expressed through a relative, rather than absolute, risk premium. It is straightforward to rewrite (13) as an equation for the generalized relative risk premium $\Pi = \Theta / \mu$ (and hence, $Y^* = (1 - \Pi) \mu$). Dividing (13) by μ and slightly rewriting yields

$$\Pi = \frac{\Theta}{\mu} = \frac{1}{2} V_r \frac{m_2}{\mu^2} - \frac{1}{6} V_r F_r \frac{m_3}{\mu^3} \quad (15)$$

where relative risk aversion is given by

$$V_r = V_a \mu = -\frac{U''(\mu)}{U'(\mu)} \mu > 0$$

and relative skewness affection by

$$F_r = F_a \mu = -\frac{U'''(\mu)}{U''(\mu)} \mu > 0$$

Note that without further assumptions, V_r and F_r depend on μ . But under constant relative risk aversion as in equation (6), we would have $V_r = \rho$ and $F_r = \rho + 1$; thus the relative risk premium would be constant.

2.3 Combining earnings postponement and uncertainty

Now, let us apply the risk compensation argument in the context of the Mincer framework. We know that there are two compensation arguments, earnings postponement and uncertainty compensation. In a non-stochastic world, as lifetime post-school income Y_s^f

would be equivalent with a lifetime income of Y_0 . However, after s years of schooling the individual in our case faces a stochastic income Y_s , with expectation $E[Y_s] = \mu_s$ and second and third moments around the mean m_{2s} and m_{3s} . We now seek the risk premium on this fixed post-school income that would compensate for the earnings uncertainty. We express this as a mark-off on the expected income in the risky situation, i.e., we replace Y_s^f by $(1 - \Pi_s)\mu_s$ where, as before, Π_s is the risk premium. Lifetime equal utility requires

$$\int_s^\infty U((1 - \Pi_s)\mu_s) e^{-\delta t} dt = E \left[\int_s^\infty U(Y_s) e^{-\delta t} dt \right] \quad (16)$$

Clearly, the discounting factors drop out as income is as yet time-independent. Expanding the left hand side around μ_s in first-order, and the right hand side around μ_s up to the third order, the earlier result simply re-appears, and we can write

$$\Pi_s = \frac{1}{2} \frac{m_{2s}}{\mu_s^2} V_r - \frac{1}{6} \frac{m_{3s}}{\mu_s^3} V_r F_r \quad (17)$$

We now combine compensation for schooling and for risk by writing the non-stochastic earnings option as a Mincer mark-up on the riskless no-schooling alternative. Thus, we write

$$E[Y_s] = \mu_s = (1 - \Pi_s)^{-1} (1 - M_s)^{-1} Y_0 \quad (18)$$

Suppose, we apply the CRRA simplification, such that

$$M_s = \frac{1 - e^{-\delta s}}{1 - \rho} \quad (20)$$

and we take Taylor expansions

$$-\ln(1 - M_s) = -\ln\left(1 - \frac{1 - e^{-\delta s}}{1 - \rho}\right) \approx -\ln\left(1 - \frac{\delta s}{1 - \rho}\right) \approx \frac{\delta s}{1 - \rho} \quad (22)$$

$$-\ln(1 - \Pi_s) \approx \Pi_s \quad (23)$$

Furthermore, rewrite the second and third-order terms in equation (17):

$$\frac{m_{2s}}{\mu_s^2} = \frac{E[(Y_s - \mu_s)^2]}{\mu_s^2} = E\left[\left(\frac{Y_s - \mu_s}{\mu_s}\right)^2\right] \quad (24)$$

$$\frac{m_{3s}}{\mu_s^3} = \frac{E[(Y_s - \mu_s)^3]}{\mu_s^3} = E\left[\left(\frac{Y_s - \mu_s}{\mu_s}\right)^3\right] \quad (25)$$

Then, the earnings function would read

$$E(\ln Y_s) = \ln Y_o + \frac{\delta}{1 - \rho} s + \frac{1}{2} \rho \frac{m_{2s}}{\mu_s^2} - \frac{1}{6} \rho(\rho + 1) \frac{m_{3s}}{\mu_s^3} \quad (26)$$

which is a simple equation in schooling years, variance term (24) and skewness term (25).

Hence with observations on relative variance and relative skewness we could estimate a Mincer earnings equation augmented with risk compensation, the *Risk Augmented Mincer equation* (RAM). Wages respond positively to risk and negatively to skewness when individuals decide on an education based on their knowledge of the second and third moment of the wage distribution associated with an education. If we don't assume CRRA,

the parameters of (26) will not be constant but depend on income levels. However, as a linearization, it would still be a good starting point for empirical work⁴.

By necessity, the model is a simplification. It is an extension, with stochastic rather than deterministic post-school earnings, of the framework specified by Jacob Mincer, still the basis for all routine estimates of returns to education. Mincer (1974, 9-11) did not even bother to spell out the strong underlying assumptions on market structure, individual abilities and information, except for a brief reference in the introduction. More elaborate estimates of the causal effect of schooling on earnings are often based on Card (1999). Card's model endogenises schooling, allows for unobserved heterogeneity in individuals' cost and benefits but does not allow for any imperfect information by individuals: they know all their cost and their prospective returns with certainty. Card's problem is an ignorant researcher looking at an omniscient individual. His model serves his purpose, but it is not the single best model to serve all purposes. Clearly, in our model we make very strong assumptions. But so did Mincer and so did Card. In fact, our model is quite similar to Levhari and Weiss (1974), the seminal paper on human capital investment under uncertainty. They also use a two period model with uncertain second period returns to investment made in the first period. As shown in Appendix 1, our basic equation can also be derived from their model.

⁴ In Diaz Serrano, Hartog and Skyt Nielsen (2004) we have extended the model to include permanent and transitory shocks in a longitudinal framework. In Hartog and Vijverberg (2002) we estimate parameters of the utility function in a model with new shocks at every year of experience.

A key assumption of our approach is that individuals cannot insure the risk of their investment. To us, this is obvious. We simply do not observe individuals commencing a college education and at the same time buying insurance or an optimal investment portfolio that completely eliminates the risk of their venture. Davis and Willen (2000) assume they would do so and compute optimal portfolios for some occupations. They find completely unrealistic values. For example, a 40-year old truck driver (in 1982) should hold a portfolio of \$550,000, including a short position in one portfolio of \$141,000. Our view is shared by e.g. Blanchard and Fisher (1989:283) and by Shaw (1996:626) who states: “The methods of reducing riskiness that are available in financial markets, namely, diversification, exchange, and insurance, are not options for reducing the riskiness of returns to human capital investments”. Palacios-Huerta (2003) studies the relationship between human capital risk and financial investment in a lifecycle consumption framework and reports supporting evidence. He finds that at the aggregate level, the mean-variance frontier does not improve if returns from financial assets are added to returns from human capital: adding an optimal financial investment portfolio does not improve the pay-off to risk taking. In the converse case (adding human capital to financial assets) the frontier does improve. For separate demographic groups, the results vary by level of education. Shaw (1996) reports a similar result, based on her own analysis and reference to earlier work: the covariance between human and financial wealth is zero, leaving no scope to reduce human capital risk by adequate financial investment.

Note that in the simple formulation of the model given above, the only uncertainty is the post-school wage rate. It's hard to see how this risk can be insured. Of course there is also substantial uncertainty about the further development of earnings during the individual's career. One might argue that an individual does not have to accept uncertain earnings over the lifecycle as an inescapable event, as it is always possible to apply consumption smoothing. However, this is of course not without cost. Consider uncertainty about income in a simple two-period model: all an individual knows is the probability distribution of income in the next period. A risk neutral agent (a bank?) would be willing to replace the individual's uncertain future income by its expectation $E\{Y\}$, which would then be the individual's sure consumption. But why would this agent leave all the utility gain to the individual? A bank certainly would charge a fee. The fee would probably be declining in $E\{Y\}$, as banks are more interested in attracting larger sums (they also pay higher interest on larger deposits). But the fee will no doubt be increasing in the variance of Y , as banks will not accept higher risk without compensation. They even dislike volatility, as is shown by the premium they pay on long-term deposits, or the penalty they charge for withdrawals. Home-made consumption smoothing is neither costless. Suppose an individual has random income Y in every period and seeks a stable consumption level C , by saving $Y-C$ if positive and borrowing $C-Y$ in the opposite case. Assume the bank will accommodate all the individual's actions, but it will do so at different interest rates for saving and borrowing. In a symmetric probability distribution, ignoring carry-over between periods (accumulating savings, or debts), the expected amount saved will be equal to the expected amount borrowed. But expected interest on savings will be less than expected interest on loans, and this will reduce C below $E\{Y\}$. The gap $E\{Y\} - C$ is a good measure of the cost of risk. Expected savings

and expected loans will increase with the variance of the distribution of Y . Hence the cost of risk will also increase in the variance. Intuitively, one expects it to decrease in skewness, as savings are favoured over loans. But these are of course just the predictions needed for the basic hypothesis. One may sense that this intuitive argument survives in a multiple period model where savings (and debts) can be accumulated, but with the outcome conditioned by the serial correlation of the income draws⁵. Of course, the literature on stochastic dynamic programming has been developed precisely because an individual cannot smooth consumption by replacing uncertain future income by its expected value (see e.g. Carroll, 2001; Low, 2005). There can be little doubt that lifetime welfare is declining in wage volatility.

3. Empirical specifications

Several specifications have been used to obtain measures of risk and skew, and below we will present a survey of estimation results. The first analysis, by King (1974), used variance and skew of earnings in occupational cells and estimated equation (5) at the aggregate level. Feinberg (1981) used a short panel to estimate each individual's coefficient of variation over 6 years. McGoldrick (1995) introduced the two-step procedure. With cross-section data, she first estimated an earnings function

⁵ Cochrane (2001) points out that in a lifetime welfare maximising framework variance as such is not relevant: it's the covariance of an asset with consumption that is essential. Levhari and Weiss (1974) also stress the key importance of the covariance between marginal utility of consumption and returns to education. We still denote the residual variance of wages as risk, as we abstract from optimal consumption profiles over the life cycle.

$$\ln Y_{ij} = X_i\beta + \sum_j \alpha_j d_j + \varepsilon_{ij} \quad (6)$$

where the subscripts i and j denote individuals and the education cell the individual belongs to respectively. The d_j are dummy variables for education cells (fixed effects). The variables included in X are years of education, age and age squared and, depending on specification, dummies for gender and ethnicity. Generally, no other explanatory variables in X are included, as the common variables that may be available (such as industry, firm and job characteristics) are all unknown to the individual when deciding on education⁶. The education fixed-effects α_j are included in order to control for the effect of omitted variables that may bias the measures of risk and skew within an education cell. Estimated residuals are used to compute measures of R and K

$$R_j = \frac{1}{N_j} \sum_i (e_{ij} - \bar{e}_j)^2 \quad K_j = \frac{1}{N_j} \sum_i (e_{ij} - \bar{e}_j)^3 \quad (7)$$

where e_{ij} is the exponential of the estimated residuals ε_{ij} in equation (6) and N_j is the number of observations in cell j . In (7), R and K are simply estimated as the second and third moment of the distribution of $\exp(\varepsilon_j)$. In the second step the estimated values for R and K are included in the wage equation

$$\ln Y_{ij} = X_i\beta + \gamma_R R_j + \gamma_K K_j + \varepsilon_{ij} \quad (8)$$

⁶ If the variables in X , like industry, firm size, etc would reflect perfectly compensating differentials associated with free choice, they would have to be included as the residuals would then be standardized on equal-utility wages. This assumption is too strong to accept.

Dummies for education cells cannot be included in (8) since R and K are already fixed in a given education cell.

Hartog and Vijverberg (2002) introduced a specification that is closer to the model in section 2. First, estimate

$$\text{Ln}W_{ji} = X_i\beta + \varepsilon_{ji} \quad (9)$$

where i indicates the individual and j indicates the occupation-schooling group that the individual belongs to. Years educated is one of the variables in the matrix X . Define σ_j^2 as the variance of the disturbance ε_{ji} in occupation/education cell j . Use the estimated parameter vector $\hat{\beta}$ and the estimated variance $\hat{\sigma}_j^2$ to predict the wage rate for each individual through:

$$\hat{W}_{ji} = \exp\left(X_i\hat{\beta} + \hat{\sigma}_j^2/2\right) \quad (10)$$

Finally, calculate wage deviations $W_{ji} - \hat{W}_{ji}$ and from these the relative variance R_j and relative skewness K_j , defined as

$$R_j = \frac{1}{I_j} \sum_{i=1}^{I_j} \left(\frac{W_{ji} - \hat{W}_{ji}}{\hat{W}_{ji}} \right)^2 \quad (11)$$

$$K_j = \frac{1}{I_j} \sum_{i=1}^{I_j} \left(\frac{W_{ji} - \hat{W}_{ji}}{\hat{W}_{ji}} \right)^3 \quad (12)$$

In (10), the variance term is added to the mean to reflect that the disturbances of the earnings distributions are approximately lognormal, as is commonly assumed. Were the distribution

indeed lognormal, equation (10) would hold exactly.⁷ R and K are the sample estimates of relative variance and relative skew, as defined in equations (24) and (25). In practice, (7) and (11)-(12) are equivalent: in Danish panel data (Diaz-Serrano, Hartog and Nielsen, 2004) the measures in (11) and (12) correlate better than 0.99 in each of 17 years with those in (7).

4. Estimates of the Risk Augmented Mincer Equation

In Table 1, we present estimates of the Risk Augmented Mincer Equation. *Grosso modo* one could say these are first generation estimates, where risk and skewness are measured within occupations. Whenever relevant, the regressions include a parabolic age (experience) profile. Sometimes there are additional controls, in particular in the second stage (the first stage should not control for effects that the individual cannot anticipate when choosing an education, such as e.g. firm size). The estimates strongly support the basic hypothesis of risk compensation. The coefficient for risk is positive in all but one case and significant, usually at high levels in all studies except the study with union interaction. The elasticity is mostly in the interval 0.1 – 0.2. The coefficient of skewness is negative, except in 4 out of the 19 cases, usually but not always at high levels of statistical significance. The elasticity is small, mostly below 0.10. The study with union interaction, by Moore (1995), focuses on benefits that unions bring their members. Compensation for wage risk is insignificant in both unionised and non-unionised jobs.

⁷ Hartog and Vijverberg (2002) test for log normality and mostly reject it. Still, adding the variance reduces the bias in the estimate of the mean.

Interestingly, in jobs covered by union contract workers have less wage variability and more hours variability, suggesting that unions manage to reduce wage risk, but pay for it with higher hours risk. They also bring higher compensation for wage risk than in the non-union sector, but the difference is statistically not significant.

The early studies and the replications suffer from flaws that have to be addressed before any firm conclusions can be drawn. First, earnings risk is measured at the level of occupations, and thus will be sensitive to selective mobility. Individuals who are not successful in an occupation may try their luck elsewhere⁸. With overrepresentation of workers with good draws, observed earnings overestimate the earnings of all those who tried this occupation. With truncation of the earnings distribution at the low end, observed variance of earnings is an underestimate of risk, while skewness is overestimated. Thus the risk coefficient will be overestimated. The effect on the coefficient of skewness cannot be predicted as both expected earnings and skewness are overestimated. Better estimates can be obtained with observations grouped by education, as one cannot escape bad draws after the education has been completed. If risk induced mobility between occupations is important, one must predict that education based estimates give lower

⁸ Johnson (1977) and McGoldrick and Robst (1996) report a higher compensation in occupations with less mobility to other occupations: lock-in effects are compensated.

Table 1 The Risk Augmented Mincer equation: occupations

Author	country	data/year	N	R, K	Elasticity R	Elasticity K	Controls
King (1974)	Table 2, row 3						
	USA	1960 Census	37 occupations (professions, 4 years college)	Standard deviation. Third moment	1.22 ^a (8.71)	-4.44 E-3 ^a (4.77)	Ability (Project Talent)
Johnson (1977)	USA	1970 Census	55-107 occupations	Standard deviation	0.11 ^{b)}	-	Within age-education groups
Feinberg (1981)	Table 1, column (1)						
	USA	1971-1976 PSID panel	1419 individuals	Individual coefficient of variation over 6 years	0.01 (2.53)	-	IQ, education, occupation; higher compensations for the more risk averse
McGoldrick and Robst (1996)	p. 230						
	USA	PSID 1979- 1984	528 women 937 men	Standard deviation residual earnings (time, time sq)	M: 0.17 (11.57) F: 0.50 (6.87)		Mobility, 7 occupations
Ma (2005)							
	China	1991 Urban Household Survey	41 education- occupation	$exp(\varepsilon_j)$	M:0.05 (1.90) F: -0.23 (8.14)	0.05 (7.92) -0.006 (0.96)	
Table 8.6			51 educations- occupations	$exp(\varepsilon_j)$	M:0.18 (3.70) F: 0.65 (10.75)	0.02 (0.80) -0.09 (3.26)	
Table 8.5		2000					
Hartog et al (2003)	Table 1,2						
	Netherlands West	OSA 1999	40 occupations	$exp(\varepsilon_j)$	M:0.18 (10.02) F:0.15 (3.96)	-0.05 (8.18) -0.06 (3.32)	6 industry dummies
	Germany East	SOEP 1992	41 occupations	$exp(\varepsilon_j)$	M: 0.11 (4.02) F: 0.25 (6.06)	-0.02 (3.07) -0.05 (4.42)	6 industry dummies
	Germany Portugal	SOEP 1992	28 occupations	$exp(\varepsilon_j)$	M: 0.12 (5.03) F: 0.08 (2.64)	-0.01 (0.91) -0.02 (1.57)	6 industry dummies
		Quadros Pesoal 1992	70 occupations	$exp(\varepsilon_j)$	M: 0.37 (33.35) F: 0.76 (30.65)	-0.17 (24.08) -0.03 (2.32)	6 occupation dummies, 7 industry dummies, tenure, firm size, ownership and age

Table 1 continued

Author	country	data/year	N	R, K	Elasticity R	Elasticity K	Controls
Hartog et al (2003) Table 1,2							
	Spain	Estructurra Salarial 1995	83 occupations	$exp(\varepsilon_j)$	M: 0.29 (58.11) F: 0.10 (11.65)	-0.02 (8.45) -0.02 (35.83)	Bargaining regime, public/private, 6 industry dummies, 6 occupation dummies, city size
Hartog and Vijverberg (2004) Tables 3 and 4							
	USA	NBER-CPS 1995-1999	129 education- occupation 104 education- occupation	Relative variance, skew Interquartile ranges	M: 0.16 (15) F: 0.05 (4) M: 0.44 (20) F: 0.08 (3)	-0.06 (26) -0.17 (14) -0.26 (12) -0.45 (15)	Elasticities, t-values averaged
Diaz-Serrano (2000) Table 7.3							
	Spain	EPF 1990	83 occupation	$exp(\varepsilon_j)$	M: 0.02 (7.4) F: 0.04 (4.2)	-0.02 (2.8) -0.10 (4.8)	Family status, region, industry, skill level
Moore (1995) Table 4							
	USA	PSID 1978- 1987	856 individuals	Individual's coefficient of variation	Union -0.08 (0.39) Non-union 0.09 (?)		Risk interacted with union dummy

t-values refer to estimated coefficient, not to elasticity

^{a)} Regression coefficients: units and means not reported

^{b)} Relative effect of one standard deviation on mean earnings, averaged over 18 age-education categories. All coefficients significant at conventional levels

Table 2 The Risk Augmented Mincer equation: education

Author	country	data/year	N	R, K	Elasticity R	Elasticity K	Controls
Diaz-Serrano, Hartog, Nielsen (2004) Table 2							
	Denmark	1984-2000	75 educations	$exp(\varepsilon_j)$	0.03 ^{a)}	-0.005 ^{a)}	Men, aged 30-40 clustered
Berkhout, Hartog and Webbink (2006) Table 1, all							
	Netherlands	LSO 1997 Elsevier/SEO	66 educations	$exp(\varepsilon_j)$	0.2 (3.69)	-0.1 (2.57)	Clustered
-	-	1996-2001	100 educations	$exp(\varepsilon_j)$	0.08 (4.53)	-0.04 (3.25)	Starting salaries, tertiary education clustered
Diaz-Serrano and Hartog (2006)							
Table 3, Model 1	Spain	1995	53 educations	$exp(\varepsilon_j)$	0.222 (16.50)	-0.034 (2.42)	Clustered
Model 2				relative error	0.202 (16.96)	-0.008 (10.00)	
Model 3				Interquartile ranges	0.035 (11.67)	-0.016 (2.50)	
Berkhout and Hartog (2006)							
	Netherlands	Elsevier/SEO 1996-2006	111 educations	$exp(\varepsilon_j)$	M: 0.05 (2.19) F: 0.04 (2.43)	-0.13 (5.50) -0.10 (6.43)	Regression on mean earnings by education, personal characteristics clustered; starting salaries
Hartog and Vijverberg (2007)							
	USA	NBER-CPS 1995-1999	DOT-SEO cells	Cell means	M: 1.07 (9.34) F: 0.86 (5.41)		Region, ethnicity, working conditions

^{a)} Elasticity averaged over 17 annual estimates (all significant); mean R and K from IZA DP 963
t-values refer to estimated coefficient, not to elasticity

estimated risk compensation than occupation based estimates, while the effect on estimated skewness affection cannot be predicted.

The studies in Table 1 also fail to recognize a complication in the estimated standard errors. As noted by Moulton (1986), with R and K measured at group level, errors may be correlated within these groups and estimated standard errors should be corrected for this clustering within cells⁹.

Table 2 presents second generation estimates, with risk and skewness measured by education and with standard errors corrected for clustering, or with mean earnings regressed on mean risk and skewness. The basic conclusion is unaffected: a positive effect of risk and a negative effect of skewness, both statistically significant. We cannot compare the magnitudes of the coefficients (elasticities) between Table 1 and Table 2, as they are based on different datasets and different controls, so we cannot test the prediction that education based measures lead to lower elasticities than occupation based measures. As we noted above, there is no difference between measuring risk and skewness with the exponential specification or as relative measures (equation (7) versus (11)-(12)). In Diaz-Serrano and Hartog (2006), we have also tested another specification, with R and K based on interquartile ranges. As R and K will be sensitive to outliers, we have calculated the percentile distribution of the residuals in the first stage, and defined R as the difference between the 75th and the 25th percentile and K as $(P_{75} - P_{25}) / (P_{50} - P_{25})$.

⁹ The Moulton problem is not relevant for Johnson (1977) and King (1974), who estimated on aggregate data by occupation and for Feinberg (1981), who estimated on panel data.

As can be seen both in Table 1 and in Table 2, we still obtain significant effects with the proper signs.

On top of the Moulton problem, we have to realise that R and K are measured from a first stage regression and will be subject to sampling errors. Murphy and Topel (1986) have presented a method for consistent estimation of the variances of the second stage parameter estimates. In Diaz Serrano and Hartog (2006), the corrections for clustering and for generated regressors are combined, by replacing the conventional OLS covariance matrix estimator in Moulton's adjustment by the covariance estimator proposed by Murphy and Topel. The Moulton correction for clustering is important, but the additional correction for generated regressors has negligible effect on estimated standard errors.

With R and K fixed for a given occupation or education, second stage regressions cannot include a fixed effect. Thus, it may be that R and K do not measure any dimension of risk but simply some fixed effect of occupation or risk that is not picked up by the controls. In Diaz Serrano, Hartog and Nielsen (2003) we have regressed education fixed effects from the first stage earnings functions on R and K (both permanent and transitory, see below). The effect of R and K on these fixed effects was not statistically significant, thus indicating that R and K are not just representing education fixed effects through the back door. In Diaz Serrano and Hartog (2006) we applied the same test to our estimates for Spain, with the same result. We can be pretty confident that R and K are not just unidentified education effects.

It's worth pointing out that our results appear not due to the effect of omitted variables on the quality of jobs. In Hartog and Vijverberg (2007) we report that the estimates are insensitive to the inclusion of measures of job disamenities such as physical burdens or exposure to toxic conditions and the like.

5. Bias from unobserved heterogeneities?

An immediate objection to using the observed distribution of earnings residuals is the possible confounding of risk and heterogeneity. The (residual) distribution of earnings will be affected by many variables that for an individual may not pose any risk at all. Probably the most prominent of these variables is individual ability. But individuals may also differ in risk attitudes and they may even face individually different risk. To address these issues, in Jacobs, Hartog and Vijverberg (2008) we set up a simple model with a safe fixed wage job (with wage W_s) and a risky job with stochastic wage, at expectation $E(W_r)$. We did not refer specifically to educations as the problem is identical in a situation where individuals can choose a job or an occupation without first spending time in school (the required Mincer compensation for postponing earnings can simply be added to earnings). We consider only risk and ignore skewness.

If individuals are identical in all aspects (risk aversion ρ , risk σ^2 and “ability” expressed as expected productivity in the risky job (in the safe job they have equal productivity anyway), in equilibrium the expected risky wage will just compensate for risk and allocation will be arbitrary. Risk is properly estimated as observed wage variance in the risky job; the coefficient of risk compensation is estimated as the observed wage gap divided by wage variance, without bias, from the equilibrium condition $E(W_r) = W_s + \rho\sigma^2/2$. This is a stripped down version of the model in section 2. Note that ρ is estimated from the wage gap divided by half the variance.

To assess the effects of heterogeneity, we distinguish three cases. First, we assume that individuals only differ in risk attitudes, and have identical abilities and risk. This poses no special problem. The wage gap between risky and safe jobs is now determined by the risk attitude of the marginal individual, and this is just what is estimated in the OLS regression. Second, if individuals only differ in the magnitude of risk and are otherwise identical, residual earnings variance underestimates true risk of the marginal worker, as only low risk individuals enter the risky job. As a consequence, risk compensation is overestimated. If individuals differ in both risk and risk attitude, the case is rather unusual. The threshold for entering the risky job is now defined as a critical value of the product of risk and risk attitude: the threshold for entry is a contour of combinations of the two variables that separate entry from non-entry in the risky job. There is no single critical value of any of these variables and by consequence they cannot be estimated. The coefficient that is estimated in the Risk Augmented Mincer equation can now be interpreted as the maximum value of risk aversion for the individual that has the mean

value of risk among those individuals who actually chose the risky job. The third case is where individuals only differ in ability, reflected in their expected earnings in the risky job; assuming individuals know their ability and the researcher does not, we have a garden variety of selectivity. The observed wage gap between the two jobs is now an overestimate of the wage premium, as the mean value of the risky wage also includes the effect of ability for those who chose the risky job (only more able individuals enter the risky job, as the ability premium should serve as a premium for accepting higher risk). Observed wage variance in the risky job is an overestimate of risk, as it also includes the variance of ability between individuals. With risk aversion estimated as the ratio between wage gap and observed variance, we cannot predict the sign of the bias, as both numerator and denominator are biased upward. To assess this case, we revert to simulation. We find that the ambiguity in the sign of the bias cannot be resolved by restriction to reasonable or plausible parameter values: depending on parameter values it may just as well be positive as negative. Considering the plausibility of parameter values we tend to conclude that underestimation of risk aversion is probably more likely than overestimation. But the key implication is that one cannot dismiss the risk compensation estimates as unreliable because they always produce an overestimated coefficient.

Below, we will discuss extensively whether ability is known to the individual and may generate an unobserved heterogeneity for the researcher, or whether individuals are also poorly informed. Here, we just note that observations on ability are available in several datasets, and that we are in a position to assess the effect of omitted ability bias in our estimates. Interestingly, the very first tests of the risk compensation hypothesis already

allowed for ability differences. King (1974) controlled for ability as measured in the Project Talent. Significant effects of risk and skewness were found while controlling for five ability measures: mathematics, English, reading comprehension, abstract reasoning and arithmetic reasoning. The basic conclusion was also upheld after splitting the sample in two classes of competing groups, technical and non-technical (a distinction confirmed by discriminant analysis on the basis of the five abilities). Estimation separately within the two groups, with their own measures of risk and skewness, confirmed the basic results. Feinberg (1981), testing on a panel with individual intertemporal variance as his measure of risk, found support in regressions that include IQ (a sentence completion test, on which no further information is given).

In Berkhout, Hartog and Webbink (2006) we analyse starting salaries for graduates from tertiary education in The Netherlands. Instead of using R and K for all individuals with the same education, we calculate these measures for ability quartiles as measured by the individuals' average grades for their secondary school final exam. This is a relevant stratification, as it is based on information for which we are sure that the individuals themselves also have it at their disposal. The model underlying the risk compensation hypothesis assumes that students perceive their risk and then decide on education. If desired, they could use the same information as we do, by stratifying on secondary school grades. As the results in Table 3 indicate, after controlling for ability in this way the basic results are upheld. In fact, if we compare the results in detail (differentiating by gender and two types of tertiary education, i.e. higher vocational and university), we find

regression coefficients for R and K that barely differ. There is no bias from omitting ability that can be picked up by secondary school grade quartile.

Table 3. Regressing log wage on risk R and skewness K; R and K measured within the individual's secondary school grade quartile (*SEO/Tertiary starters*)

	R	t	K	T	N
All tertiary					
All	2.67	3.95	-3.19	3.33	45020
Men	1.96	4.07	-2.66	3.98	19659
Women	2.44	3.01	-2.14	1.98	25361
Vocational					
All	2.99	4.80	-3.66	4.01	21431
Men	1.86	3.43	-2.55	3.61	8600
Women	2.97	4.59	-2.68	2.72	12831
University					
All	2.80	4.89	-3.83	4.40	23 589
Men	1.65	3.30	-1.92	2.19	11 059
Women	2.57	1.77	-3.07	1.74	12 530

t values based on robust standard errors clustered by education type. Grades only in second stage regression; first stage regressions separately by grade quartile. Source: Berkhout, Hartog and Webbink (2006).

Panel data are interesting for at least two reasons. By considering individual wage variation over time, unobserved heterogeneity is eliminated. Some of the first generation studies used panel data, and measured risk as the individual variance over time. Feinberg (1981) and McGoldrick and Robst (1996) each used 6 waves of the PSID and found statistically significant effects of individually measured risk, although the magnitude of the effect in the former study was very small. Moore (1995) measures risk individually from ten waves of the PSID and finds no significant risk compensation within either the union or the non-union sector, but instead finds that unions provide wage insurance (lower variance) at the price of higher hours variability.

Second, with panel data we can separate earnings variation between individuals from variation for given individuals over time. With a large and long Danish panel (1984-2000), Diaz Serrano, Hartog and Nielsen (2004) separate the earnings residual in an individual (random) effect and an annual transitory shock. We take the variance of the individual effects of individuals within the same education as the permanent risk of that education and the variance of the transitory component within an education as transitory risk, using 75 educations. If consumption smoothing holds, we expect no compensation for transitory risk, if individuals know their ability we expect no compensation for permanent risk. We insert the values in the wage equation for the pooled panel and we report the results in Table 4. The compensations for permanent and for transitory risk are both highly significant, though the elasticities are quite modest in magnitude. As to skewness, compensation for the permanent component is highly significant, with the right sign. For the transitory component, however, the prediction of a negative sign is squarely

rejected, with a significant positive effect. Again, the panel data estimates do not suggest an overestimation of compensation because of unobserved heterogeneity. We also have estimated the standard equation on each of the 17 years of our panel data. The average coefficient value from these 17 estimates is about equal to the estimate for the permanent component for both R and K, while the panel data estimate for the transitory component is much larger than annual average¹⁰. The results for the permanent component are at variance with the hypothesis that individuals know their own fixed effect and need no compensation for it, the results for transitory risk are at variance with full and costless consumption smoothing.

Comparable results are obtained with the NLSY panel for the United States, with the same risk decomposition (permanent risk from individual random effects, transitory risk from the annual transitory shocks), but with compensation for risk only tested on the wage equation for the final year (Raita, 2005). Risk and skewness are measured in education-ability cells, with two levels of education and 10 ability deciles (AFQT), resulting in 17 observations (as not all cells contain sufficient number of observations). Again one finds that risk compensation has the predicted sign, both permanent and transitory, although insignificantly so in one case. For skewness, the predicted effect for transitory shocks is found but permanent skewness has no significant effect.

¹⁰ Annual averages are in table 6 of our IZA Discussion Paper 963 (Diaz-Serrano, Hartog and Skyt Nielsen, 2003).

Table 4: Wage compensation for transitory and permanent shocks

Denmark ^a 1984-2000		
	Risk	Skewness
<u>Permanent</u>	0.3322 (17.90) <i>0.0390</i>	-0.0481 (17.84) <i>-0.0104</i>
<u>Transitory</u>	1.5727 (6.00) <i>0.0472</i>	4.1831 (8.74) <i>0.0128</i>
USA ^b 1979-2000		
<u>Using residual variation</u>		
<u>Permanent</u>	0.001 (0.99)	0.076 (1.32)
<u>Transitory</u>	0.309 (12.05)	-0.976 (11.94)
<u>Using earnings variation</u>		
<u>Permanent</u>	0.010 (1.92)	0.002 (1.15)
<u>Transitory</u>	0.297 (2.55)	-0.106 (3.99)

Notes:

a. Controls include years of education, age, age squared, and dummies for industry and occupation. Each cell contains coefficient, t-value in parentheses and elasticity in italics. Compensation estimated on pooled panel data. Source: Diaz Serrano, Hartog and Nielsen (2004)

b. Controls included for marital status, region, some occupation dummies, IQ test score (AFQT); the data are from NLSY 1979 (1979-2000). Compensation estimated on wages in 2000. Source: Raita, 2005, Tables 5.6, 5.7.

Unobserved individual heterogeneity is also filtered out if we use individuals' expectations. Schweri, Hartog and Wolter (2008) analyse data on the probability distribution of wages that Swiss students expect for themselves from different potential schooling scenarios: medians and dispersion measures elicited with the Dominitz and Manski method (see next section). Expected risk variables show a positive effect, expected skewness variables a negative effect on expected median wage. In fact, we even

get similar values for the elasticities: a risk elasticity of 0.12 and a skewness elasticity of -0.011 . These values are in the middle of the interval of values that we found in the empirical literature based on market data. The regression coefficients are only modestly reduced if we include individual fixed effects in the regressions.

One may have doubts that R and K properly measure the uncertainty they are meant to measure. In that respect it is reassuring that if we use dispersion measures derived from the percentile distribution of residual earnings, the results are not essentially different (Hartog and Vijverberg, 2007). Supporting evidence is also provided by Pereira and Martins (2002). They measure risk as the difference between returns to education in the highest and the lowest decile in quantile regression. Regressing the Mincer rate of return on risk across 16 countries, they find a significant positive effect with implied elasticity of 0.2. Even though there are only 16 observations, the evidence is neatly in line with the other results we report here.

6. Self-selection: what do students know?

The hypothesis of risk compensation in wages is based on the assumption that supply decisions by potential students affect market wages. Students are supposed to act on the financial uncertainty associated with an education. The estimations we presented use observed residual variances as an indicator of risk. It's by now a natural reflex of labour economists to assume a selectivity problem here, on the hypothesis that individuals are

well informed on their abilities and their income prospects. Selectivity corrections are based on the assumption that students, when they take their decisions, have better information on future returns than the researcher and precisely these decisions make the observed distribution of residuals a biased estimate of the actual uncertainty: we observe realised, not potential uncertainty. A proper econometric modelling and a proper estimation technique should retrieve the true dispersion that individuals base their decision on. If we just use the observed distribution of residuals, we confuse unobserved heterogeneity and risk. In section 5 we indicated that selectivity by no means leads to unambiguous biases and that we cannot discard the results as obviously spurious. Here, we will consider the issue of student information more generally.

There are very few empirical studies to inform us on this distinction between risk and heterogeneity. Chen (2008) uses the NLSY 1979-2000 and estimates wage equations and schooling choice. The wage equation contains a permanent and a transitory effect, the schooling choice contains unobserved heterogeneity. The transitory component cannot be foreseen by the individual, but individual uncertainty on the permanent component is reduced through its correlation with unobserved heterogeneity in the schooling choice equation: the unobserved heterogeneity factors allow the individual a better prediction of the permanent effect in wages (a lower variance) than the outside observer. Chen finds that wage uncertainty (the unforeseen transitory component in wages plus the incompletely foreseen permanent component) does not deviate much from observed wage inequality. Observed residual variances in wages, for less than high school, high school, some college and college are 0.526, 0.411, 0.500 and 0.512, respectively, while

calculated wage uncertainties are 0.485, 0.406, 0.497 and 0.472. That would make observed inequality a good indicator of uncertainty¹¹. In an earlier version of the paper, the selectivity correction on the variance of the permanent component was shown to be negligible (Chen, 2003). Chen and Khan (2007) focus on the average effect of education on wage inequality (residual variance): college versus high school or less. OLS estimation gives a ratio of 1.08, Heckman two-stage 1.09, an IV estimate gives 0.94 and pairwise matching gives 1.29. The absolute effect is also quite small. Straightforward OLS estimation of an earnings model yields residual standard deviation of 0.397 for college and 0.370 for high school graduates, while two-stage Heckman estimation gives 0.455 and 0.445 (o.c., Table 5 in the 2005 working paper version). Heckman, with several co-authors, has developed an elaborate framework to identify counterfactual distributions of earnings for various schooling choices. The method rests on the notion that if information on results that only become observable to the outsider in realised outcomes has an impact on schooling choices it must have been in the agent's information set when making the choice. In a survey paper, Cunha and Heckman (2007, p.892) conclude: "For a variety of market environments and assumptions about preferences, a robust empirical regularity is that over 50% of the *ex post* variance in the returns to schooling are forecastable at the time students make their college choices". Arcidianoco (2004) focuses on information on ability; with his econometric specification he estimates that 50-60% of the variance of an ability indicator related to major in university is noise.

¹¹ The numbers are not equal to those in Chen's Table 4, but have been corrected for obvious printing errors.

Rather than extracting evidence from observed behaviour through an imposed econometric model, one might simply question individuals about their information set. We should point out that here we are *not* interested in the true risk associated with an education. What matters in our case is the *belief* of potential students on the risks they are facing, as this is what they act on. We encounter a methodological divide here. Under a strict instrumentalist approach, it's irrelevant what individuals believe: all that matters is the quality of the prediction. Alternatively, one may argue that if individuals are inevitably to base their choices on expected outcomes, one might as well study these expectations directly, as this can only improve our understanding of individuals' behaviour.¹²

The direct approach to eliciting perceptions on the risk of returns to education has been pioneered by Dominitz and Manski (1996). Their study is rather exploratory in methodology, with only 71 high school students and 39 college students, but it is very carefully executed and gives unusual insights. Students are asked to answer questions in a computer-executed survey, with feedback to correct errors. First, they are asked the median salary for a given educational scenario (complete high school), with an explanation of what a median is. Next, the program calculates 75% of the median salary and asks for the probability that the student's future salary will not surpass this level, and it asks for the probability of surpassing 125% of the median, with feedback on answers that violate the rules of probability. Assuming a normal distribution of future salaries, the answers are sufficient to calculate the individual's parameters of this distribution.

¹² See Blaug (1980) on the Lester-Machlup controversy (Chapter 7) and on Friedman's methodological stance (Chapter 4).

Students appear to have widely divergent anticipations. For example, among male high school students asked for expected median earnings at age 30 in the scenario where they will have completed a bachelor degree, the 10th decile expects 25 000 dollars whereas the 90th decile expects 56 000 dollars. On the dispersion in their future distribution of earnings, the same groups, under the same scenario, expect interquartile ranges of 28 000 and 58 000 dollars. Perceived returns also differ widely. At age 30, the male high school students at the 10th percentile expect an earnings advantage from college of 10 000 dollars, at the 90th percentile they expect a gain of 30 000 dollars. Students were also asked to state the actual dispersion of earnings by education. Generally, they overestimate the interquartile range. Interestingly, predictions of their own median expected salary correlate positively with their perception of the actual median: “Respondents who believe current median earnings to be high (low) tend also to expect their own earnings to be high (low)” (o.c., p 25).

Wolter and Weber (2004) apply the Dominitz and Manski approach to Swiss students. They also find large dispersion in expected medians (a coefficient of variation across individuals of about 0.20) and considerable uncertainty as reflected in the interquartile range. Americans’ individual uncertainty is larger than actual interquartile ranges, whereas for Swiss students individual uncertainty is smaller than actual dispersion. The Swiss students are asked for individual wage expectations under two scenarios, secondary education and tertiary education. These estimates are not significantly different from actually observed medians for these groups (the mean signed error is not significantly different from zero). Most interestingly, the deviations do not differ between respondents

with different actual schooling choice (students in high school, in a business college or a university of applied science), neither for men nor for women, neither for age 30 nor for age 40. This does not point at selectivity related expectations. Rather, the outcomes are compatible with all students anchoring their expectations on actually observed wages.

Schweri, Hartog and Wolter (2008), also argue against a pervasive role for superior private information. Using another set of data on Swiss students collected with the Dominitz and Manski method, they report that individuals' expectations on the distribution parameters of earnings for specified education choices are related to individuals' perceptions on these parameters for employees already in the labour market and who have realised these choices. That is, an individual's expectations mirror just what they see in the labour market as realised outcomes¹³. Even more interestingly, deviations between an individual's predictions for himself and his perceptions of realised market outcomes for those who have chosen the particular schooling scenario are not systematically related to variables that could reveal private information. Neither family background nor ability (secondary school grade for math) explain the differences. This is similar to results reported by Nicholson and Souleles (2001). They use survey data from students at a particular medical school in the US to assess the formation of income expectations and their relevance for explaining the choice of medical specialty. Income expectations are strongly anchored on perceived contemporaneous incomes, without exclusively determining them. For every 1000 dollar increase in the actual

¹³ McKenzie, Gibson and Stillman (2007) underscore the value of direct measurement of expectations. They also apply the Dominitz and Manski method and find that potential emigrants from Tonga underestimate earnings in New Zealand, probably because actual emigrants under-report to restrict transfer obligations to relatives back home.

contemporaneous income of the specialty that the student plans to enter, the student's expectation for that specialty goes up by 590 dollars. For every 1000 dollars difference between the student's estimate of contemporaneous specialty income and actual specialty income, the student's expected income goes up by 840 dollars. Misperceptions in actual specialty income end up almost dollar for dollar in income expectations. By contrast, the effect of "ability" is very small: performing in the top quartile of the exam at the end of the second year in medical school only increases expected income by 5.9%.

Results by Brunello, Lucifora and Winter-Ebmer (2004) are also at variance with anticipations from the private information hypothesis. They collected information on expected earnings from university students in ten European countries (business and economics). The expected wage premium over high school graduates at labour market entry was unrelated to any variable except age: not to parental background, not to channel of information about future earnings (university publication, career center, special reports, press, personal communication), not to reason for choosing their selected university, not to self-assessed relative ability. The expected premium of university over high school graduation after ten years of experience is only significantly lower for older students, women and students with longer expected time to complete the degree (and again not related to self-assessed ability). These results only refer to university students, not to those who choose not to go to university, but they do not point to systematic patterns in expected benefits, except for the effect of expected length of study, which possibly reflects an update of initial expectations during the study.

Perhaps the strongest blow to the notion of superior information about one's future position in the wage distribution is from Webbink and Hartog (2004). First-year students can predict differences in mean salaries by fields of education but they cannot predict their own starting salary after graduation, four years ahead. The correlation between an individual's starting salary after graduation predicted upon entering tertiary education and the realisation after graduation is a mere 0.06.

We conclude that direct measurement of potential students' information points to a large dispersion between individuals and to substantial perceived uncertainty by individuals on their own prospects. The large dispersion indicates that students do not all just use observed distributions for graduates already in the market as their own expected distribution. But the differences in their expectations are neither compatible with an obvious pattern of private information: they are unrelated to self-assessed ability. Even more telling: Dutch first-year students appeared unable to predict their position in the distribution of starting salaries. There is evidence, however to support the hypothesis that students anchor their expectations on their perceptions of actual rewards in the labour market, but differ in these perceptions. This is certainly a hypothesis worthy of further investigation. Econometric modelling to extract information from observed behaviour lends more support to the relevance of self-selection, although the magnitude of the correction for private information is uncertain and not necessarily large.

7. Supporting evidence

To consider the credibility of the results, we may consider some circumstantial evidence. First of all, we may note that compensation for risks in the labour market is well established, both for instability of employment and for health and morbidity hazards (Rosen, 1986). In fact, employer behaviour exhibits mirror effects of the risk compensation we analyse here. Fresh graduates pose a risk for employers to the extent that they cannot accurately assess their potential productivity. But as graduates have no alternative for putting their education at work, we expect employers to shift the risk of unknown abilities to employees. Using variance and skewness of students' grades in different disciplines as indicators of employer uncertainty we find that starting wages are lower in fields with high variance and higher in fields with high skewness (Berkhout, Hartog and van Ophem, 2009).

More specifically, we will cite evidence that individuals care about financial risk when choosing a career, that structural parameters estimated for risk compensation are in line with estimates in life cycle consumption models and that more risk loving workers experience higher wage growth.

Evidence that individuals care about the financial risk when choosing a labour market position is given by King (1974) and by Saks and Shore (2003). They note that with risk aversion declining in wealth, one should expect that students from wealthier backgrounds choose more risky occupations. This is precisely what they find in American data. Lazear (2005) reports that individuals are more likely to start their own business firm if the first industry in which they were employed after leaving university has high wage variance,

suggesting they select themselves into positions to match their relatively low risk aversion. Diaz-Serrano (2005) and Dohmen et al. (2007) document such selection more explicitly. Diaz-Serrano measures earnings uncertainty as residual variance in Italian panel data after removing individual fixed effects and finds that earnings uncertainty correlates significantly with the stated reservation price for a lottery ticket: more risk averse individuals have less earnings uncertainty. Dohmen et al use a measure of individual risk attitude obtained from a survey question (validated in experimental data) and measure earnings risk as residual earnings variance by occupation. They report significant correlation between risk and individual risk attitude: more risk averse individuals sort into jobs with less earnings variance. Negative results are reported by Kodde (1985), who uses Dutch questionnaire data on individual students' expectations. He finds no clear effects of the gap between an individual's expected highest and expected lowest earnings, after university education, to attend university.

Belzil and Leonardi (2007) measure risk attitudes from answers on the reservation price for a hypothetical lottery. Conditional on completing senior high school (in Italy), the probability to go on to university appears to be decreasing in risk aversion, with an effect as large as that of parental education. However, conditional on junior high school the decision to attend senior high school is increasing in risk aversion. That would imply that individuals see a university education as risky and senior high school education as an insurance. Hartog and Diaz Serrano (2007) find that Spanish high school graduates are less inclined to go to university in provinces where the return to university is more risky (as measured by residual earnings variance), and that the effect is indeed mitigated by

declining risk aversion. Flyer and Rosen (1997) derive that someone who in her lifecycle anticipates to devote less time to market activity should behave as if she were more risk averse. Indeed they find in CPS two-digit occupation data that the coefficient of variation of women's wages correlates negatively with the percentage of women in the occupation: women flock into less risky occupations.

Shaw (1996) argues that on-the-job training is a risky investment and that less risk averse individuals will invest more; hence, less risk averse individuals should have higher wage growth. She measures risk attitudes from the share of wealth held in risky assets and from direct survey questions. Indeed, less risk averse individuals have higher wage growth. If the share of financial wealth invested in risky assets increases from the sample mean of 0.1 to 0.4, the share of human capital devoted to investment in creating new human capital would also increase fourfold. The 3-year growth rate of the wage would then increase from 6.6% to 7.8 %, i.e. an increase by 18%. We can use some of her results to calculate that compensation for earnings risk as emphasized in this paper has an elasticity of about 0.5: moving from the class of above-average risk takers to the class of average risk takers, residual log wage variance falls by 50% while the average wage falls by 30%, moving to the class of no-risk takers, residual log wage variance falls by 70%, while the average wage falls by 40%. Shaw's results provide strong evidence in favour of our approach. Her results are not based on *assumed* risk taking behaviour, they hold for individuals who differ in stated risk attitudes. Individuals who indicate that they take more financial risk in order to obtain larger gains have higher average wage and higher residual wage variance. The same relationship holds in growth rates: less risk averse individuals

have higher wage growth rate and higher variance of residual wage growth¹⁴. In a replication study for the US, Germany, Spain and Italy (Budria et al, 2009), we confirm that risk attitudes are relevant for wages (support for the original Shaw specifications is mixed though, and we find that in particular the wage level is sensitive to attitudes towards risk taking, more so than wage growth). Brown and Taylor (2005) report support for the Shaw model in British data.

Hartog and Vijverberg (2002) estimate structural parameters in wage compensation equations like (26) for US data. The estimated coefficient of relative risk aversion is 0.64 for men and 0.46 for women, which is in the low end of the interval of estimates from various sources, such as consumption–savings models, TV games and direct surveys. The coefficient of relative skewness affection is estimated at 1.03 for men and 2.18 for women, also well within the interval of results found elsewhere in structural models (e.g. using a CRRA utility function and thus directly connecting risk aversion and prudence, Hubbard et al, 1994 take relative risk aversion at 3 and relative prudence at 4 as representing empirical findings, while Keane and Wolpin, 2001, report values of 1.5 and 2.5 respectively; generally, the range of values found in the empirical literature is very wide). The interesting observation is that the implied preference parameters are very much in line with similar values found in quite different applications, and that attitudes towards financial risk in the labour market are not fundamentally different from attitudes revealed in consumption-saving decisions. Skyt Nielsen and Vissing-Jorgenson (2006) estimate a model of educational choice for 50 Danish post-high school educations, characterising returns to education from a long panel of individuals, correcting the

¹⁴ The effect of the increase in the risky asset share is given in Shaw's Table 2. Changes in mean wages by risk class are given by the regression results at the bottom of page 641. Variances by risk class are given in footnote 22.

earnings data for selectivity. They find a clear negative effect of earnings risk on choosing an education, in particular for the permanent component of the variance. They use a utility function with constant relative risk aversion and estimate the coefficient at 5. This is much higher than what Vijverberg and Hartog (2002) find, with a more flexible utility function, but still easily within the wide range of commonly found values.

8. Alternative explanations

One might conjecture that the observed relationship reflects something other than risk compensation. One argument might be that earnings distributions obey the lognormal distribution, in which a relationship between mean and variance is implied. It is well known that when a variable W has a lognormal distribution with parameters μ and σ^2 , there is a linear relationship between the mean and standard deviation of W : $E[W] = \exp(\mu + 0.5\sigma^2)$ and one may write the variance of W as $Var(W) = (E[W])^2 \exp(\sigma^2 - 1)$. Thus mean and variance of W will both increase if the log variance increases, and even the relative variance (the coefficient of variation) will increase. Thus, one might argue that we only have reproduced a property of the lognormal distribution. There are several arguments to counter this interpretation. First, lognormality does not explain anything: even if we were to observe that a variable perfectly obeys a lognormal distribution, we still want to know why this is so. Risk compensation might be precisely the argument. Second, lognormality is not an iron law

and statistical tests often reject it (as we show in Hartog and Vijverberg, 2002 for our type of data). Third, in Diaz-Serrano, Hartog and Nielsen (2003) we find risk compensation in earnings also if we measure risk as the intertemporal variation in earnings from panel data; in this case there is even less reason to believe that we just reproduce an iron law of distribution. Fourth, a reasonable model of schooling choices and resulting earnings distributions by education does not imply a necessarily positive relationship between mean and variance. Suppose, schooling just slices off successive segments of an ability distribution and transforms ability monotonically into marketable skills. Then, mean income by schooling will increase, but dispersion measures will depend on ability dispersion within the segments and the schooling transformation function; such a model may easily exhibit earnings variance declining in schooling level. Finally, the most compelling counter-argument is that it would require skew also to be positively related to the mean, as skew for a lognormal variable equals $(2 + \exp \sigma^2) \sqrt{\exp \sigma^2 - 1}$. However, we consistently find a negative relationship.

Could it be that an underlying skill price effect drives the result? If earnings emerge as the reward for a skill, with the price of the skill determined in the labour market, an increase in the market price of the skill, with stable underlying skill distribution, will increase both mean and variance of earnings. But if so, the third moment of the earnings distribution should also go up.

Tournament theory (Lazear and Rosen, 1981) also implies a relationship between mean and variance: if shocks to performance become more important, both mean and variance

of wages will be affected. With risk neutral agents, optimal effort is independent of the variance of the shocks to output¹⁵; increasing shock variance does not affect output or mean wage. The spread between the two wages increases as the shock variance increases¹⁶. Hence, as one would expect, if agents are risk neutral, with wage spread increasing and mean wage unaffected, there is no compensation for increasing risk. In case of risk averse agents, the relationship between mean and dispersion may be positive or negative, depending on parameter values (see Appendix 2).

Job search theory has a systematic explanation for the persistence of wage differences among observationally identical individuals, pointing to the cost of mobility and of searching for information. A distribution with a large dispersion presents ample opportunity to find a high wage, and one might think the theory yields the same predictions as the risk compensation model. This is not the case however: in job search models a positive association between mean and dispersion of observed wages *may* arise, but a negative relationship is a distinct possibility. In the simplest possible model, unemployed workers sample from a wage offer distribution with a reservation wage. As Mortensen (1986, p 864) shows, an increase in the dispersion of the wage offer distribution increases the reservation wage. If one assumes the wage offer distribution to be normal, an increase in the truncation level from a higher reservation wage will increase the mean of the observed wage distribution (realised offers) while the effect on

¹⁵ Equation (9) in Lazear and Rosen (1981) and page 31 in Lazear (1995).

¹⁶ Equation (10), Lazear and Rosen (1981): increasing shock variance reduces $g(0)$.

the variance cannot be unambiguously signed.¹⁷ Thus, this model does not unambiguously predict the observed positive relationship between mean and variance of observed wages. This basic mechanism remains in more complicated models.¹⁸ In a more informal sense one might point to the appreciation that individuals will have for a higher dispersion in the job offer distribution, as this will bring them higher expected wage. This would have a negative effect on mean wages, as supply increases. If the higher wage is to be realised from repeated search, after entering the labour market, and if the higher variance in the offer distribution translates into higher variance in the observed wage distribution, one may perhaps expect a lower starting wage in high variance educations, as individuals appreciate the opportunity to search in a distribution with great opportunities, and a higher mean wage after sufficient experience, when search has paid off. Thus, while in this case predictions are identical for advanced careers, the effect on early careers stages has the power to discriminate. The SEO/Elsevier data analysed in Berkhout, Hartog and Webbink (2006) are restricted to starting wages, 1 to 2 years after graduating from tertiary education. As Table 3 above shows, starting wages respond positively to risk (variance), which does not support the merger of risk preferences and search theory. The role of skewness in search theory has not been investigated. Intuitively, one expects opposing effects: the opportunity to arrive in the high end of the distribution, by prolonged search, will push up mean earnings. But as individuals like this opportunity, they will accept a lower wage. Again, one might differentiate by career stages. At young ages, the effect of increased supply dominates

¹⁷ See Maddala (1983), p.365. The conditional mean M of the truncated normal distribution is positive in threshold c . The conditional variance $V= 1 - M(M-c)$ cannot be signed as M is positive in c but $M-c$ is negative, as the derivative of M to c is smaller than 1.

¹⁸ In Gautier and Moraga Gonzales (2003), wage dispersion is non-monotonic in labour market tightness: it first increases and then decreases with increasing tightness.

and the wage will be depressed. Later, when individuals have moved up through search, the wage should be boosted. But as pointed out above, the estimated effects for starters in the labour market have no different signs¹⁹.

9. What have we learned, what remains to be done?

Our empirical results have quite convincingly shown that wages in an occupation/education relate positively to the residual variance and negatively to the residual skew of wages within that occupation/education. These results are compatible with risk averse individuals demanding compensation for risk and willing to pay for favourable odds to obtain really high wages. We have given several supporting arguments that make this a plausible link.

The magnitudes of the elasticities are not large. Interestingly, including the risk compensation terms in the Mincer equation has generally no effect on the estimated rate of return to education. This indicates that the estimated return does not suffer from omitted variable bias if risk is ignored, counter to the suggestions of Weiss (1972), Olson, White and Shefrin (1979) and Low and Ormiston (1991). The outcome is compatible with the finding that (residual) earnings dispersion within schooling categories has no

¹⁹ One commentator has remarked that workers can continue searching on the job and that there is repeated sampling rather than a single draw. This would imply that risk is not properly measured by observed wage variance. However, the argument is more complex. Repeated sampling keeps chopping off the lower end of the offer distribution, with presently earned wages as threshold. If the variance of the offer distribution increases, this will increase the mean of observed wages. But as the effect on observed variance cannot even be predicted for a single draw, the effect on observed variance in this case will be even harder to predict. Hence, we cannot predict the relationship between mean and variance of observed wages.

robust standard pattern in relation to the level of schooling (schooling and risk are uncorrelated, as was demonstrated empirically in Raita (2005, Chapter 4)).

Although we believe that risk compensation in wages is genuine, a true test on causality would be most welcome. However, so far we have been unable to find a natural experiment or a good instrument for risk. A relevant concern is the correlation of risk attitude with other variables that may enhance earnings. If high-risk activities attract individuals with more ambition, greater drive, lower discount rate, then the reward for these traits will show up in higher earnings that give the appearance of compensation for risk. Note however that this argument does not seem to hold for ability, to the extent that the methods used to control for ability are sufficient (our use of school grades, IQ measures used in some regressions, using individual expectations and including individual fixed effects as regressor).

More can be done, both in theoretical modeling and empirical testing. For example, it might be useful to filter out compensation for choice variables from the wage equation and use a standardised wage net of compensation for disamenities. If people prefer working for a small firm rather than a big firm then the wage reduction may just equalise the utility from working for different sized firms and the associated wage variation should not be part of risk. But this assumes that all such variations are fully voluntary and individuals can freely choose from their wage frontier, which may not be true. Many such variables are not known in advance and can neither be freely adjusted at zero cost. In that sense they may well be part of the individual's uncertainty and be rightly included in the

wage dispersion measure. Still, from a theoretical perspective, one might prefer to measure potential wages, standardized for such effects to the extent possible.

We also think that there are strong arguments to accept the observed residual wage distribution as the agents' source of information on risk, and that there is no reason to apply a correction for unobserved heterogeneity. Indeed, in those cases where (some) unobserved heterogeneity was accounted for (controlling for ability, using individuals' predicted wages rather than market wages), the basic conclusions were upheld. But this is a very controversial issue. At the same time it points to a most interesting direction for further research: uncovering of information that potential students use when they decide on their schooling. Heckman has set a challenging standard by claiming that at least half of the ex post variance in earnings is foreseen by individuals ex ante. Direct survey information from students suggests much more ignorance. Investigating students' information sets is not only relevant for the issue tackled in this paper, but has far wider relevance in the economics of education.²⁰

²⁰ Nicholson and Souleles (2001) report that individually expected income is a better predictor of specialty chosen by medical students than actual income, whether contemporaneous or realized. Predicting specialty choice with actual contemporaneous average income by specialty, uncorrected for selection bias, is correct in 57.3% of the cases. Using, instead, individually expected incomes with selectivity correction for actually chosen specialty yields correct predictions in 85.6%. Using realised average incomes as predictors of future incomes yields 50.6% correct predictions, whereas using expectations on the same subsample gives 72.8 correct predictions. These differences are also affected by the difference between using individual expectations and group averages, with much less variation.

References

- Arcidianoco, P. (2004), Ability Sorting and the Returns to College Major, *Journal of Econometrics*, 121 (1-2), 343-375
- Arrow, K. (1965), *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjo Hahnsson Foundation.
- Belzil, C. and M. Leonardi (2007), Can risk aversion explain schooling attainments? Evidence from Italy, *Labour Economics*, 14(6), 957-970,
- Berkhout, P. and J. Hartog (2006), Starting wages respond to employer risk, Working Paper University of Amsterdam
- Berkhout, P., J. Hartog and H. van Ophem (2009), Starting wages respond to employer risk, Discussion Paper Tinbergen Institute 09.071/3, Amsterdam (revision and extension of Berkhout and Hartog 2006)
- Berkhout, P., J. Hartog and D. Webbink (2006), Risk compensation under worker heterogeneity, Working Paper Universiteit van Amsterdam, forthcoming in *Southern Economic Journal*
- Blanchard, O. J. and S. Fisher (1989), *Lectures on macroeconomics*, Cambridge, Mass: MIT Press
- Blaug, M. (1980), *The methodology of economics*, Cambridge: Cambridge University Press
- Brown, S. and K. Taylor (2005), Wage Growth, Human Capital and Financial Investment, *Manchester School*, 73(6), 686-708.
- Brunello, G., C. Lucifora and R. Winter-Ebmer (2004), The wage expectations of European business and economics students, *Journal of Human Resources*, 39 (4), 1116-1142
- Budria, S., L. Diaz Serrano, A. Ferrer Carbonell and J. Hartog (2009), Risk attitude and wage growth: replication and reconstruction, IZA Discussion Paper
- Card, D. (1999), The causal effect of education on earnings, in O. Ashenfelter and D. Card (editors), *Handbook of Labor Economics*, Chapter 30, volume 3A, Amsterdam: North-Holland
- Carroll, C. (2001), A theory of the consumption function, with and without liquidity constraints, *Journal of Economic Perspectives*, 15 (3), 23-45

Chen, S. (2003), Does college attendance reduce wage volatility?, Discussion Paper, SUNY Albany

Chen, S. (2008), Estimating the variance of wages in the presence of selection and unobserved heterogeneity, *Review of Economics and Statistics*, 90 (2), 275-289

Chen, S. and Khan (2007), Estimating the causal effect of education on wage inequality using IV methods and sample selection models, Working Paper SUNY-Albany

Cochrane, J. (2001), *Asset pricing*, Princeton: Princeton university Press

Cunha, F. and J. Heckman (2007), Identifying and estimating the distributions of ex post and ex ante returns to schooling, *Labour Economics, Special issue: Education and risk*, 14 (6), 870-893

Davis, S. and P. Willen (2000), "Occupation-level income shocks and asset returns: their covariance and implications for portfolio choice", Working Paper 7905, NBER.

Diaz Serrano, L. (2000), Human capital, progressive taxation and risk aversion, PhD Dissertation, Universitat Rovira i Virgil, Spain

Diaz Serrano, L. (2005), On the negative relationship between labor income uncertainty and homeownership: risk aversion versus credit constraints, *Journal of Housing*, 14 (2), 109-126

Diaz Serrano, L. and J. Hartog (2006), Is there a risk-return trade-off across educations? Evidence from Spain, *Investigaciones Economicas* 30 (2), 353-380

Diaz Serrano, L., J. Hartog and H. Skyt Nielsen (2004), Compensating wage differentials for schooling risk in Denmark, Discussion Paper Maynooth/Amsterdam/Aarhus, an earlier version was published as IZA DP 963 in 2003; revision 2007; forthcoming in *Scandinavian Journal of Economics*, 110 (4), 2009

Dohmen, T., A. Falk, D. Huffman en U. Sunde (2006), Individual risk attitudes: new evidence from a large representative experimentally validated survey. IZA Discussion Paper 1730

Dominitz, J. and C. Manski (1996), Eliciting student expectations of the return to schooling, *Journal of Human Resources*, 31, 1-26.

Feinberg, R.M. (1981), Earnings-risk as a compensating differential, *Southern Economic Journal*, 48, 156-163.

Flyer, F. and S. Rosen (1997), The new economics of teachers and education, *Journal of Labor Economics*, 15 (1, part 2), S104-S139

- Garrett, T.A. and R.S. Sobel (1999), Gamblers favor skewness, not risk: further evidence from United States' Lottery games, *Economics Letters*, 63, 85-90.
- Gautier, P. and J.L. Moraga-Gonzales (2003), Strategic wage setting and coordination frictions with multiple applications, Working paper Tinbergen Institute
- Golec, J. and M. Tamarkin (1998), Bettors love skewness, not risk, at the horse track, *Journal of Political Economy*, 106 (1), 205-225.
- Gollier, C. (2001), *The economics of risk and time*, Cambridge, Mass: The MIT Press
- Hartog, J. and L. Diaz Serrano (2007), Earnings risk and demand for higher education, *Journal of Applied Economics*, X (1), 1-28
- Hartog, J., H. van Ophem and S. Bajdechi (2007), Simulating the risk of investment in human capital, *Education Economics* 15 (3), 259-275
- Hartog, J., E. Plug, L. Diaz Serrano and J. Vieira (2003), Risk compensation in wages, a replication, *Empirical Economics*, 28, 639-647
- Hartog, J. and W. Vijverberg (2002), Do wages really compensate for risk aversion and skewness affection?, Bonn: IZA Discussion Paper 426
- Hartog, J. and W. Vijverberg (2004), On compensation for risk aversion and skewness affection in wages, Working Paper University of Amsterdam, shorter version in *Labour Economics, Special Issue on Education and Risk*, 14 (6), 938-956 (2007)
- Hartog, J. and W. Vijverberg (2007), Schools, skills and risk, *Economics of Education Review*, 26 (6), 758-770
- Hubbard, R., J. Skinner and S. Zeldes (1995), Precautionary saving and social insurance, *Journal of Political Economy*, 103 (2), 360-399
- Jacobs, B., J. Hartog and W. Vijverberg (2008), Self-selection bias in estimated wage premiums for earnings risk, *Empirical Economics*, forthcoming (available at the journal's website)
- Johnson, W. (1977), Uncertainty and the distribution of earnings, in: F. Juster (ed.), *The distribution of economic well-being*, Cambridge, Mass: Ballinger for NBER, 379-396
- King, A.G. (1974), Occupational choice, risk aversion and wealth, *Industrial and Labor Relations Review*, 586-596.
- Keane, P. and K. Wolpin (2001), The effect of parental transfers and borrowing constraints on educational attainment, *International Economic Review*, 42 (4), 1051-1103

- Kodde, D. (1985), Microeconomic analysis of demand for education, PhD thesis Erasmus University Rotterdam
- Lazear (1995), *Personnel Economics*, Cambridge, Mass: MIT Press
- Lazear, E (2005), Entrepreneurship, *Journal of Labor Economics*, 23 (4), 649- 680
- Lazear, E. and S. Rosen (1981), Rank-order tournaments as optimal labor contracts, *Journal of Political Economy*, 89 (5), 841-864
- Levhari, D. and Y. Weiss (1974), The effect of risk on the investment in human capital. *American Economic Review*, 64 (6), 950-963.
- Low, H. (2005), Self-insurance in a life-cycle model of labour supply and savings, *Review of Economic Dynamics*, 8 (4), 945–975
- Low, S. and M. Ormiston (1991), Stochastic earnings functions and risk, *Southern Economic Journal*, 57 (4), 1124-1132
- Ma, X, (2005), Research on the risk of individual investment in education--using the Urban Chinese in the 1990s as an example, PhD dissertation Peking University
- Maddala, G. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge: Cambridge University Press
- McGoldrick, K. (1995), Do women receive compensating wages for earnings risk? *Southern Economic Journal*, 62, 210-222.
- McGoldrick, K. and J. Robst (1996), The effect of worker mobility on compensating wages for earnings risk, *Applied Economics*, 28, 221-232.
- McKenzie, D., J. Gibson and S. Stillman (2007), A Land of Milk and Honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?, Bonn: IZA DP 2788
- Mincer, J. (1974; reprinted 1993), *Schooling, experience and earnings*, Aldershot UK: Gregg Revivals
- Moore, M.J. (1995), Unions, employment risks, and market provision of employment risk differentials, *Journal of Risk and Uncertainty*, 10(1), 57-70.
- Mortensen, D. (1986), Job search and labor market analysis, Chapter 15 in O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics*, Volume 2, Amsterdam: North-Holland

Moulton, B. R. (1986), Random group effects and the precision of regression estimates, *Journal of Econometrics*, 32, 385-397

Murphy, K. and R. Topel (1985), Estimation and inference in two-step econometric models, *Journal of Business and Economic Statistics*, 3, 88-97

Nicholson, S. and N. Souleles (2001), Physician income expectations and specialty choice, Cambridge: NBER Working Paper 8536

Olson, L. H. White and H. Shefrin (1979), Optimal Investment in Schooling When Incomes are Risky, *Journal of Political Economy*, 87 (3), 522-539

Palacios-Huerta, I.(2003), An empirical analysis of the risk properties of human capital returns, *American Economic Review* 93 (3), 948-964

Pereira, P and P. Martins (2002), Is there a return-risk link in education? *Economics Letters*, 75 (1), 2002, 31-37

Pratt, J. (1964), Risk Aversion in the small and in the large, *Econometrica*, 32, 122-36.

Raita, S. (2005), The risk of investment in human capital, PhD thesis Tinbergen Institute Universiteit van Amsterdam

Rosen, S. (1986), The theory of equalizing differences, in O. Ashenfelter and R. Layard (editors), *Handbook of Labor Economics*, Chapter 12, volume 1, Amsterdam: North-Holland

Saks, R. and S. Shore (2003), Risk and career choice, paper presented at the AEA Meetings, San Diego, January 2004

Schweri, J., J. Hartog and S. Wolter (2008), Do students expect compensation for wage risk? , Working paper University of Amsterdam

Shaw, K.L. (1996), An empirical analysis of risk aversion and income growth, *Journal of Labor Economics* 14 (4), 626-653.

Skyt Nielsen, H. and A. Vissing-Jorgenson (2006), The impact of labour market risk on education choices: estimates and implied risk aversion, Working paper Aarhus/Northwestern.

Tsiang, S.C. (1972), The rationale for mean-standard deviation analysis, skewness preference and the demand for money, *American Economic Review*, 62 (3), 354-371.

Webbink, D. and J. Hartog (2004), Can students predict their starting salary? Yes!, *Economics of Education Review*, 23 (2), 103-113

Weiss, Y. (1972), The risk element in occupational and educational choices, *Journal of Political Economy*, 80 (6), 1203-1213

Wolter, S. and B. Weber (2004), Returns to education: are students' expectations rational? Working Paper Swiss Coordination Center for Research in Education, Aarau, Switzerland

Appendix 1. Risk compensation in Levhari and Weiss (1974)

Equation (5) in Levhari and Weiss (1974), in their notation, reads

$$E \left\{ \frac{\partial U}{\partial c_1} [f_\lambda - (1+r)y_0] \right\} = 0$$

or

$$\frac{E \frac{\partial U}{\partial c_1} f_\lambda}{y_0 E \frac{\partial U}{\partial c_1}} = 1+r$$

Subtracting $E(f_\lambda)/y_0$ on both sides, and using $E(xy) = E(x)E(y) + \text{cov}(x,y)$, we can write

$$\frac{E f_\lambda}{y_0} = 1+r - \frac{\text{cov} \left(\frac{\partial U}{\partial c_1}, f_\lambda \right)}{y_0 E \frac{\partial U}{\partial c_1}}$$

Kodde (1985, Chapter 7) shows, using a second-order Taylor expansion around $E(\mu)$ that we may write

$$\text{cov} \left(\frac{\partial U}{\partial c_1}, f_\lambda \right) = \frac{\partial^2 U}{\partial c_1^2} f_\mu f_{\lambda\mu} \sigma^2 - \left| \frac{\partial^3 U}{\partial c_2^3} f_\mu^2 f_{\lambda\mu\mu} + \frac{\partial^2 U}{\partial c_2^2} f_{\mu\mu} f_{\lambda\mu\mu} \right| \sigma^4 / 4$$

where σ^2 is the variance of μ . Assuming $f_{\lambda\mu\mu} = 0$, as Levhari and Weiss (1974) implicitly do, we can write

$$\begin{aligned} \frac{E f_\lambda}{y_0} &= 1+r - y_0 \frac{\partial^2 U / \partial c_1^2}{E \partial U / \partial c_1} f_\mu f_{\lambda\mu} \frac{\sigma^2}{y_0^2} \\ &= 1+r + \rho f_\mu f_{\lambda\mu} \frac{\sigma^2}{y_0^2} \end{aligned}$$

where ρ is relative risk aversion evaluated at expected values. Thus, with risk aversion positive, and $f_{\lambda\mu}$ positive as assumed by Levhari and Weiss, under uncertainty the

expected return on human capital surpasses the Mincer rate by a term that is proportional to risk aversion and relative variance.

Appendix 2. Mean and dispersion in the tournament model (Lazear and Rosen, 1981)

Under risk aversion, equilibrium condition (24) in Lazear and Rosen (1981) can be written as

$$\mu = \frac{1}{C'(\mu)} \frac{V}{1+a\sigma^2} \text{ with } a = sC''\Pi > 0$$

Totally differentiating and rewriting yields

$$\frac{d\mu}{d\sigma} = - \frac{2a\sigma\mu/(1+a\sigma^2)}{1+\mu C''(\mu)/C'(\mu)} < 0$$

This holds because $C' > 0$ and $C'' > 0$ by assumption. One might also write the equilibrium condition as

$$\mu C'(\mu) = \frac{V}{1+a\sigma^2}$$

With the LHS increasing in μ , an increase in σ^2 has to be matched by a decrease in μ .

The equilibrium wage dispersion, according to footnote (9) in Lazear and Rosen (1981) obeys

$$C'(\mu) = wb\sigma^{-1} \text{ with } w = w_1 - w_2 \text{ and } b = 1/\sqrt{2\Pi}$$

Totally differentiating this condition (to μ , w and σ) yields,

$$\begin{aligned} \frac{dw}{d\sigma} &= \frac{\sigma}{b} \left\{ C''(\mu) \frac{d\mu}{d\sigma} + wb\sigma^{-2} \right\} \\ &= \frac{\mu C''(\mu)}{b} \left\{ \mathcal{E}_\sigma^\mu + \frac{1}{\mu C''(\mu)/C'(\mu)} \right\} \end{aligned}$$

where \mathcal{E}_σ^μ is the elasticity of μ to σ . Thus

$$\frac{dw}{d\sigma} > 0 \text{ if } \mathcal{E}_\sigma^\mu > - \left[\frac{C''(\mu)}{C'(\mu)} \mu \right]^{-1}$$

i.e. if the elasticity of effort is between zero and the inverse of what might be called relative effort aversion. The elasticity is endogenous, characterizing the dislocation of equilibrium effort. If we substitute from equation (1) above, we can derive

$$\frac{dw}{d\sigma} = \frac{C'(\mu)}{b} \left\{ 1 - \frac{2a\sigma^2 \mu C'(\mu)}{1 + \mu C''(\mu)/C'(\mu) V} \frac{1}{V} \right\}$$

This is ambiguous in sign, depending on exogenous parameter values.