A joint initiative of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute for Economic Research

Area Conference on

Economics of Education

03 - 04 September 2010 CESifo Conference Centre, Munich

Starting Wages Respond to Employer's Risk

Peter Berkhout, Joop Hartog and Hans van Ophem

CESifo GmbH
Poschingerstr. 5
81679 Munich
Germany

Phone: +49 (0) 89 9224-1410
Fax: +49 (0) 89 9224-1409
E-mail: office@cesifo.de
Web: www.cesifo.de

1.

STARTING WAGES RESPOND TO EMPLOYER'S RISK

peter berkhout

EIB Amsterdam

joop hartog

hans van ophem

Universiteit van Amsterdam

JEL code: J31
Keywords: wages; risk compensation; ability; incomplete information

Abstract
Firms hiring fresh graduates face uncertainty on the future productivity of workers. Theory suggests that starting wages reflect this, with lower pay for greater uncertainty. We use the dispersion of exam grades within a field of education as an indicator of the unobserved heterogeneity that employers face. We find solid evidence that starting wages are lower if the variance of exam grades is higher and higher if the skew is higher: employers shift the cost of productivity risk to new hires, but pay for the opportunity to catch a really good worker. Estimating the extent of risk cost sharing between firm and worker shows that shifting to workers is larger in the market sector than in the public sector and diminishes with experience.

Hartog acknowledges financial support from the Spanish Ministry of Science and Innovation through grant SEJ2007-66318. Comments by George Baker, John Jones and Wim Vijverberg are gratefully acknowledged, as well as comments received at seminars at CUNY, Columbia University Teachers College and SUNY Albany. We are grateful to Dinand Webbink for data assistance.

First draft 22032006; this draft 07122009; substantial revision of IZA DP 3026.
File: Starting wage employer riskIZARevisedTI060809

1. Introduction

An employer hiring a new employee fresh from school has no more than imperfect information on the worker's qualities. The diploma itself, some information on school grades, extracurricular activities, a job interview and perhaps a psychological test cannot fully resolve the uncertainty about future productivity. Firms may be expected to bill the workers for the cost of dealing with this uncertainty. Workers fresh from school have no successes yet to support a bargaining position and will have to accept that employers put a discount on starting wages in accordance with the risk they face. Thus, we predict that starting wages will be lower in fields where employers face more uncertainty on any individual's productivity. However, we also predict that starting wages will be higher if employers perceive more skew in the productivity distribution: they appreciate the chance to catch an individual with very high productivity. We find clear support for these predictions.

We use the distribution of exam grades within a field[1] to measure uncertainty. Exam grades differ among individuals in a given field because of differences in abilities, efforts and no doubt other factors. We assume that the heterogeneity that is reflected in the variance of exam grades is correlated with the heterogeneity that is relevant for employers. If the variance of exam grades across graduates in economics is larger than across graduates in physics, we assume that employers can make less accurate predictions on the productivity of an individual economist than on the productivity of an individual physicist. We do not require that employers consult grade variance data. Employers have their own ways of learning about the productivity risks among graduates in different fields. We only assume that these risks are adequately proxied by school grade variance (and skew). Our core hypothesis is that wages will respond negatively to the variance of exam grades in a field (workers pay a risk premium) and positively to the skew of the exam grades in a field (firms appreciate the upside risk of hitting upon a very good worker). This hypothesis is supported by a large sample of starting salaries for graduates from tertiary education in The Netherlands. We scrutinise our interpretation by adding robustness checks and by considering alternative explanations. Our stance is that we do not see a consistent alternative explanation for our findings.

We present formal modeling to derive our core prediction that wages of starting graduates will reflect the productivity risk that employers face. We will also specify conditions for productivity risk to be indicated by the distribution of exam grades. Worker payment for productivity risk has been highlighted in some earlier papers. Freeman (1977) introduced the idea that risk neutral firms are willing to insure starting workers against wage drops as information develops on their productivity. Harris and Holmstrom (1982) further developed this model to a market with risk neutral firms, risk averse workers and symmetric incomplete information on individual worker productivity. They show that wages are reduced by an insurance premium that diminishes with work experience (as information accumulates) and that is increasing in the (perceived) variance

[1] In the Netherlands, studies are not organised by major: first-year students right away specialise in an academic discipline (economics, physics, etc).

of productivity. Also, the variance of the wage increases with experience as wages come to reflect individual productivity. Harris and Holmstrom demonstrate that their model is in line with several stylized facts on wages, but offer no new direct testing[2].

Rothschild and Stiglitz (1982) and Aigner and Cain (1977) also predict that wages will respond negatively to higher perceived variance in unobserved ability. The former argue on the basis of mismatch between imperfectly observable individual ability and the optimal ability level for a given job, the latter on the basis of noise in an imperfect indicator of productivity. Neither of these papers offers any empirical evidence. To the best of our knowledge, there is no empirical work comparable to ours (cf Waldman, 2007). Our paper is the first that directly tests the prediction that wages are lower if the variance in unobserved productivity is higher with an indicator of productivity variance.

Allowing for skew is not routinely considered in labour market applications, but is well known in the lifecycle consumption-savings literature (as “prudence”). We include skew because we want to mirror our treatment of worker risk in educational choice (Hartog, 2009). That approach focuses on compensation to workers for the financial risk of an education (an education gives access to a distribution of wages, not to a single wage rate), and provides a good backdrop for the present analysis: in a sense it is the mirror image of the case we study here. Complementarity of these two cases means that the results of the tests reinforce each other.

We will proceed by presenting formal modeling in the next section. Data and basic results follow in sections 3 and 4. Section 5 considers alternative explanations and robustness. Section 6 presents estimates for a model of risk cost sharing between employers and employees. Section 7 concludes.

2. Formal arguments

In section 2.1, we will formally derive how employers adjust starting wages for the risk they face. In section 2.2, we present a set of conditions for the distribution of exam grades to be informative on productivity risk.

2.1 A model for risk shifting

It is commonly assumed that workers are risk averse and firms are risk neutral. This is probably pushing the case too far. It is quite likely that on average workers are more risk averse than firms, but no doubt firms are also risk averse. Small firms may have every reason to behave as risk averters, as they often lack the resources to survive bad draws. But large firms are also observed to engage in buying all kinds of insurances, for failing debtors, worker safety hazards, currency fluctuations, etc[3]. There is sufficient evidence to

[2] Waldman (2007) gives a good survey of the literature on wages under imperfect information.
[3] To witness: Dutch electronics multinational Philips sells its chips division because sales and profits vary too much over the business cycle (NRC, August 4, 2006).

assume that firms are willing to pay to avoid a fair gamble. We will thus assume that firms are risk averse.

Assume that individuals differ in potential productivity. Productivity refers to quality and quantity of the firm's output. Firms can hire two types of workers, experienced and inexperienced. Productivity of experienced workers has been revealed with experience and is fully known to firms and there is a well-defined market wage function to reward this productivity. This is obviously a simplification, as in practice information develops gradually. In their decision what type of employees to hire, firms maximize the utility of profits by maximizing it across productivity q_{ij}. Productivity both depends on the individual (i) and the firm (j). Thus, for firms we have

$$\max U(\Pi_{ij}) = U(pq_{ij} - w(q_{ij})) \tag{1}$$

where: Π_{ij} = the profit earned on individual i if employed by firm j, maximized across the productivity q_{ij};
p = price of output;
$w(.)$ = the market wage depending on the productivity of individual i if employed by firm j.

Solving this problem determines the optimal productivity q_{ij}^{o} for firm j. We ignore that a bargaining situation may arise between the firm and the worker: we simply assume that the market dictates the wage for a worker of known productivity.

If firms hire inexperienced workers, i.e. fresh graduates, they do not know individuals' true productivity. Instead, the firm has some perception $\tilde{q}_{ij}$, the perceived productivity of individual i if employed by firm j. The information is based on e.g. impressions collected during the application process, the evaluation of labour market activity while in school, extracurricular activities, etc. The firm deliberately searches for workers with optimal productivity as defined above; we assume that the firm's search and selection procedure leads to candidates with the proper expected productivity, but that perceived productivity deviates randomly from the optimum, i.e.

$$E(\tilde{q}_{ij}) = q_{ij}^{o} \tag{2}$$

In its decision to hire an individual or not the firm considers:

$$U(p\tilde{q}_{ij} - \tilde{w}(\tilde{q}_{ij})) \tag{3}$$

This is the utility of profit on a worker with perceived productivity $\tilde{q}_{ij}$ who is offered a wage commensurate with the uncertainty facing the firm. As the firm can hire experienced workers with known productivity, it will only hire inexperienced workers with uncertain productivity if the pay-offs are identical:

$$E[U(p\tilde{q}_{ij} - \tilde{w}(\tilde{q}_{ij}))] = U(pq_{ij}^{o} - w(q_{ij}^{o})) \quad . \tag{4}$$

Firms will use the wage function $\tilde{w}(\tilde{q}_{ij})$ to establish this equality between expected utility from hiring a new graduate with uncertain productivity or hiring an experienced worker with known productivity.

As shown in the Appendix, we can use Taylor series expansions to rewrite equilibrium condition (4) as an equation for the expected wage of an inexperienced worker:

$$E(\tilde{w}(\tilde{q}_{ij})) = E(w(\tilde{q}_{ij})) - \frac{1}{2}\frac{\partial^2 U / \partial \tilde{q}^2\big|_{q_{ij}}}{E\left(\partial U / \partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)}\sigma^2 - \frac{1}{6}\frac{\partial^3 U / \partial \tilde{q}^3\big|_{q_{ij}}}{E\left(\partial U / \partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)}\kappa^3$$
$$- \frac{\operatorname{cov}\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}, \tilde{w}(\tilde{q}_{ij})\right)}{E\left(\partial U / \partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)} \tag{5}$$

Equation (5) indicates that the expected wage offered by the firm for an inexperienced worker deviates from what the market offers for expected productivity if this productivity were known by subtracting a premium to compensate for risk and skew of unknown productivity. With second derivative of utility to productivity negative and third derivative positive, a higher variance reduces the offered wage and a higher skew increases it, as the derivative to the wage is negative[4].

Making the assumptions: $\tilde{w}_{ij} = \tilde{w}(\tilde{q}_{ij}) = E(\tilde{w}(\tilde{q}_{ij})) + \varepsilon_{ij}$, $E(\varepsilon_{ij}) = 0$, and $E(w(\tilde{q}_{ij})) = \beta' x_{ij}$, that the (fractions of the) derivatives with respect to utility can be approximated by constants, and that the last term can be considered purely random[5] (ξ_{ij}), we arrive at an estimable model:

$$\tilde{w}_{ij} = \beta' x_{ij} + \alpha_1 \sigma^2 + \alpha_2 \kappa^3 + (\varepsilon_{ij} + \xi_{ij}) \tag{6}$$

---

[4] For the second derivative to be negative we need assume $\frac{\partial^2 \Pi}{\partial q^2} < 0$, which seems uncontroversial. The third derivative is positive if $\frac{\partial^3 U}{\partial \Pi^3}$ is positive, a necessary condition for decreasing absolute risk aversion.

[5] Not necessarily with zero mean. A nonzero mean will be part of the constant in $\beta' x_{ij}$

Equation (6) is essentially the equation we have estimated. It is based on the assumption that the effects of uncertainty about productivity are fully shifted to the worker. However, we may generalise this and assume that the incidence of risk is not identical in all labour markets. To allow for varying risk sharing, we introduce a factor $0 \leq \theta_j \leq 1$ and rewrite the equilibrium condition as:

$$E[U(p\tilde{q}_{ij} - \tilde{w}(\tilde{q}_{ij}))] = U(pq_{ij}^o - w(q_{ij}^o))$$
$$\approx E[U(p\tilde{q}_{ij} - w(\tilde{q}_{ij}))] - \theta_j \left(\frac{1}{2} \frac{\partial^2 U}{\partial \tilde{q}_{ij}^2} \Big|_{q_{ij}^o} \sigma^2 + \frac{1}{6} \frac{\partial^3 U}{\partial \tilde{q}_{ij}^3} \Big|_{q_{ij}^o} \kappa^3 \right) \quad (7)$$

If $\theta_j = 1$, the employer is fully compensated by a lower wage, if $\theta_j = 0$, the employer is not compensated at all. Going through the same steps as before, we end up with the following result:

$$\tilde{w}_{ij} = \beta' x_{ij} + \alpha_1 \theta_j \sigma^2 + \alpha_2 \theta_j \kappa^3 + \varepsilon_{ij} + \xi_{ij} \quad (8)$$

θ_j will depend on the distribution of market power in a given market. High unemployment or a higher level of concentration in the industry the firm operates in might cause θ_j to be closer to one. To model this, assume:

$$\theta_j = \frac{\exp(\gamma' z_j)}{1 + \exp(\gamma' z_j)} \quad (9)$$

where z_j is a vector of variables representing the market power of firm j. With this specification we guarantee that θ_j is between 0 and 1.

Our model and earlier models mentioned above have in common that wages are predicted to fall with increasing (perceived) productivity risk of workers. Rothschild and Stiglitz (1982) essentially derive their conclusion from assumptions on the production function, while the other models focus on risk and insurance motives. Freeman (1977) and Harris-Holmstrom (1982) assume risk neutral firms selling insurance to risk neutral workers, while our model and Aigner and Cain (1977) assume risk averse firms that shift the cost of risk to workers. None of the earlier models considers the effects of skew in the productivity distribution. We consider asymmetry as relevant in itself, as firms will have a different interest in upside and downside risk. Also, we think there is added value in the common treatment of employer and employee risk as developed in the Risk Augmented Mincer equation (Hartog, 2009).

2.2 Measuring productivity risk

From their own experience and possibly other sources, employers have collected information on expected productivity of workers and on variance and skew of productivity within given fields of education. We, as outside observers, have no access to this information. But we will assume that the distribution of school performances provides some relevant information about the distribution of future productivities. School performances depend on abilities, motivation, effort, interest and possibly many other variables. As worker productivity depends at least in part on the same variables, this commonality can be exploited to use the distribution of school performance as a proxy for the variability of productivity.

Let's assume linear relations between ability and productivity:

$$q_{si} = \beta_s a_{si} + \upsilon_{si} \tag{10}$$

where a is ability and q is productivity; both ability and productivity are specific for education s. The stochastic term υ indicates that other factors than ability also affect productivity. Then, distribution moments follow as[6]

$$E(q_{si}) = \beta_s E(a_{si}) \tag{11}$$

$$\sigma^2_{qs} = \beta^2_s \sigma^s_{as} + \sigma^2_{\upsilon s} \tag{12}$$

Now, assume employers know β_s, σ^2_{as} and $\sigma^2_{\upsilon s}$: they cannot observe an individual's ability, but they know their production function and the moments of the distributions. In the labour market, employers shift productivity risk to workers. We want to estimate this as the regression coefficient of wages on risk as given by (12).

As outside observers, we have no data on productivity risk. Instead, we use school grade variance as a proxy. We assume that school grades are produced by the same school type specific ability and unidentified other factors:

$$g_{si} = \alpha_s a_{si} + \varepsilon_{si} \tag{13}$$

and hence,

$$\sigma^2_{gs} = \alpha^2_s \sigma^s_{as} + \sigma^2_{\varepsilon s} \tag{14}$$

For school grade variance to be a good proxy for employers' risk productivity risk, we need

[6] We do not spell out the condition for the third moment, as this is similar.

$$\alpha_s^2 \sigma_{as}^s + \sigma_{\varepsilon s}^2 = \quad \beta_s^2 \sigma_{as}^s + \sigma_{\upsilon s}^2 \tag{15}$$

or

$$\sigma_{as}^2 = \frac{\beta_s^2}{\alpha_s^2} \sigma_{as}^2 + \frac{\sigma_{\upsilon s}^2 - \sigma_{\varepsilon s}^s}{\alpha_s^2} \tag{16}$$

The quality of the proxy might vary quite irregularly across educations s. But we can say that grade variance is a good proxy for employers' productivity risk if, for each education s, the noise variances and the slopes α and β are always close together. This means that across educations, β / α is close to a constant (the effects of ability on grades and on productivity move together) and so are the stochastic disturbances. In educations where ability has a strong effect on productivity it should also have strong effect on grades[7], in educations where the productivity relation is noisy, the grade relation should also be noisy. We cannot assess the empirical status of this condition. We can only say that it is quite conceivable that it holds and maybe it is even plausible.

If the gap between the noise variances is close to a constant across educations, we have a systematic over- or underestimation of risk, and an associated under- or overestimate of the regression coefficient. If the noise in the grade function is larger than the noise in the production function, we overestimate risk and underestimate the reaction coefficient. This holds a fortiori if employers do not seek to shift noise in the production function as this reflects external factors that the employer absorbs as typical risk born by the entrepreneur.

3 Data

We use data from the Elsevier/SEO survey, held among graduates from tertiary education. A new cohort of graduates has been interviewed every year since 1996, with focus on outcomes in the first 20 months in the labour market. Dutch tertiary education is basically divided into two levels: higher vocational education (in Dutch abbreviated as HBO) and university education (WO). HBO-education prepares students for specific (categories of) professions. It is taught at about 60 special institutes evenly spread over the Netherlands. On average, 50,000 students graduate each year from HBO. WO-education is considered to be of a somewhat higher intellectual level and has a more general academic character; it also requires a higher level of secondary education for direct admission. It is taught at 14 universities. Approximately 23,000 students graduate every year. At HBO-level students can choose between 250 different courses of study, while at WO-level they may choose

[7] Note that possibly different scaling between grades and productivity can be accommodated by measuring grades at the same scale as productivity. It's only the variation of the slopes across educations that counts.

between 260 different specializations. Most of them, however, produce only small numbers of graduates, making statistical analysis based on variances in grades within specialisations unreliable. About 80 percent of the student population is concentrated in the 113 largest degree fields. The survey is restricted to these 113 degree fields (studies) which divide evenly over HBO and WO. This means the survey is representative of 80 percent of the yearly outflow of graduates at HBO- and WO-level. Every year a sample of on average 7,500 observations is drawn. The special feature of the survey is the large number of studies within tertiary education and the focus on starting salaries

We pool 9 cohorts, from 1999 until 2007, with a time dummy to distinguish them. Earnings are defined as log net hourly wages at the time of the survey, i.e. on average 20 months after graduation (reported earnings are divided by reported hours); salaries are self-reported and may contain the associated noise. For our empirical purposes, we excluded all respondents who are self-employed, part time employed and all those for whom data on control variables are unavailable. To eliminate outliers, we discarded both the highest and the lowest 1% of the sample. All correlations between explanatory variables are low, and we need not worry about multicolinearity. The data, based on about 45 000 observations from 113 schooling types, are characterised in the Appendix.

In the Elsevier/SEO data individuals were asked for their average exam grade, across all courses, in tertiary education (grading uses a standard 0-10 scale; passing requires a minimum of 5.5, though not necessarily for all courses, as compensation is sometimes allowed). As explained above, we take the dispersion (and skew) of exam grades, for all students with a given type of tertiary education, as an indication of dispersion (and skew) of the productivity distribution.

The dataset allows us to use many variables, i.e. to control for the situation in the labour market (region, unemployment/vacancy ratio, time in the labour market since graduation), personal characteristics (age, gender, parents' education, individual grades) and job characteristics (job level, industry, type of contract). These variables should be sufficient to predict expected productivity. We also control for the workers' risk when selecting an education, by including variance and skew of the earnings residual in the chosen education. The argument here is that potential students will only select an education if they are compensated for the earnings risk of that education. Formal modelling, a survey of empirical results and assessment of pitfalls and critique are given in Hartog (2009), examples in Hartog and Vijverberg (2007) and Diaz Serrano, Hartog and Skyt Nielsen (2008). Worker risk is measured as the variance of the residuals from a Mincer earnings function within that education[8]. Workers demand and indeed get compensation for risk and are willing to take a pay reduction for a better chance of very high incomes and hence accept lower wages for higher skew. We include this compensation for worker risk because compensation for employer risk and employee risk neatly mirror each other, as theory predicts. We take the symmetry of the results on both sides of the market as additional support for the analysis presented here.

[8] In Hartog (2009) we extensively discuss the selectivity problem and the potential ability bias when using residual variance as a measure of risk.

Mean wages are plotted against variance and skewness of exam grades in Figure 1. The plots indicate weak correlations (negative and positive, respectively)[9]. We have extensively looked at the data, to see if there are any obvious patterns among types of education (such as typically high or low variances for all humanities or for all sciences) that might be reason for suspicion. We did not see any. Some of the studies that are intellectually more demanding (like biomedical science and pure mathematics) have high variance, but among the high variance fields we also find less demanding fields, like sociology and languages. A field with restrictive entry like medicine has grade variance in the higher end of the distribution but a similar field like dentistry has low variance.

4 Basic results

Tables 1, 2 and 3 give the basic results. As announced, we include the results on worker risk, as they give a strong background to the compensation for employer risk. Worker risk is measured by the residual earnings variance within an education type from a Mincer equation across all educations (the regression contains a dummy for field of education and cohort dummies). As usual, we find a positive effect for the variance (Erisk) as workers do not like risk, a negative effect for skew (Eskew), as workers like positive asymmetry (with some extra probability of a very high outcome). Estimation results for Erisk and Eskew are robust in magnitude and statistical significance (except for some variation across industries) and will not be commented on in detail. With education completed, workers have no alternative and employers can shift the risk from the heterogeneity they cannot observe to workers. The results we report here on employer's risk are independent of employee's risk: regression coefficients vary only marginally and significance levels are unaffected when employee risk is excluded or included[10].

We start, in Table 1, with an OLS estimation at the level of studies, with the mean earnings in a field of study as the dependent variable. Our basic prediction is strongly supported: in studies with higher variance in grades, starting salaries are lower, with higher skew in grades they are higher. The results are not sensitive to including or excluding compensation for worker risk.

In Table 2 we present estimates with individual earnings as the dependent variable and standard errors adjusted for clustering, as there may be correlation for errors within fields of education, the so-called Moulton problem. We allow for clustering by education type, as we measure our independent variables of prime interest by type of education[11].

Again, the basic prediction is supported, with similar coefficients, but sometimes lower precision. The advantage of the estimation on micro data is the possibility to include

[9] The remarkable outlier is dentistry (as employee: we exclude the self-employed); we have no explanation.

[10] Full analysis of worker risk compensation in the present dataset, with formal modelling and references is given in Berkhout, Hartog and Webbink (2010).

[11] We experimented with different types of education clustering, based on our a priori notions on related shocks through sensitivity to the same or related product or labour markets (e.g clustering all language educations). Reducing the number of clusters by assumed market similarities proved immaterial for the results.

additional controls for other variables that influence earnings. As Table 2 shows, in panel A, the coefficients are sensitive to controls for personal and job characteristics but not for time and region. Precision also falls and significance levels become low. This suggests that some of the variation in earnings initially ascribed to risk as reflected in dispersion in school grades is due to heterogeneity among individuals and their jobs. Note that risk is not recalculated when we add controls and this must imply that school grade heterogeneity is correlated with the controls we use.

In Table 3, we present estimates for a random effects model. Random effects at the level of the studies control for unobserved variables that explain differences in earnings and that may bias the estimates of our risk coefficients. Note that we cannot estimate a fixed effects model because variance and skew are fixed for a given type of education. As the first column of panel A shows, the coefficients for grade variance and skew do not change much, but statistical significance increases[12]. Adding controls now has far less effect than in Table 2. The reduction in magnitude of the coefficients is smaller and they remain significant at conventional level. Thus, controlling for unobserved heterogeneity between the studies increases support for our hypothesis. The small effect of introducing random effects for education types on estimated coefficients suggests that our measures of employer risk are not picking up effects of omitted variables that determine wage differentials between types of education.

Finally, we note that the effect of individual school grades, while significant in the OLS estimates and in the most extensive RE specifications, is quite small. School performance is graded on 0-10 scale, with 6 or more required for a pass. In the OLS model, a one unit grade point difference in average exam score increases starting wages by about 1 percent. Overall, the standard deviation of individual grades is about 0.5. Thus, a one standard deviation increase in individual grades raises starting wage by 0.5 percent. In the RE model the effects are even smaller.

5. Alternative explanations and robustness checks

Grading practices may differ markedly between types of education. In particular, so called grade inflation (the tendency to give all students high grades) would truncate the grade distribution, reducing variance and skew because the grade distribution hits the upper ceiling of the grade scale. To assess this argument, we should first of all observe the differences in grading practices between education types as rooted in differences in teacher leniency and tolerance and the like. The pattern should then fit our observations. Thus, truncation (upward drift of the grade distribution) should be stronger in types of education with high wages, as high wage educations should have low variance (and high skew). Considering the heterogeneity among the high wage educations, this does not strike us as an obvious pattern. For example, university education has an average wage of 2.318. Educations with wages above 2.34 are mechanical engineering, technology and management, business science, econometrics, fiscal law, medicine, dentistry, veterinary

[12] If we calculate the standard errors in the random effects model without clustering, t-values are not below 2.4.

science and public administration (see Data Appendix). We cannot imagine good reasons to assume that these fields are connected by high grading practices. Neither do we suspect an effect of grading practices in particular schools: these fields are taught in many institutions.

Another potential explanation is based on differences in intellectual requirements between fields of study: some studies can only be completed for students with high analytical ability. This will generate self-selection at entry[13] and selective drop-out based on those requirements. In the end, the population of graduates may be rather homogenous. As these high ability graduates may also be expected to obtain high earnings, we would see a negative correlation between earnings and grade variance. With truncation at the low end of the ability distribution, a positive correlation between earnings and grade skew would also result. The problem is of course to measure the differences in intellectual requirements (or "difficulty") of studies. We considered applying a distinction based on our own perceptions but discarded this as too subjective. Instead we selected a rather homogeneous "average ability" group of students and used the difference between average grade in secondary school and in the tertiary fields of education as an indicator. We selected students in the middle of the grade distribution at the final exam of secondary school: only students close to the overall mean exam grade are retained (we used the middle third of the distribution, symmetrically about the median). In addition, we required the variance in the individual's exam grades across subjects to be small (we used the middle third of the distribution). Thus, we have a fairly homogeneous group of students, about 9% of the sample. We then calculated the differences between their mean exam grades and the mean exam grade in their tertiary study. Based on these differences we split studies between "difficult" and "easy", as two roughly equally large groups. The difficult studies have the larger average gap between tertiary grade and secondary grade. Although the resulting distinction does not always match our own perceptions, we used this distinction to add a dummy for difficult studies to our estimation equation. As Tables 2 and 3 (panels B) show, the coefficient on this dummy is not statistically significant and inclusion has no effect on any of the other coefficients. In fact, observing the variance of grades by type of educations also provides evidence against this explanation. As we noted at the end of section 3, intellectually demanding studies do not stand out with low grade variance: among vocational graduates, variance is high for medical imaging and radiotherapy, applied informatics and applied physics, among university graduates it is high for chemistry, computer science, biology, pure mathematics, electrical engineering, econometrics and biomedical science.

An explanation might also be found in the organisation of the studies, in combination with labour market protection. Some studies are rather strictly organized, with attendance requirements (e.g. in laboratories), regular assignments and active student monitoring. This may increase the homogeneity of the population of graduates, by selective drop-out and elimination of differences in effort and other study habits, with low grade variance as a result (the effect on grade skew is less clear). If these studies happen to be the studies leading to high-paid jobs, there would be a negative correlation between earnings and

---

[13] Dutch universities do not select at entry, anyone with the proper secondary school diploma must be admitted. In higher vocational education, schools may restrict entry.

grade variance. Medicine would be an example. In addition, in some occupational fields government regulations lead to a monopolistic market structure, with high protected earnings. If such a monopoly coincides with strict organization and student monitoring, again as in medicine, this would also generate the negative correlation between earnings level and grade variance. To check this argument, we added a dummy for 7 studies preparing for a job in a highly regulated market: physiotherapy, medical techniques and radiotherapy, dental hygiene, pharmacy, notary, medicine, dentistry. All the medical studies are strictly organised and regulated, notary is regulated but the study is not strictly organised. The dummy has the expected positive effect on earnings, but inclusion has no effect on any other coefficient (Tables 2 and 3, second panel).

Could the results be explained from investments in human capital? Human capital theory predicts correlation between intercept and slope of experience-earnings profiles: in fields where workers invest more in on-the-job-training, intercepts will be lower and earnings growth rates will be higher. This leads to an alternative explanation of our results if two conditions are met. First, intercepts and growth rates in different fields of study should correlate. Second, the low intercept/high growth rate fields should coincide with the fields that have high variance in exam grades (high risk). It turns out that the first condition is met, but the second is not. We use a nationally representative cross-section sample of earnings data by field of study (LSO 2002, collected by CBS, the Dutch national statistical office; the data are described in detail in Berkhout, Hartog and Webbink, 2010). We selected fields of study on which we have observations both in our own data set and in LSO (containing least 10 observations): 26 fields for males at higher vocational level and 20 at university level, 20 fields for females at higher vocational level and 17 at university level. From the LSO data we estimate an OLS earnings equation where log wage is explained from an intercept, age and age squared, separately for each field, for males and for females. From the regression we predict average annual earnings growth for the first 10 years of experience. Figure 2 shows the results for males with university education. The human capital prediction of correlation between intercept and slope of the earnings profile is strongly supported (Figure 2A). But low-intercepts or high growth rates have no relationship with variance or skew of exam grades. Regression analyses corroborate these findings; when variance or skew are related to intercepts or growth rates, t-values are mostly well below 1 and never above 1.5. Results for the other three groups (not shown here) are similar. This rules out that our results should in fact be explained by patterns of investments in on-the-job training or any other theory that focuses on the relationship between intercept and slope of the earnings function, such as incentive theories.

There is often some presumption that more able individuals select innately more risky activities or that more "difficult" activities also inherently have more risky outcome. This, of course would work the wrong way, as one would expect that abler individuals through more difficult studies, would earn higher wages. But then higher grade variances would be associated with higher wages, not lower. A similar argument holds that in some fields, productivity is easier to predict than in other fields. But this argument is too imprecise to get us anywhere. Does it mean that the variance of productivity is inherently smaller for, say, philosophers than for mechanical engineers? How do we know? Or does it mean that better predictors are available of the extent of variation? This may well be picked up by our grade variances. Reference to measurement errors, no doubt substantial in our dataset, with all these self-reported data, is also too vague: we need a precise prediction on the pattern of measurement errors that can explain our results.

Employment protection is obviously a relevant issue. If an employee can be fired at will, the employer risk from unknown productivity will be lower than under a permanent contract. By itself the argument is not at variance with our story, as it simply specifies a cost component of risk: firing and hiring cost. To disrupt our interpretation, different regimes of employment protection should hold for different graduates. That does not seem very likely. Nowadays, most graduates will be offered an initial contract with limited duration (or a probation period) and firing before the contract expires is quite unusual.

We also considered, in Table 4, the effects within subpopulations, with variance and skew calculated for the relevant sub-populations. We only present RE estimates, as we believe that allowing for unobserved heterogeneity among educations is indeed called for (as above, OLS estimates have higher standard errors). Statistical precision falls within subgroups, in particular for grade skew. Some wrong signs occur, but we never find significant violation of the hypothesis in the form of a wrong sign that is statistically significant. Separate estimation for men and women does not affect the results. If we allow the labour market to be segmented by "ability", we find robust results for grade variance but insignificance for grade skew. We have created ability quartiles on the basis of average grade in the secondary school final exam (quartile 1 is the lowest grade segment). School grades are influenced by ability and drive and both are relevant to employers. Our results suggest that the effect of risk is not driven by a spurious effect of ability/ambition as reflected in school grades. Distinction by time between graduation and time of survey (work experience longer or shorter than the mean) does not affect the results. Distinction between difficult and easy studies again produces no violation of the basic predictions[14]. Separate estimation for six employment sectors produces no significant violations; results are significantly confirmed for three sectors (Government, Services and Manufacturing), but insignificant for Education and Care. Education and Care are industries with excess demand, so this might reflect a disequilibrium result. However, separating the vocational educations in Education and Care from the other industries did not improve results.

If we distinguish between university and higher vocational education, the results within these subgroups are no longer statistically significant. An explanation might be that a university education offers a more general education than higher vocational education, reducing the risk in hiring a vocational graduate. The effect may be reinforced by the entry selection that many vocational institutions apply, as opposed to universities: their graduates might simply be more homogeneous groups.

[14] We also checked the effect of deleting any cohort t, for t=1 to 11 from the sample; this proved immaterial for the estimation results.

We can express the magnitude of the effects in elasticities, by multiplying the coefficients by the mean of the independent variable. Table 5 collects results. It's immediately clear that all elasticities are small and not very sensitive to specification. The wage elasticity for grade variance is between 0.05 and 0.10, the wage elasticity for grade skew is between 0.01 and 0.02. The variance of these elasticities within subgroups is modest. Mean values of grade variance and grade skew vary remarkably little between the subgroups, and hence the estimated regression coefficients give a good indication of the variation. The wage differential between the education with maximum and minimum grade variance is 10 % (the variances are 0.455 and 0.137), between education with maximum and minimum grade skew is 6.5% (the skew is .244 and -.013; we use the regression coefficients in column A4, Table 3). Thus, the effect on wage differentials is not negligible.

6. Risk sharing

As noted above, we can allow for risk cost sharing between worker and firm, rather than impose that the worker will bear the full burden. In equation (9), we have allowed for a sharing parameter θ that is made to depend on parameters that affect relative bargaining power. θ is a parameter between 0 and 1, where the value of 1 indicates that the employer is fully compensated for productivity risk with a lower wage and the value 0 indicates that the employer fully carries the cost of risk without any compensation in wages. In particular, we estimated the following model

$$Ln\, w_{ij} = \beta' x_{ij} + \theta_j \left(\delta_1 \sigma_j^2 + \delta_2 \kappa_j^3 \right) + \varepsilon_{ij}$$
$$\theta_i = \frac{\exp Z_i\, \alpha}{1 + \exp Z_i\, \alpha} \qquad (14)$$

Results are given in Table 6. We have experimented a little with assigning variables to the linear part of the wage equation and the non-linear part through the risk sharing parameter θ. Due to high levels of multicolinearity, we could not have all variables in both the linear and the non-linear part. Estimated coefficients on the linear part of the wage equation are not much different from our earlier estimates. Grade variance and grade skewness come out with the right signs, and have even higher significance levels than before. Base values are lower in absolute sense, but now are no longer constant. In the risk cost sharing part, a positive coefficient indicates that the variable contributes to shifting the cost of risk to the worker, by reducing the wage rate for a given risk. Thus, graduates from vocational education have larger wage reduction for given risk than university graduates. In the private sector, wages are more reduced for given employer risk than in the public sector: with the government as reference sector, and education and health care as mostly public sectors, the private sectors all have higher coefficients. Business services show the highest degree of shifting the cost of risk to workers, the education sector shifts least. Interestingly, as shown in Figure 3, the year dummies trace the profile of the national unemployment rate. If unemployment goes up, the year dummy goes up, indicating that with higher unemployment, the worker has to pay a higher share of the risk.

Our model in section 5.1 makes a distinction between starting workers and experienced workers, with extreme assumptions on information about their productivity: jumping from no individual information to full information. In practice, of course, information develops gradually, as e.g. in Harris and Holmstrom (1982). We can use these models to predict that gradually, as information on individual productivity develops, workers will have to pay less to compensate employers' risk: θ should decline with experience. This is indeed exactly what we find. In the linear model (as reported in Table 3), we have split the sample in two groups: work experience since graduation less than 18 months or greater than 18 months. For the total sample we estimated θ of 0.317 (standard error 0.076). For the short experience sample, with 19 634 observations, we find 0.307 (0.052), for the more experienced sample (25 340 observations) we find 0.243 (0.040). We have also estimated a continuous specification, with experience since graduation in months m. A parabolic specification in the θ logit yields $0.027m-0.0016m^2$; only the quadratic term is statistically significant[15]. The parabola peaks at 8.4 months and is negative after 16.9 months. Clearly, the price that workers pay for the uncertainty that employers face when hiring them falls over time.

7. Conclusion

From an analytical model, we derived the prediction that firms would pay lower starting wages if they face larger variance in individual productivity and higher wages if they have more favourable odds of hiring individuals with very high productivity. The model formalises the argument that firms shift the risk associated with uncertain productivity of labour market entrants to the worker, but are willing to pay for the probability of hitting upon exceptionally good workers. The prediction on the effect of the variance is similar to that in Harris and Holmstrom (1982), the prediction on skew is derived only in our own model. We use the distribution of exam grades to obtain information on the distribution of individual productivity for given education. An analytical model on the link between productivity risk and exam grade variance did not yield unambiguous results. Exam grade variance is only a monotonic indicator of productivity risk if certain parameter conditions hold.

Our empirical results are unambiguous. We find that starting wages are lower for fields of education where the variance of exam grades is higher and that starting wages are higher in fields where the skew of exam grades is higher. These results are robust within sub-populations and also survive some tests against alternative interpretations. The results are stronger if we estimate a random effects model than when we apply OLS. This strengthens our case, as the RE model controls for unobserved heterogeneity among educations. In a non-linear specification we have estimated the extent of risk cost sharing between employers and workers. We find that vocational graduates pay more for the productivity risk that employers face than university graduates and that employers in the

[15] With a simple linear specification, estimated coefficients are statistically not significant.

market sector shift a larger share to workers than public sector employers. As experience grows, workers pay less for employers' productivity risk. We have estimated standard errors with allowance for clustering of educations, but actually we find clustering a rather arbitrary procedure, as it is unclear a priori what clustering structure is called for. Dropping clustering would decrease our standard errors and increase statistical significance considerably.

In all specifications, we include worker's financial risk, as reflected in the distribution of residuals within an education. In all specifications, we find a positive effect of residual variance and a negative effect of residual skew. We also take this as (circumstantial) evidence for our approach as it indicates general support for the notion that risk has an effect on wages.

As we see no obvious alternative explanation for our findings, we take the empirical results as clear support for our interpretation: employers shift part of the cost of productivity risk to workers.

References

Aigner, D. and G. Cain (1977), Statistical theories of discrimination in labor markets, *Industrial and Labor Relations Review*, 30 (5), 175- 187

Altonji, J. and C. Pierret (2001), Employer learning and statistical discrimination, *Quarterly Journal of Economics*, 116 (1), 313-350

Berkhout, P., J. Hartog and D. Webbink (2010), Risk compensation under worker heterogeneity, *Southern Economic Journal*, 76 (3), 762-790

Diaz Serrano, L., J. Hartog and H. Skyt Nielsen (2008), Compensating Wage Differentials for Schooling Risk in Denmark, *Scandinavian Journal of Economics*, 110 (4), 711-731,

Farber, H. and R. Gibbons (1996), Learning and wage dynamics, *Quarterly Journal of Economics*, 111 (4), 1007-1047

Freeman, S. (1977), Wage Trends as Performance Displays Productive Potential: A Model and Application to Academic Early Retirement, *Bell Journal of Economics*, 8, 419-443.

Harris, M. and B. Holmstrom (1982), A theory of wage dynamics, *Review of Economic Studies*, 49 (3), 315-333

Hartog, J. (2009), A Risk Augmented Mincer Earnings equation? Taking stock, IZA Discussion Paper 4493, forthcoming in *Research in Labor Economics*, 2011

Hartog, J. and W. Vijverberg (2007), On compensation for risk aversion and skewness affection in wages *Labour Economics, Special Issue on Education and Risk***,** 14 (6), 938-956

NRC Handelsblad (2006), Philips verkoopt chips divisie, August 4, 2006

Rothschild, M and J. Stiglitz (1982), A model of employment outcomes illustrating the effect of the structure of information on the level and distribution of income, *Economics Letters*, 10, 231-236

Waldman, M. (2007), Theory and evidence in internal labor markets, MPRA Paper 5113, online at http://mpra.ub.uni-muenchen.de/5113/

Table 1. OLS on aggregate data (113 studies)

Ln hourly wage	**Coef.**	**t-val**
Intercept	2.225	59.04
Erisk	3.054	5.27
Eskew	-2.992	-3.04
Grade variance	-0.422	-2.61
Grade skew	0.442	2.58

N=113, R^2 = 0.249

Table 2. OLS on micro data

	Regressions							
Ln hourly wage	**A1**	**A2**	**A3**	**A4**	**B1**	**B2**	**B3**	**B4**
Erisk	3.512	3.428	2.888	2.801	3.288	3.214	2.774	2.757
(t-val)	(3.12)	(3.05)	(2.45)	(5.55)	(3.21)	(3.13)	(2.52)	(5.24)
Eskew	-3.454	-3.342	-2.721	-2.898	-3.021	-2.932	-2.469	-2.702
(t-val)	(-1.78)	(-1.72)	(-1.46)	(-4.10)	(-1.88)	(-1.82)	(-1.50)	(-3.80)
Grade variance	-0.504	-0.471	-0.349	-0.223	-0.534	-0.500	-0.399	-0.224
(t-val)	(-2.09)	(-1.98)	(-1.54)	(-2.36)	(-2.40)	(-2.30)	(1.96)	(-2.50)
Grade skew	0.495	0.439	0.231	0.118	0.508	0.453	0.268	0.102
(t-val)	(1.97)	(1.80)	(1.03)	(1.33)	(2.12)	(1.96)	(1.27)	(1.15)
Individual grades	0.017	0.015	0.019	0.010	0.012	0.011	0.015	0.009
(t-val)	(3.93)	(3.40)	(4.58)	(3.41)	(3.24)	(2.71)	(3.74)	(3.16)
Difficult					-0.014	-0.014	-0.019	-0.000
(t-val)					(-0.87)	(-0.89)	(-1.31)	(-0.01)
Regulated					0.094	0.093	0.079	0.031
(t-val)					(1.63)	(1.67)	(1.81)	(1.66)
R^2	0.033	0.061	0.081	0.203	0.046	0.073	0.090	0.204

N = 44974; t-values based in standard errors clustered by field of education.

A = Regression without dummies for difficult and regulated studies.
B = Regression with dummies for difficult and regulated studies.

1 = intercept only.
2 = as 1, plus labour market variables (year dummies, regions, time since graduation).
3 = as 2, plus personal characteristics (age, gender, parental education).
4 = as 3, plus job characteristics (job level, industry, contract type).

Difficult = dummy for difficult studies (see text).
Regulated = dummy for regulated studies (see text).

Table 3 Estimation with random effects on micro data

	Regressions							
Ln hourly wage	**A1**	**A2**	**A3**	**A4**	**B1**	**B2**	**B3**	**B4**
Erisk	3.741	3.681	3.548	2.978	3.092	3.065	2.979	2.718
(t-val)	(5.44)	(5.71)	(5.48)	(6.36)	(5.88)	(6.22)	(6.05)	(7.23)
Eskew	-3.656	-3.556	-3.496	-2.709	-2.649	-2.596	-2.605	-2.269
(t-val)	(-3.45)	(-3.43)	(-3.43)	(-3.78)	(-3.22)	(-3.15)	(-3.24)	(-3.67)
Grade variance	-0.546	-0.542	-0.506	-0.336	-0.516	-0.512	-0.480	-0.324
(t-val)	(-3.33)	(-3.39)	(-3.33)	(-3.30)	(-3.39)	(-3.46)	(-3.42)	(-3.38)
Grade skew	0.495	0.459	0.377	0.256	0.423	0.390	0.312	0.232
(t-val)	(2.18)	(2.07)	(1.88)	(2.03)	(2.05)	(1.93)	(1.74)	(1.94)
Individual grades	0.002	0.000	0.005	0.011	0.002	0.000	0.005	0.011
(t-val)	(0.74)	(0.08)	(1.67)	(4.62)	(0.73)	(0.06)	(1.65)	(4.54)
Difficult					-0.016	-0.015	-0.015	-0.005
(t-val)					(-1.29)	(-1.25)	(-1.32)	(-0.63)
Regulated					0.119	0.113	0.104	0.054
(t-val)					(2.97)	(3.01)	(3.34)	(2.51)
St dev RE	0.069	0.061	0.056	0.024	0.066	0.060	0.056	0.024
St dev error	0.202	0.199	0.198	0.188	0.202	0.199	0.198	0.188
R^2-within	0.000	0.028	0.038	0.139	0.000	0.028	0.038	0.140
R^2-between	0.376	0.414	0.486	0.772	0.459	0.488	0.551	0.782
R^2-overall	0.047	0.076	0.096	0.223	0.056	0.085	0.103	0.225

N = 44974, t-values based in standard errors clustered by field of education.

A = Regression without dummies for difficult and regulated studies.
B = Regression with dummies for difficult and regulated studies.

1 = intercept only.
2 = as 1, plus labour market variables (year dummies, regions, time since graduation).
3 = as 2, plus personal characteristics (age, gender, parental education).
4 = as 3, plus job characteristics (job level, industry, contract type.)

Difficult = dummy for difficult studies (see text).
Regulated = dummy for regulated studies (see text).

Table 4 Random effects estimation on micro data, subpopulations

Ln hourly wage	**Erisk**	**Eskew**	**Grade var.**	**Grade skew**	**Ind. grades**	**R^2 (overall)**	**N**
All	2.978	-2.709	-0.336	0.256	0.011	0.223	44974
(t-val)	(6.64)	(-3.78)	(-3.30)	(2.03)	(4.62)		
University	2.895	-2.933	-0.331	0.208	0.014	0.243	23578
(t-val)	(5.91)	(-3.98)	(-2.04)	(1.07)	(3.62)		
Higher vocational	2.496	-2.727	-0.038	-0.056	0.008	0.157	21396
(t-val)	(4.66)	(-2.90)	(-0.39)	(0.56)	(2.47)		
Men	2.354	-2.543	-0.364	0.258	0.022	0.226	19645
(t-val)	(8.50)	(-3.81)	(-3.64)	(2.53)	(6.75)		
Women	3.468	-3.178	-0.308	0.250	0.004	0.228	25329
(t-val)	(7.14)	(-4.11)	(-2.86)	(1.86)	(1.32)		
Ability 1	3.098	-1.633	-0.308	0.250	0.004	0.172	7891
(t-val)	(4.46)	(-1.03)	(-2.86)	(1.86)	(1.08)		
Ability 2	-0.089	1.534	-0.306	-0.029	0.012	0.195	8222
(t-val)	(-0.05)	(0.60)	(-3.40)	(-0.17)	(2.32)		
Ability 3	3.396	-2.64	-0.321	0.202	0.013	0.216	12021
(t-val)	(6.36)	(-4.13)	(-2.07)	(1.22)	(2.77)		
Ability 4	1.572	-1.896	-0.247	0.237	0.013	0.246	16840
(t-val)	(1.84)	(-1.40)	(-1.63)	(1.34)	(2.68)		
Government	0.717	-0.212	-0.357	0.255	0.011	0.245	4255
(t-val)	(0.96)	(-0.21)	(-2.84)	(2.05)	(1.96)		
Education	3.595	-3.025	-0.009	-0.062	0.004	0.249	6863
(t-val)	(10.03)	(-4.93)	(-0.08)	(-0.54)	(0.07)		
Services	-0.089	-0.199	-0.345	0.551	0.027	0.237	11967
(t-val)	(-0.15)	(-0.32)	(-2.38)	(4.11)	(6.48)		
Care	4.026	-3.598	-0.296	-0.153	-0.002	0.273	7822
(t-val)	(8.28)	(-3.35)	(-1.70)	(-0.93)	(0.34)		
Manufacturing	-0.056	-0.067	-0.445	0.379	0.021	0.252	4342
(t-val)	(-0.08)	(-0.07)	(-3.09)	(2.98)	(4.32)		
Other	0.424	-0.324	-0.213	0.262	0.012	0.171	9725
(t-val)	(0.54)	(-0.27)	(-1.56)	(1.47)	(3.19)		
Experience below mean	2.695	-2.427	-0.260	0.219	0.008	0.207	26554
(t-val)	(6.44)	(-3.78)	(-2.67)	(1.85)	(2.37)		
Experience above mean	3.167	-3.009	-0.396	0.232	0.017	0.229	18420
(t-val)	(8.58)	(-4.34)	(-3.73)	(1.92)	(5.24)		
Difficult	3.171	-2.904	-0.366	0.200	0.013	0.224	26551

	(t-val)	(6.45)	(-3.29)	(-2.89	(1.13)	(3.58)		
Not difficult		2.226	-2.212	-0.233	0.298	0.009	0.225	18423
	(t-val)	(3.27)	(-2.06)	(-2.27)	(2.61)	(2.90)		

t-values based in standard errors clustered by field of education.

Table 5 Elasticities

	Table 1 Aggregate	**Table 2 OLS**	**Table 3 RE**
Erisk	0.156	0.137	0.146
Eskew	-0.054	-0.052	-0.049
Grade variance	-0.108	-0.056	-0.085
Grade skew	0.026	0.007	0.015

Table 2 and table 3: Based on column A4

Table 6 Estimating equation (14)

Wage variables[1]: β	coefficient	t-value
Intercept	1.830	3.448
Erisk	2.941	6.79
Eskew	-3.122	-4.73
Individual grade	0.011	3.64
Male	0.021	5.57
HBO-level	-0.078	-7.83
Employer risk: δ		
Grade variance (δ_1)	-0.887	-9.63
Grade skew (δ_2)	0.619	4.03
Risk sharing[2] θ_i: α		
Intercept	-1.728	05.84
HBO level	0.48	2.03
Industry (reference:government)		
Education	-3.246	-1.18
Business services	2.028	5.74
Financial services	0.838	3.19
Health and personal care	-1.501	-2.42
Manufacturing	1.045	4.45
Other	1.908	5.80

N = 44947 R^2= 0.204 t-values based on standard errors clustered by fields of education

Characteristics θ_i .

Mean = 0.317; Minimum = 0.0003; Maximum = 0.959; Standard Deviation = 0.276

[1] also contains gender, age, parent education, months since graduation, dummy for Ph.D. student, higher vocational grade and job below level of eduction

[2] also contains 3 regional dummies, vocational graduate, job below level of education and months since graduation

Figure 1A Mean Ln hourly wage and grade variance, by education

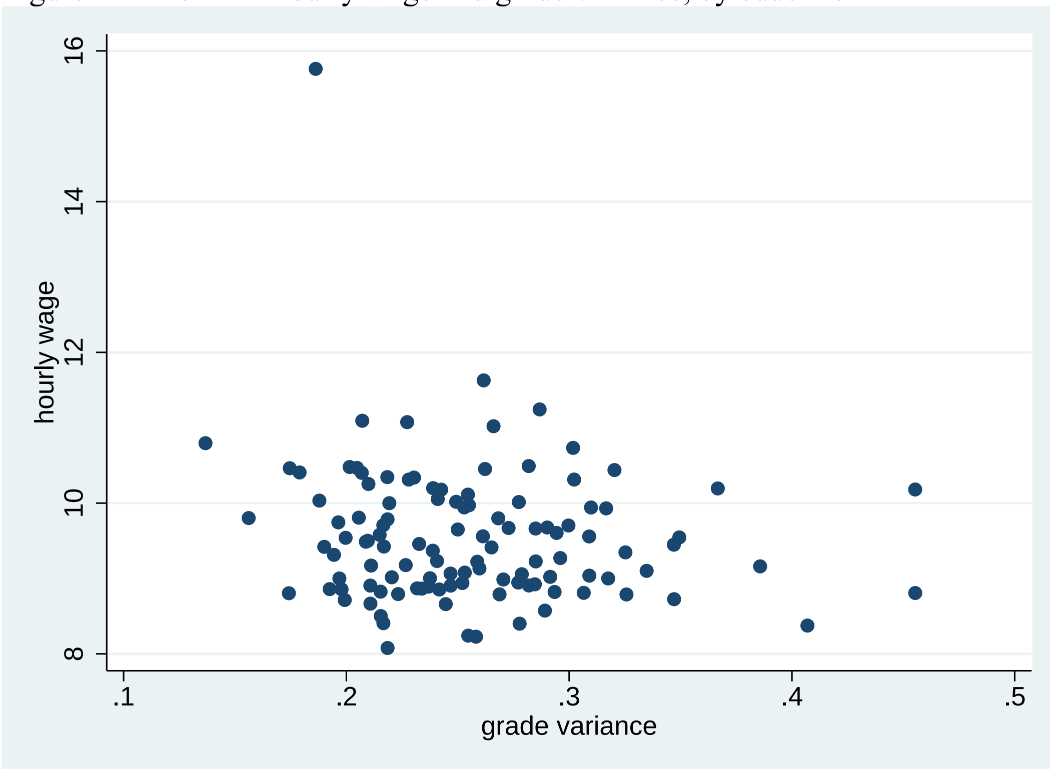
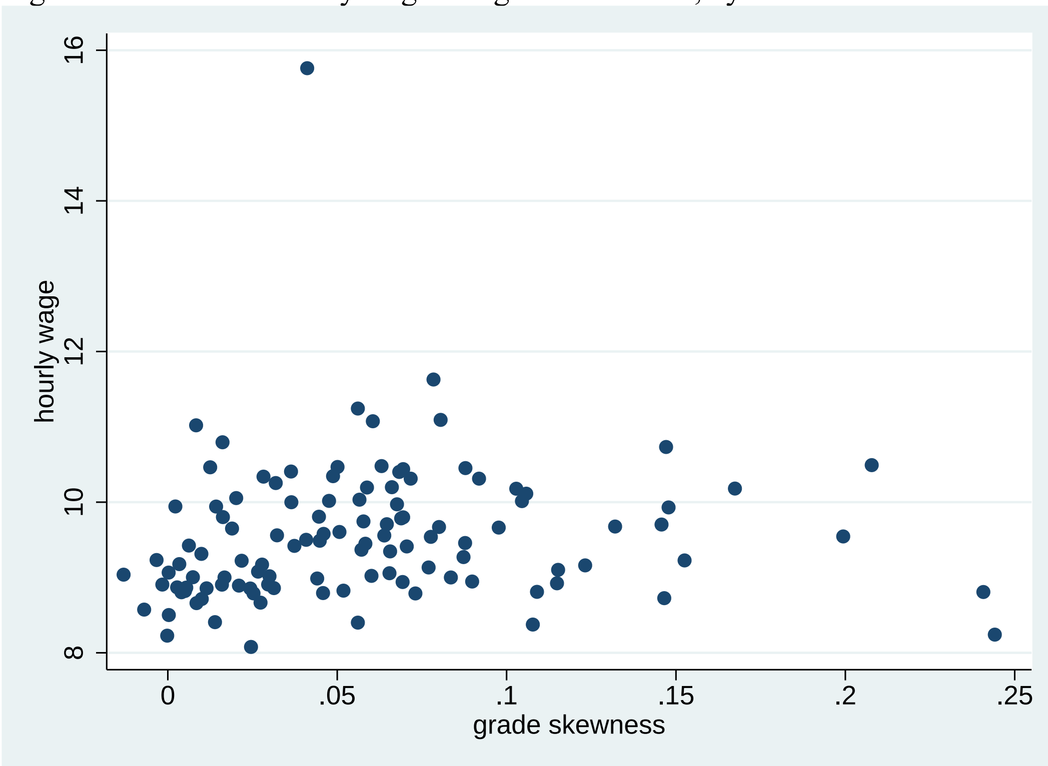


Figure 1B Mean Ln hourly wage and grade skewness, by education



Figure

Figure 2: Human capital as potential alternative explanation (males, university)

Figure 2A: Average 10 year earnings growth related to intercept

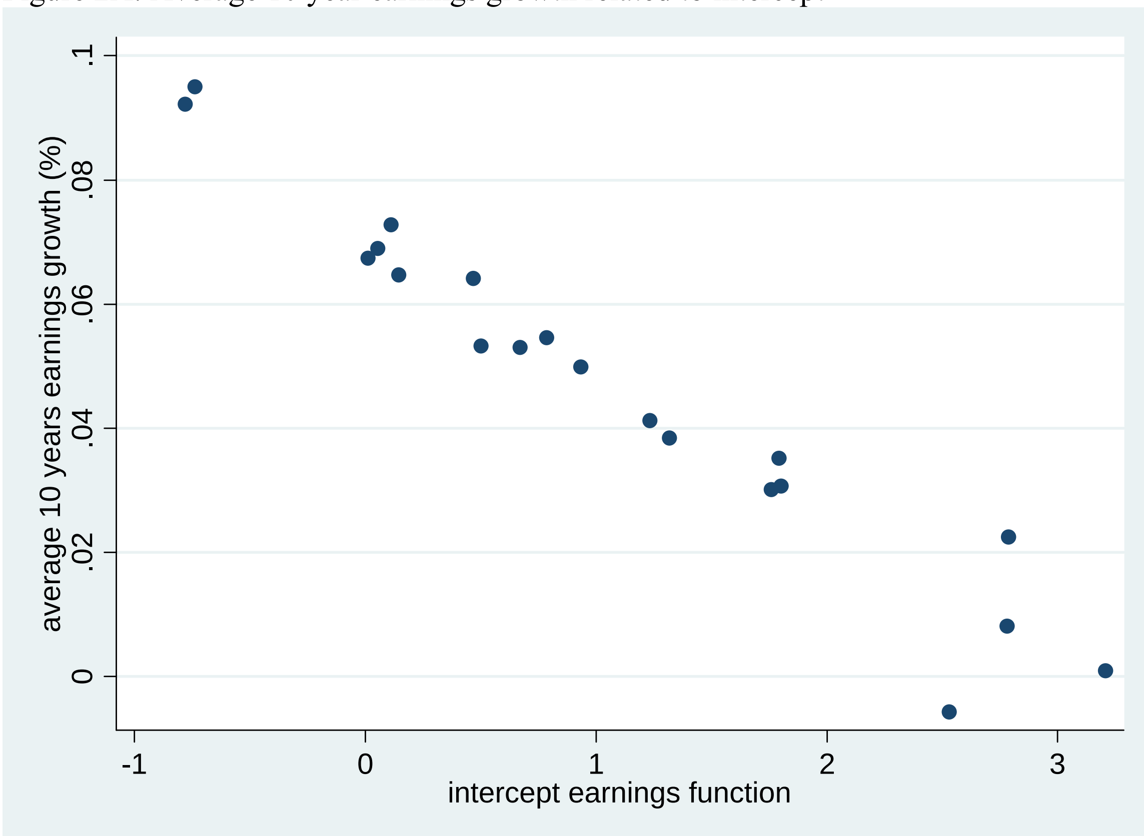


Figure 2 (continued)

Figures 2B-2E: relation between intercept or slopes and variance and skew of exam grades

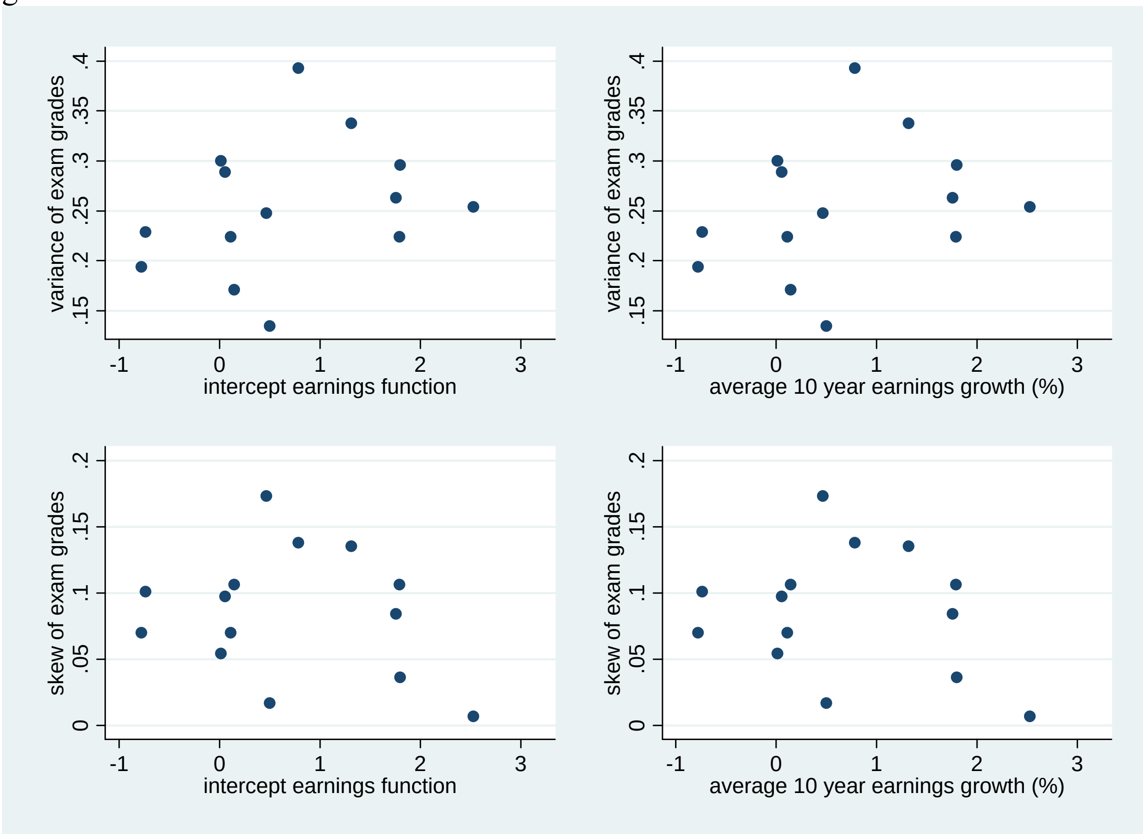
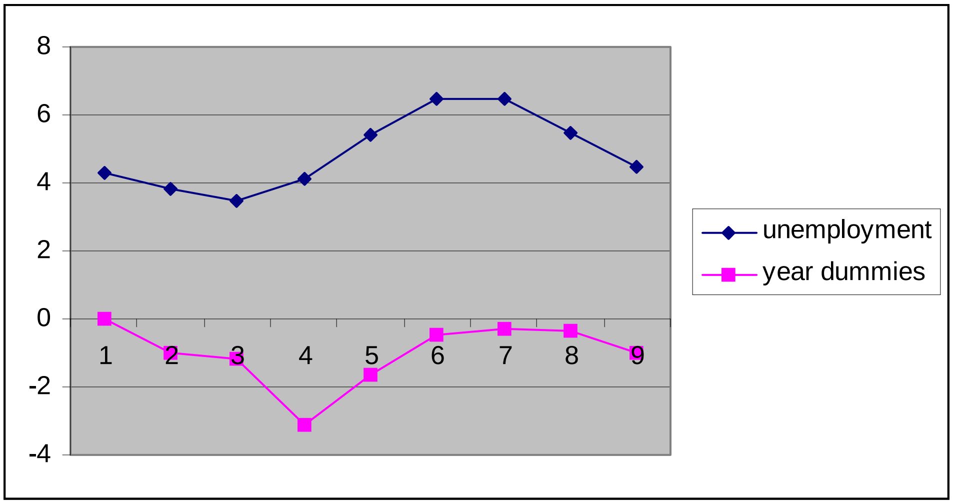


Figure 3. Risk shifting year dummies and national unemployment rate
(*source*: dummies regression (25);unemployment rate: CPB)



Appendix A. The wage equation for inexperienced workers

We can rewrite the right hand side of (4) in section 2.1 by substituting from a Taylor expansion of utility from hiring a worker with uncertain productivity at the experienced market wage function around the utility from hiring an experienced worker, at the experienced market wage function:

$$U(p\tilde{q}_{ij} - w(\tilde{q}_{ij})) = U(pq_{ij}^{o} - w(q_{ij}^{o})) + \frac{\partial U}{\partial \tilde{q}_{ij}}\Big|_{q_{ij}^{o}} (\tilde{q}_{ij} - q_{ij}^{o}) + \frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^{o}} (\tilde{q}_{ij} - q_{ij}^{o})^2 + \frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^{o}} (\tilde{q}_{ij} - q_{ij}^{o})^3 + R_4 \tag{A1}$$

where R_4 collects the higher order terms, ignored from now on. Since $\tilde{q}_{ij}$ is a random variable we need to assume that firm j actually considers expected profits in its decisions, so we get:

$$\begin{aligned} E[U(p\tilde{q}_{ij} - w(\tilde{q}_{ij}))] &= E[U(pq_{ij}^{o} - w(q_{ij}^{o}))] + \frac{\partial U}{\partial \tilde{q}_{ij}}\Big|_{q_{ij}^{o}} E[(\tilde{q}_{ij} - q_{ij}^{o})] + \frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^{o}} E[(\tilde{q}_{ij} - q_{ij}^{o})^2] \\ &+ \frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^{o}} E[(\tilde{q}_{ij} - q_{ij}^{o})^3] + E[R_4] \approx U(pq_{ij}^{o} - w(q_{ij}^{o})) + \frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^{o}} \sigma^2 + \frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^{o}} \kappa^3 \end{aligned} \tag{A2}$$

where σ and κ have been defined implicitly by the expected values they replace. Thus, rewriting,

$$U(pq_{ij}^{o} - w(q_{ij}^{o})) \approx E[U(p\tilde{q}_{ij} - w(\tilde{q}_{ij}))] - \frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^{o}} \sigma^2 - \frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^{o}} \kappa^3 . \tag{A3}$$

To develop the left hand side of (4), start from a first order Taylor expansion[16]

$$U(p\tilde{q}_{ij} - \tilde{w}(\tilde{q}_{ij})) = U(p\tilde{q}_{ij} - w(\tilde{q}_{ij})) + \frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})} (\tilde{w}(\tilde{q}_{ij}) - w(\tilde{q}_{ij})) \tag{A4}$$

Taking the expectation, as in (4):

$$E(U(p\tilde{q}_{ij} - \tilde{w}(\tilde{q}_{ij}))) = E(U(p\tilde{q}_{ij} - w(\tilde{q}_{ij}))) + E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})} (\tilde{w}(\tilde{q}_{ij}) - w(\tilde{q}_{ij})) \right) \tag{A5}$$

Equating RHS and LHS of (4), yields, after substituting (A5) and (A3),

[16] We use here only a first order expansion as we now expand in the wage rather than in stochastic productivity.

$$E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}\tilde{w}(\tilde{q}_{ij})\right)=E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}w(\tilde{q}_{ij})\right)-\left(\frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^o}\sigma^2+\frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^o}\kappa^3\right) \quad \text{(A6)}$$

Since $w(.)$ is determined on the labour market it is independent of the firm's utility and therefore we can write:

$$E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}\tilde{w}(\tilde{q}_{ij})\right)=E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}\right)E(w(\tilde{q}_{ij}))-\left(\frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^o}\sigma^2+\frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^o}\kappa^3\right)$$

(A7)

The expectation on the left hand side of (11) is of the form: $E(\xi_1\xi_2)$ and is therefore equal to $\operatorname{cov}(\xi_1,\xi_2)+E(\xi_1)E(\xi_2)$ or:

(A8)

$$E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}\right)E(\tilde{w}(\tilde{q}_{ij}))=$$

$$E\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})}\right)E(w(\tilde{q}_{ij}))-\left(\frac{1}{2}\frac{\partial^2 U}{\partial \tilde{q}_{ij}^2}\Big|_{q_{ij}^o}\sigma^2+\frac{1}{6}\frac{\partial^3 U}{\partial \tilde{q}_{ij}^3}\Big|_{q_{ij}^o}\kappa^3\right)-\operatorname{cov}\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})},\tilde{w}(\tilde{q}_{ij})\right)$$

Rearranging terms:

$$E(\tilde{w}(\tilde{q}_{ij}))=E(w(\tilde{q}_{ij}))-\frac{1}{2}\frac{\partial^2 U/\partial \tilde{q}_{ij}^2\big|_{q_{ij}}}{E\left(\partial U/\partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)}\sigma^2-\frac{1}{6}\frac{\partial^3 U/\partial \tilde{q}_{ij}^3\big|_{q_{ij}}}{E\left(\partial U/\partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)}\kappa^3$$

$$-\frac{\operatorname{cov}\left(\frac{\partial U}{\partial \tilde{w}(\tilde{q}_{ij})}\Big|_{w(\tilde{q}_{ij})},\tilde{w}(\tilde{q}_{ij})\right)}{E\left(\partial U/\partial \tilde{w}(\tilde{q}_{ij})\big|_{w(\tilde{q}_{ij})}\right)}$$

(A9)

Appendix B. Data description

Key variables

Individual Grade	Mean of individuals' average exam grade, self-reported
Gradevar	Variance of individuals' average exam grade, by field of study
Gradeskew	Skew (third moment) of individuals' average exam grade, by field of study
Erisk	Residual earnings variance, by field of study
Eskew	Residual earnings skew (third moment), by field of study
Residuals from earnings function: ln wage on dummies for education, cohort and region	

Correlations

	Hourly wage	Ln hourly wage	Erisk	Eskew	Grade var.	Grade skew
Hourly wage	1.000					
Ln hourly wage	0.995	1.000				
Erisk	0.349	0.333	1.000			
Eskew	0.038	0.056	0.739	1.000		
Grade variance	-0.058	-0.052	0.359	0.438	1.000	
Grade skew	0.122	0.132	0.275	0.345	0.658	1.000

Key data by education

VOCATIONAL	Grade var	Grade skew	Erisk	Eskew	Hourly wage	DumDiff	Ind. Grade	N
Business Economics/Business Sciences	0,247	0,016	0,036	0,010	2,185	1	6,782	565
Commerce	0,285	0,115	0,048	0,020	2,186	1	6,927	501
Business Informatics	0,194	0,010	0,046	0,013	2,232	1	6,984	549
Communication	0,198	0,011	0,043	0,013	2,181	1	7,052	480
Accountancy	0,270	0,044	0,041	0,010	2,194	1	6,752	431
International Business & Languages	0,245	0,008	0,033	0,010	2,157	1	7,004	368
Tourism & Leisure	0,218	0,025	0,039	0,008	2,089	0	7,011	421
Hotel Management	0,269	0,025	0,046	0,016	2,173	1	7,026	409
Small Business and Retail Management	0,291	0,060	0,055	0,018	2,199	1	7,025	218
Management, Economics & Law	0,242	0,024	0,045	0,016	2,180	1	6,896	461
Logistics & Economics	0,237	0,021	0,044	0,019	2,183	1	6,911	541
Facility Services	0,211	0,027	0,043	0,015	2,159	1	7,010	549
Journalism	0,285	0,153	0,050	0,014	2,222	1	7,118	462
Business Management	0,193	0,031	0,025	0,004	2,182	1	7,055	118
Fiscal Economics	0,347	0,058	0,055	0,017	2,239	1	6,895	220
European Professions	0,252	0,069	0,069	0,037	2,191	1	7,033	153
Leisure Management	0,258	0,000	0,048	0,016	2,108	1	7,013	154
International Business & Management	0,211	-0,002	0,064	0,017	2,187	1	7,306	62
Real Estate	0,217	0,014	0,052	0,012	2,129	1	6,966	73
Personnel & Labour	0,220	0,030	0,037	0,007	2,197	1	7,184	514
Socio-Cultural Studies	0,199	0,010	0,049	0,017	2,165	0	7,193	423
Social Work & Services	0,209	0,045	0,043	0,016	2,249	0	7,302	550
Social Pedagogy	0,232	-0,003	0,034	0,008	2,183	0	7,185	766
Socio-Legal Services	0,211	-0,028	0,034	0,010	2,218	0	7,181	381
Information Management	0,234	0,005	0,045	0,016	2,183	1	6,969	327
Creative Therapy	0,200	0,078	0,049	0,014	2,255	0	7,377	110
Medical LaboratoryTechnician	0,278	0,056	0,036	0,020	2,128	1	7,066	473
Nursing	0,259	0,022	0,036	0,016	2,222	0	7,168	708
Physiotherapy	0,266	0,008	0,083	0,026	2,399	1	6,993	534

Speech Therapy	0,296	0,087	0,058	0,021	2,227	0	7,256	412
Nutrition & Dietetics	0,253	0,027	0,052	0,022	2,206	1	7,000	467
Ergotherapy	0,261	0,032	0,053	0,021	2,258	0	7,150	535
Medical Imaging & Radiotherapy	0,347	0,147	0,016	0,003	2,167	1	7,092	184
Oral Hygiene	0,179	0,036	0,068	0,039	2,343	1	7,218	85
Environmental Management	0,238	0,007	0,039	0,008	2,199	1	6,929	380
Agri-Business	0,227	0,003	0,042	0,017	2,217	1	6,900	306
Animal Husbandry	0,215	0,000	0,060	0,025	2,140	1	6,955	322
Food Technology	0,197	0,017	0,029	0,004	2,200	1	6,837	95
Primary School Teacher	0,294	0,051	0,041	0,022	2,261	0	7,330	683
Physical Education Teacher, Grade 1	0,175	0,013	0,070	0,016	2,348	1	6,996	374
Dutch Language Teacher	0,196	0,058	0,062	0,023	2,275	0	7,237	251
Economics Teacher (General & Business)	0,217	0,006	0,053	0,019	2,241	1	7,011	307
Special Needs Teacher	0,273	0,080	0,032	0,008	2,264	0	7,471	277
Social Studies Teacher	0,174	0,004	0,044	0,010	2,176	1	7,046	87
Education Teacher	0,241	-0,003	0,037	0,006	2,223	0	7,409	208
Math/Physics Teacher	0,367	0,059	0,054	0,015	2,322	1	7,170	350
Geography/History Teacher	0,253	0,002	0,077	0,021	2,296	1	7,025	424
Arts & Crafts Teacher	0,215	0,052	0,108	0,049	2,175	1	7,300	65
English/French/German Language Teacher	0,320	0,069	0,083	0,018	2,345	0	7,295	469
Visual Arts & Design Teacher	0,455	0,241	0,099	0,057	2,176	0	7,390	295
Music Teacher	0,455	0,167	0,095	0,028	2,320	1	7,814	153
Drama Teacher	0,407	0,108	0,049	0,028	2,125	1	7,692	39
Chemical Technician	0,289	-0,007	0,037	0,011	2,149	1	6,975	223
Structural Engineering	0,223	0,046	0,035	0,010	2,174	1	7,062	438
Electrical Engineering	0,260	0,077	0,031	0,007	2,214	1	7,144	353
Civil Engineering	0,293	0,005	0,035	0,011	2,175	1	7,079	399
Chemical Engineering	0,279	0,065	0,029	0,006	2,202	1	7,091	532
Applied Informatics	0,325	0,066	0,046	0,015	2,235	1	7,123	500
Mechanical Engineering	0,277	0,090	0,041	0,015	2,192	1	7,053	385
Applied Physics	0,318	0,084	0,024	0,003	2,197	1	7,229	83
Fashion Management & Technology	0,255	0,244	0,030	0,005	2,108	1	7,168	110

Car Mechanics	0,309	-0,013	0,054	0,027	2,204	1	7,018	84
UNIVERSITY	**Grade var**	**Grade skew**	**Erisk**	**Eskew**	**Hourly wage**	**DumDiff**	**Ind. Grade**	**N**
Dutch	0,268	0,069	0,052	0,019	2,283	0	7,336	451
English	0,309	0,064	0,073	0,032	2,258	0	7,351	345
Other Languages	0,290	0,132	0,072	0,034	2,268	0	7,323	285
Philosophy/Theology	0,310	0,014	0,076	0,036	2,296	1	7,716	141
History	0,300	0,146	0,070	0,029	2,273	1	7,387	475
Language & Culture	0,265	0,071	0,064	0,036	2,243	0	7,200	358
Art History & Archeology	0,282	0,030	0,046	0,011	2,187	0	7,435	216
Corporate Communications	0,190	0,037	0,050	0,028	2,241	1	7,056	293
European Studies	0,250	0,019	0,090	0,081	2,262	1	6,945	64
Film, Television & Theatre Studies	0,247	0,000	0,061	0,015	2,204	1	7,259	108
Information Science	0,233	0,088	0,031	0,004	2,247	0	7,258	209
Chemistry	0,307	0,109	0,056	0,030	2,173	0	7,388	449
Computer Science	0,349	0,199	0,052	0,023	2,252	0	7,490	204
Biology	0,326	0,073	0,061	0,037	2,174	0	7,389	669
Pharmacy	0,262	0,078	0,049	0,009	2,452	0	7,283	417
Pure Mathematics/Physics	0,386	0,123	0,060	0,022	2,214	1	7,545	434
Agricultural Science	0,215	0,046	0,047	0,020	2,260	0	7,246	303
Chemical/Technical Agri-Sciences	0,210	0,041	0,053	0,020	2,251	1	7,214	610
Architecture	0,217	0,065	0,033	0,008	2,270	1	7,224	665
Mechanical Engineering	0,262	0,088	0,045	0,012	2,345	0	7,430	590
Electrical Engineering	0,302	0,092	0,059	0,035	2,334	0	7,433	358
Chemical Engineering	0,277	0,105	0,050	0,017	2,304	0	7,391	471
Civil Engineering	0,188	0,057	0,036	0,012	2,307	1	7,216	619
Technology & Management	0,137	0,016	0,046	0,019	2,378	1	7,133	606
Industrial Design	0,156	0,016	0,039	0,012	2,280	0	7,165	328
Aerospace Engineering	0,230	0,028	0,028	0,004	2,336	1	7,478	113
Applied Computer Science	0,249	0,048	0,038	0,017	2,301	0	7,462	235
Applied Mathematics/Physics	0,285	0,098	0,050	0,016	2,269	0	7,506	545
Economics	0,228	0,072	0,045	0,014	2,333	1	7,089	1.254
Business Science	0,205	0,050	0,048	0,017	2,348	1	7,072	631

Econometrics	0,302	0,147	0,060	0,022	2,374	0	7,319	434
Fiscal Economics	0,227	0,061	0,036	0,010	2,402	1	6,948	162
Business Studies	0,201	0,063	0,049	0,014	2,349	0	7,137	610
Dutch Law	0,243	0,103	0,041	0,014	2,320	1	7,124	908
Notarial Law	0,255	0,106	0,040	0,011	2,314	1	7,018	410
Fiscal Law	0,207	0,081	0,042	0,015	2,405	1	6,995	443
Healthcare	0,241	0,020	0,058	0,020	2,307	0	7,182	625
Medicine	0,287	0,056	0,048	0,013	2,420	1	7,237	889
Dentistry	0,186	0,041	0,129	-0,001	2,757	1	7,162	111
Biomedical Science	0,335	0,115	0,050	0,020	2,209	0	7,328	487
Veterinary Science	0,282	0,208	0,034	0,006	2,351	0	7,114	223
Sociology	0,317	0,148	0,049	0,010	2,294	1	7,236	388
Psychology	0,255	0,068	0,068	0,030	2,300	0	7,299	902
Political Science	0,218	0,049	0,055	0,018	2,337	0	7,299	369
Education Science	0,210	0,032	0,062	0,028	2,328	0	7,173	574
Applied Education Science	0,239	0,066	0,051	0,020	2,323	0	7,304	334
Cultural Anthropology	0,239	0,057	0,064	0,030	2,235	0	7,269	316
Communication	0,206	0,045	0,056	0,026	2,284	1	7,155	544
Socio-Cultural Science	0,219	0,036	0,053	0,026	2,301	0	7,178	662
Public Administration	0,207	0,068	0,044	0,013	2,341	1	7,152	822
Human Geography & Planning	0,218	0,069	0,042	0,014	2,280	1	7,092	919
TOTAL VOCATIONAL (weighted)	**0,258**	**0,044**	**0,045**	**0,016**	**2,217**	**0,499**	**7,104**	**21396**
TOTAL UNIVERSITY (weighted)	**0,245**	**0,074**	**0,050**	**0,019**	**2,318**	**0,565**	**7,211**	**23578**
TOTAL (weighted)	**0,253**	**0,054**	**0,047**	**0,017**	**2,252**	**0,538**	**7,141**	**44974**

SUBPOPULATIONS	**Grade var**	**Grade skew**	**Erisk**	**Eskew**	**Hourly wage**	**DumDiff**	**Ind. Grade**	**N**
All	0,253	0,054	0,047	0,017	2,252	0,538	7,141	44974
University	0,245	0,074	0,050	0,019	2,318	0,565	7,211	23578
Higher vocational	0,258	0,044	0,045	0,016	2,217	0,499	7,104	21396
Men	0,253	0,058	0,046	0,016	2,260	0,763	7,076	19645
Women	0,254	0,051	0,047	0,018	2,247	0,487	7,188	25329

Ability 1 (lowest)	0,254	0,028	0,045	0,017	2,234	0,318	7,166	7891
Ability 2	0,245	0,048	0,045	0,016	2,206	0,744	7,047	8222
Ability 3	0,258	0,069	0,048	0,017	2,265	0,716	7,101	12021
Ability 4 (highest)	0,257	0,076	0,049	0,017	2,309	0,663	7,250	16840
Government	0,240	0,057	0,046	0,016	2,294	0,721	7,123	4255
Education	0,280	0,060	0,050	0,021	2,267	0,248	7,345	6863
Services	0,244	0,057	0,045	0,015	2,245	0,829	7,052	11967
Care	0,251	0,035	0,047	0,016	2,290	0,325	7,181	7822
Manufacturing	0,255	0,065	0,045	0,016	2,250	0,662	7,079	4342
Other	0,253	0,058	0,047	0,017	2,199	0,766	7,100	9725
Experience below mean	0,253	0,052	0,046	0,017	2,231	0,593	7,149	26554
Experience above mean	0,254	0,059	0,047	0,017	2,287	0,618	7,128	18420
Difficult	0,248	0,054	0,046	0,015	2,257	1,000	7,060	26651
Easy	0,261	0,054	0,047	0,019	2,247	0,000	7,249	18423