SCHOOLS, SKILLS AND RISK

joop hartog

SCHOLAR

Universiteit van Amsterdam

Abstract

Skill development involves important choices for individuals and school designers: should individuals and schools specialise, or should they aim for an optimal combination of skills? We analyse this question by employing mean-standard deviation analysis and show how cost structure, benefit structure and risk attitudes jointly determine the optimal investment strategy.

As no one has read the draft yet, there is as yet no one to be grateful to; but it is surprisingly easy to earn an acknowledgement here.

First draft: 25-04-02; this version 04-11-04
File: skillspec270203

1. Introduction

In 1964, Gary Becker concluded a reflection on the variation in rates of return to education with the remark that the long pay-off period to education "increases the advantage of an education that is useful in many kinds of economic environments. If "liberal" education were identified with such flexible education, as it may well be, there would be an important economic argument for liberal education, as well as arguments based on intellectual and cultural considerations."(Becker, 1964, 204). The suggestion here is that a broad, non-specialised education lowers the risk in the rate of return to the education. In 1983, Sherwin Rosen analysed whether individuals should specialise or not when investing in human capital and concluded that individuals should specialise when the cost of developing the skills are separable, even when the return to the skill is a strictly positive function of the level of that skill. The result derives from optimal allocation of time (skill utilisation) once the investments have been competed.

Becker's statement is no more than a suggested hypothesis. Rosen's conclusion is based on a rigorous formal analysis but leaves out risk attitudes: individuals just maximise lifetime earnings in a riskless environment. In this paper we will analyse the specialisation question under conditions of risk in the returns to investments. As far as I am aware, there are no other economic analyses of this problem, besides my own exercises in Hartog (1981).

2. A linear world

2.1 Basic model

Consider a situation where individuals in period 1 go to school, and in period 2 (the rest of their working life) go to work. In school, they can develop two skills, 1 and 2. In a vocational school, they may have to choose between plumbing and carpentry, in secondary school they may have to choose between language or mathematics, in university economics they may choose between theory and econometrics. In several European school systems, they may have to choose between a vocational education and an academic (general) education, of equal length. We assume that any linear (non-negative) combination of skills is feasible. This either means that individuals can freely choose how to combine skill production (they have full choice over the curriculum) or it means we use the perspective of a planner who has to design the school curriculum. We assume that school attendance is given, either because individuals have decided on that in a sequential procedure or because they are under obligation from compulsory schooling laws. To them, the cost of education is fixed and independent of choices they make.

Time devoted to skill i is denoted s_i. Production of the skill is directly proportional to time spent and so are the returns. Skill i yields lifetime earnings (in the second period) of μ_i, with variance σ_i, and correlation ρ between the returns to the two skills. We assume arbitrarily $\sigma_2 > \sigma_1$ and hence require $\mu_2 > \mu_1$. Now, obviously, expected returns from schooling μ and expected variance σ^2 are given by

(1) $$\mu = s_1 \mu_1 + s_2 \mu_2$$

(2) $$\sigma^2 = s_1^2 \sigma_1^2 + s_2^2 \sigma_2^2 + 2 s_1 s_2 \rho \sigma_1 \sigma_2$$

These two conditions can be used to develop the opportunity frontier in terms of expected returns and variance for any combination of investment in the two skills, as a straight application of such analyses in finance[1]. In Figure 1, we draw the feasible combinations of μ and s that follow from different values of skill shares s_1 and s_2. The shape of the frontier is essentially determined by the correlation between the pay-offs to the two skills[2].

i) ? = 1

In this case, we can write

(3) $s^2 = (s_1 s_1 + s_2 s_2)^2$

(4) $s = s_1 s_1 + s_2 s_2$

With specialisation in skill 1 the investment yields the combination (μ_1, s_1), with specialisation in skill 2 the investment yields (μ_2, s_2) and varying investment share s_1 between 1 and 0 gives combinations along the straight line connecting the two points.

ii) ? = -1

We now write

(5) $s^2 = (s_1 s_1 - s_2 s_2)^2 = \{s_1(s_1 + s_2) - s_2\}^2$

For any given μ following from some investment share s_1 this will generate the lowest variance across the values of ? : perfect negative correlation gives maximum opportunity to diversify. In fact, setting $s_1 = \frac{s_2}{(s_1 + s_2)}$ yields zero variance. Starting at (μ_1, s_1), emanating from $s_1 = 1$, decreasing s_1 reduces standard deviation s linearly with μ, until s_1 becomes so low that the term under the square becomes negative; after turning point s = 0, and the standard deviation as the positive root increases again. At the turning point, s = 0, $\mu = (s_2 \mu_1 + s_1 \mu_2) / (s_1 + s_2)$

iii) -1 < ? < 1

---

[1] For a simple introduction, see Bodie, Kane and Marcus (1989).

[2] By limiting the curves to the interval (s_1, μ_1), (s_2, μ_2) we incorporate the restriction $0 \leq s_1, s_2 \leq 1$: going short in skill shares is hard to imagine.

The frontier smoothly moves inside the triangle as ? increases from –1 to +1. Minimum variance now requires setting $s_1 = \frac{s_2}{\left(s_1 + s_2\right)}$, but this does not allow to attain s = 0.

The efficient frontiers, of course, are restricted to the segments of the frontiers that are not negatively sloped.

To solve the individuals' optimum position, we look for maximum attainable utility. If individuals are homogenous in risk attitudes, only one point of the mean-risk frontier will be realised. If all individuals are risk loving or risk neutral, with non-positively sloped indifference curves, this single optimum will be the point of full specialisation in skill 2, which gives the highest return, at the highest risk. No one will invest in skill 1, unless the market would be able to support a higher return. The more interesting case is where individuals are heterogeneous in risk attitudes, a case with intuitive and econometric support (Hartog, Ferrer-i-Carbonell and Jonker, 2002; Halek and Eisenhauer, 2001).[3] With heterogeneous risk attitudes and mean-risk frontiers conditional on the correlation between skill returns as depicted in Figure 1, individuals will locate on the frontier depending on the slope of their mean-risk indifference curves. Risk loving and risk neutral individuals will only be found to specialise in high risk skill 2. Risk averse individuals will locate on the frontier where their indifference curve is tangent to the frontier. Even risk averse individuals may fully specialise in high risk skill 2, if their indifference curve indeed pushes them to that corner. But if the frontier has a segment with negative slope for high values of s_1, no one will ever fully specialise in low risk skill 1, as better deals (higher return for given risk) are always available. A negative slope obtains if $? < \frac{s_1}{s_2}$, as can be easily verified.[4] We can summarise the results as follows.

With a linear technology in school and production, based on one low risk-low return and one high risk-high return skill, and schooling cost independent of skill shares:

1. *risk neutral and risk loving individuals will always fully specialise in the high risk-high return skill*
2. *with correlation in the returns to the two skills smaller than* $\frac{s_1}{s_2}$*, there will be no individuals fully specialising in the low risk-low return skill.*

Result 2 implies that with low correlation between returns to skills it is not rational to design a school curriculum with full specialisation in the low-return-low risk skill. Since the critical value is non-negative this conclusion holds a fortiori for negative correlation between skill returns. In such cases, the curriculum should always at least contain a dash of the risky skill.

The results should be conditioned on student abilities if they are observable and relevant for frontier or preferences. Level and type of ability may affect both the level and the riskiness of

[3] Hartog et al (2002) with Dutch questionnaire data find risk aversion declining in level of education. Halek and Eisenhauer (2001) find no clear and unequivocal relation with education in the USA; it declines with education for a hypothetical switch to a job with risky income, it is not significant or increasing in purchasing behaviour of life insurance. Using Italian questionnaire data, Ventura and Eisenhauer (2002) find a negative relation between risk aversion and level of schooling.

[4] In (2), substitute for s_1 from (1), differentiate to μ, require this to be negative at $s_1 = 0$ and solve for ?.

skill returns. For example, it may well be that for high ability individuals the risk premium is higher, which makes full specialisation in the risky skill more attractive than for low ability individuals.

What parts of the opportunity frontier will be realised in equilibrium depends on supply and demand conditions. If the $\{m_i, s_i^2\}$ combinations are immovable dictated by technology cum product demand conditions, the frontier will be as derived above. Preferences determine the realised equilibrium points as tangencies (or corner solutions) of (m, s^2) indifference curves and the frontier. With sufficiently heterogeneous preferences, the entire opportunity frontier may be traced out (this is an example of an identifiable solution in a hedonic model). But suppose now, that dispersions are technologically dictated and cannot be adjusted, while mean compensation can be adjusted, as there are rents to be shared, or output prices can be adjusted without complete loss of demand. Suppose, the two full specialisation points are each available in the market, and individuals have identical risk attitudes. Then, the wage differential between the two specialised jobs compensates for the risk differential:

$$m_2 - m_1 = \frac{1}{2} V_a \left(s_2^2 - s_1^2 \right)$$

Where $V_a = -U''/U'$ is absolute risk aversion (see Hartog and Vijverberg, 2003). That is, the two (m, s^2) combinations are on the same indifference curve. Suppose, that all linear combinations of the two skills are available as job options, with output a linear combination, as assumed above. Then, if the frontier is below the indifference curve connecting the two points of full specialisation, these skill combinations require a higher wage than follows from the opportunity frontier, as otherwise supply to the skill combination will not be forthcoming. Hence, employers or consumers should give up some rent to compensate workers for risk. Conversely, if the frontier is above the indifference curve connecting the two specialised jobs, employers can afford to pay less to workers and retain some rent. Clearly, there is potential scope for bargaining over the risk component in wages. Note that the slope of the critical indifference curve is given by $(V_a/2) d s^2$.

With only two skills, it is straightforward to derive an explicit solution. Let's apply the Mean-Variance model, and assume individuals maximize a utility function

$$U = U\left(m, s^2\right)$$

with $A = -u'_m / u'_{s^2}$ the required minimum compensation for an increase in risk (the reservation price of risk). Maximising utility with respect to s, using $s_2 = 1 - s$, and substituting (1) and (2) in the utility function yields

$$s_1^* = \left\{ s_2^2 - \frac{m_2 - m_1}{2A} \right\} / \left\{ s_1^2 + s_2^2 - 2 r s_1 s_2 \right\}$$

Note that if we simplify to a risk free asset $\left(s_1^2 = r = 0 \right)$ and a risky asset, the investment share in the risky asset will be

$$s_2^* = \frac{m_2 - m_1}{2As_2^2}$$

which is exactly the solution Shaw (1996) derives. In finance, $(m_2 - m_1)/s_2^2$, “excess return” relative to the variance, is called the Sharpe ratio (see e.g. Bodie, Kane and Marcus, 1989).

2.2 Adding cost

Conclusions on the rationality of specialisation obviously depend on the relative slopes of indifference curves and the opportunity frontier. Two rules stand out. One: if, at full specialisation in skill 2 (high risk), the slope of the frontier is negative, a risk neutral or risk averse individual will never specialise in that skill; a risk lover may specialise, if the market compensation for risk is too low relative to the reservation premium. Two: if, at full specialisation in skill 1 (low risk), the slope of the frontier is negative, no one, whatever her risk attitude, will specialise in that skill, as the backward bending frontier always provides a skill combination with the same risk and higher returns. These two rules are behind the basic results above, but hey also apply under the more general conditions that we will consider below. The rules indicate how rational individuals will behave in the face of given opportunities. By implication, some segments of the opportunity frontier may not be observed, as no individual would be interested.

As a first step towards generalisation, let’s add the cost of producing skills, either because individuals have to pay for the investment or because we take a social perspective and seek a social optimum. We assume that at the time of deciding, the cost of investments in skills are fully known with certainty and that individuals are interested in net returns. Then, introducing costs serves to reduce returns without affecting variance. This means that the μ axis is transformed, with transformation depending on the nature of the cost function. If cost are independent of specialisation, the result is very similar to the case of no cost: all points between μ_1 and μ_2 are reduced by the same amount and the efficient frontier slides down along the μ axis while maintaining its slope. The optimum will be affected only if the slope of an indifference curve at given s is sensitive to μ, i.e. if absolute risk aversion (defined here as minimum compensation required for a given increase in s) is not constant but depends on wealth (income). If cost are separable, and each proportional to the share invested in skill i, μ(i) is effectively redefined to μ(i) – c(i) and the entire analysis above holds in terms of these net returns to investment. But if investment cost is not separable in the shares s_i the frontiers are twisted, depending on the cost function.

Consider the cost function

(6) $$C = s_1 c_1 + s_2 c_2 + f(s_1)$$

with f(0) = f(1) = 0: cost are proportional to investment shares plus an interaction effect that moves from zero to zero as the investment share s_i moves from zero to one, and may be positive or negative for the intermediate range. Combined skill production may have cost advantages or disadvantages, or have advantages and disadvantages over different segments of the range. This means a twist in the μ-axis of Figure 1, where many configurations are possible, and generally, even the linear frontier for perfect negative correlation will become non-linear.

Assume the diversification cost term $f(s_1)$ is a smooth curve, with a single maximum or minimum. Then, non-risk averse individuals will still fully specialise in the high risk, high mean skill if at $s_2 = 1$, the net slope is still positive, or

$$(7) \qquad \frac{d\mu}{ds} - \frac{dC}{ds} > 0$$

Now,

$$(8) \qquad \frac{d\mu}{ds} = \frac{\partial\mu}{\partial s_1}\frac{\partial s_1}{\partial s} = \frac{\partial\mu/\partial s_1}{\partial s/\partial s_1}$$

$$(9) \qquad \frac{dC}{ds} = \frac{\partial C}{\partial s_1}\frac{\partial s_1}{\partial s} = \frac{\partial C/\partial s_1}{\partial s/\partial s_1}$$

Combining, we get

$$\frac{d\mu}{ds} - \frac{dc}{ds} = \frac{\partial\mu/\partial s_1 - \partial C/\partial s_1}{\partial s/\partial s_1}$$

We know that at $s_2 = 1, d\mu/ds > 0$. Also, $\partial\mu/\partial s_1 = \mu_1 - \mu_2 < 0$. By (8), this implies $\partial s/\partial s_1 < 0$ and hence, (7) requires

$$(10) \qquad \frac{\partial\mu}{\partial s_1} < \frac{\partial C}{\partial s_1}$$

or

$$(11) \qquad \frac{\partial f}{\partial s_1} > (\mu_1 - \mu_2) - (c_1 - c_2)$$

If the opposite of (11) holds, the net slope of the frontier at $s_2 = 1$ is negative; risk neutral and risk averse individuals will certainly not fully specialise in skill 2, while risk lovers may specialise, depending on slope conditions.

Similarly, if at $s_1 = 1$ the net frontier has a negative slope, no individual will fully specialise in skill 1, because there is always a better mean-risk combination available. A negative net slope obtains if

$$(12) \qquad \frac{d\mu}{ds} - \frac{dC}{ds} < 0$$

Using the same reasoning as in deriving (11), we can derive the required condition for (12) 0 hold as[5]

(13a) $\frac{\partial f}{\partial s_1} > (\mu_1 - \mu_2) - (c_1 - c_2) \quad if \quad ? < s_1/s_2$

(13b) $\frac{\partial f}{\partial s_1} < (\mu_1 - \mu_2) - (c_1 - c_2) \quad if \quad ? > s_1/s_2$

We may again summarise. Define $(\mu_1 - \mu_2) - (c_1 - c_2)$ as the maximum return gap, and $\partial f / \partial s_1$ as the diversification cost slope.

If, at full specialisation in the high risk skill:

- *the diversification cost slope is greater than the maximum return gap, non–risk averse individuals will fully specialise in that skill*
- *the diversification cost slope is smaller than the maximum return gap, risk neutral and risk averse individuals will not fully specialise in that skill, risk loving individuals may do so.*

No individual will fully specialise in the low risk skill if

- *correlation in skill returns is low ($? < s_1/s_2$), and if at full specialisation in the low risk skill, the diversification cost slope is larger than the maximum return gap,*
- *correlation in skill returns is high ($? > s_1/s_2$), and if at full specialisation in the low risk skill, the diversification cost slope is smaller than the maximum return gap.*

Note that this conclusion (and the earlier one) can be read as an advice for school curriculum design. If no one wants full specialisation in skill 1, it does not make sense to offer that. The broad implication is that risk properties and risk attitudes are relevant for curriculum design. Rosen's conclusion that skill specialisation dominates skill combination if cost are separable does not hold if skill investments have risky returns.

3. A general model of curriculum choice

If we generalise to many skills, the basic approach is unaffected. We can still consider the mean-risk frontier, the cost function and the indifference curves to determine optimum choices and eliminate inefficient ones. But with many skills, it is not easy to derive restrictions on the shape of the mean-risk frontier. One may even generalise further. A labour market separable in skills is a strong assumption, explicitly discarded by hedonic models and indeed rejected by a test on the US wage structure in terms of required aptitudes and abilities (Hartog, 1980).

We will now develop a general model of curriculum choice with uncertain pay-off without imposing linear separability on the labour market. As before, we assume two periods, schooling

[5] We now have two conditions as the slope $d\mu/ds$ in (8) at $s_1 = 1$ depends on ? relative to s_1/s_2, and so will the sign of $\partial s / \partial s_1$.

and work. In the schooling period, individuals can choose their curriculum as proportions s_i, $i=1, \dots$ I spent on skill i. The expected pay-off M is a general function of the skill shares vector s, M(s) and the variance is a function V(s). The cost of skill investment is C(s). The generalization in M(s) and V(s) acknowledges the hedonics literature denial of linear separability of skills in the labour market.

We assume individuals evaluate returns with a mean-risk utility function U(M,V). We believe this is a most attractive way of representing preferences. It has intuitive appeal, as mean and variance (our measure of risk) appear as separate inputs and it's quite conceivable that this is comfortably close to individuals' perception of uncertain situations. We consider it more attractive than expected utility as a probability weighting of the utility of potential outcomes, both in the descriptive sense of individuals perceptions and in the sense of analytical tractability. Intertemporal expected utility models generate optimum conditions featuring inter alia covariance between marginal utility of income (consumption) and returns, and such terms are not very transparent (cf the model developed by Levhari and Weiss, 1974). Defining utility on mean and variance connects directly to observable variables, which is a great advantage for empirical work.

We formulate the choice problem facing the individual as

$$\overset{\max!}{s_i, i=1,2,..\mathrm{I}} \quad \mathrm{e}^{-r} U\{M(s), \mathrm{V}(s)\} + U\{Y - C(s), 0\}$$

$$\text{(14)} \qquad \text{s.t.} \quad \sum_{i=1}^{I} s_i \leq 1$$

$$0 \leq s_i \quad , i = 1,2,...\mathrm{I}$$

Second period utility, derived from mean and variance of returns, is discounted by discount rate ? . Costs of investment C, deducted from exogenous first period income Y are riskless, and evaluated at the same utility function, at variance equal to zero. The essential Kuhn-Tucker conditions for an optimum can be derived as

$$\text{(15)} \qquad \left\{ e^{-?} \; U_m^{'} \; \frac{\partial M}{\partial s_i} + e^{-?} \; U_v^{'} \; \frac{\partial V}{\partial s_i} - U_C^{'} \; \frac{\partial C}{\partial s_i} - ? \right\} \; s_i = 0 \, , \; all \quad i$$

$U_x^{'}$ indicates marginal utility of x, l is the multiplier on the summation constraint. For interior solutions $0 < s_i < 1$, the term in braces equals zero, for corner solutions $(s_i = 0)$ it is non-positive. For any two interior skill shares, we may rewrite to

$$\text{(16)} \qquad e^{-?} \left\{ \frac{\partial M}{\partial s_i} - \frac{\partial M}{\partial s_j} \right\} - \frac{U_c^{'}}{U_m^{'}} \left\{ \frac{\partial C}{\partial s_i} - \frac{\partial C}{\partial s_j} \right\} = -e^{-?} \; \frac{U_v^{'}}{U_m^{'}} \left\{ \frac{\partial V}{\partial s_i} - \frac{\partial V}{\partial s_j} \right\}$$

At optimum skill composition of the curriculum, the difference in net marginal benefits should equal the difference in marginal risk, evaluated at the marginal rate of substitution between mean and risk, or the reservation price of risk. We can also write (16) as an equality of slopes

$$(17) \qquad -\frac{U_v'}{U_m'} = \frac{\left(\frac{\partial M}{\partial s_i} - \frac{\partial M}{\partial s_j}\right) - \frac{U_C'}{U_m'} e^{?} \left(\frac{\partial C}{\partial s_i} - \frac{\partial C}{\partial s_j}\right)}{\frac{\partial V}{\partial s_i} - \frac{\partial V}{\partial s_j}}$$

As usual, the left-hand side is the slope of the indifference curve and the right-hand side is the slope of the opportunity frontier. It's a generalization of the Sharpe ratio, the excess return per unit of risk used in the finance literature (see e.g. Bodie, Kane and Marcus, 1989).

Individual heterogeneity enters the optimum condition in several ways. Ability differences may affect marginal cost, expected returns and risk. Thus, empirical analysis of the opportunity frontier may have to differentiate between frontiers faced by different types of individuals. On the other side, individual risk attitudes, reflected in the reservation price for risk, may differ, and as a result, of course, equilibrium matches may exhibit a segmentation with the less risk averse choosing a riskier curriculum, just as in the hedonic model of labour market allocation.

There are several ways to allow for preferences for the different types of skills. A simple way is to assume that both utility when working and utility when in school are directly sensitive to the type of skill, i.e. the vector s is an argument of the utility function. In equation (17), the reservation price of risk on the left-hand side will then be augmented by

$$(17') \qquad -\frac{U'_{wsi} - U'_{wsi} + e^{?}\left(U'_{ssi} - U'_{ssj}\right)}{U'_{m}} \; \frac{1}{\frac{\partial V}{\partial s_i} - \frac{\partial V}{\partial s_j}}$$

U'_{wsi} refers to the derivative of the utility function to skill type s_i when the individual is working, U'_{ssi} to the derivative when the individual is still in school (the subscripts w and s reflect that the marginal utilities are evaluated at different values of the utility function). Hence, the reservation price of risk is corrected for differences in marginal utilities of the two skill types, expressed in monetary equivalents expressed per unit difference in marginal risk. Skill preferences may be related to ability differences, creating a correlation between the reservation price and the slope of the opportunity frontier.

Naturally, returning to the linear model of section 2 generates conformingly simplified equilibrium conditions. With equal marginal cost and a labour market linear in two skills we can explicitly solve for the relationship between the skill proportions:

$$(18) \qquad s_1 = s_2 \frac{s_2^2 - ?s_1 s_2}{s_1^2 - ?s_1 s_2} - \frac{1}{2}\frac{U'_m}{U'_v}\frac{M_1 - M_2}{s_1^2 - ?s_1 s_2}$$

With $s_1 = s_2$ and $? = 0$, we can further simplify to

$$(19) \qquad s_1 - s_2 = -\frac{1}{2}\frac{U'_m}{U'_v}\frac{M_1 - M_2}{s_2}$$

which is even closer to the Sharpe ratio from finance theory. The equality of slopes would then read

$$(20)\qquad -\frac{U_v^{'}}{U_m^{'}} = \frac{1}{2}\frac{M_1 - M_2}{(s_1 - s_2)\mathrm{s}^{2}}$$

We conclude this analysis by noting that solving optimum problem (14) yields a set of optimum values $s_i^*, i = 1,2,\ldots \mathrm{I}$, which in turn define the individuals' net returns:

$$(21)\qquad U\{M(s^*), \mathrm{V}(s^*)\} - U\{Y - C(s^*), 0\}$$

4 The option value of education

Modelling schooling choice as a single up-front decision on total schooling length ignores that school systems usually have more than one decision point. Moreover, single decision models on intended total schooling length have to interpret drop-out as a failure, and the entry decision as an ex post erroneous decision. Allowing explicitly for future decision nodes in a schooling career allows to include the option value of schooling decisions and to interpret the decision to abstain from continued education as a rational response to improved information generated by the early stages in the schooling process.

Assume, for the sake of simplicity, that schooling takes two years. Each school year has an optimal curriculum s. Entering the labour market with k years of education gives access to an earnings distribution with mean M_k and variance V_k, which has utility value $U(M_k, V_k)$. Individuals differ in ability, unknown to them when entering the first stage of schooling, affecting schooling cost. When deciding on the first stage of schooling, individuals only know that with probability p they are able individuals with schooling cost C_a, and with probability $(1-p)$ they are "dumb" individuals with the higher schooling cost C_d. Schooling cost is private information, and the wage distributions are not conditioned on this ability.

If individuals would have to make a single binding decision on an entire two-period schooling career, they would enter school if

$$(22)\qquad \begin{array}{l} \int_2^{\infty} \left[U\{M_2, V_2\} - U\{M_0, V_0\}\right] e^{-?t}\, dt \;> \\ \int_0^{2} \left[U\{M_0 0, V_0\} - pU\{Y - C_a, 0\} - (1-p)U\{Y - C_d, 0\}\right] e^{-?t}\, dt \end{array}$$

The lifetime discounted utility gain from schooling should surpass the discounted cost, the utility difference between being in school and going to work. The utility of being in school is represented as expected utility; generalising this to the representation of prospect theory, with additional weighting of probabilities and including reference points would better reflect individual decision making, but for present purposes probably generate no additional insights.

Now suppose that individuals can take a new decision at the end of the first schooling period. We will assume that at the end of the first schooling stage, the individual knows her ability; for low ability, the best option is then to quit school, for high ability the best option is to continue:

(23) $$\int_1^{\infty} U\{M_2, V_2\} e^{-?t}\, dt + \int_0^1 U\{Y - C_d, 0\} e^{-?t}\, dt < \int_0^{\infty} U\{M_1, V_1\} e^{-?t}\, dt$$

(24) $$\int_1^{\infty} U\{M_2, V_2\} e^{-?t} dt + \int_0^1 U\{Y - C_a, 0\} e^{-?t}\, dt > \int_0^{\infty} U\{M_1, V_1\} e^{-?t}\, dt$$

If we now analyse the individual's decision to enter school, we should acknowledge that at the end of the first stage the individual knows her ability, and knows that as high-ability student she will continue, as low-ability student she will not. The first stage of schooling provides the option to quit or to continue, and the optimal decision at the end of the first stage should enter the initial decision.

The individual will now enter the first stage of schooling if

(25) $$\begin{aligned} & p\left[\int_2^{\infty} U\{M_2, V_2\} e^{-?t}\, dt + \int_1^2 U\{Y - C_a, 0\} e^{-?t}\, dt\right] + (1-p)\int_1^{\infty} U\{M_1, V_1\} e^{-?t}\, dt + \\ & + \int_0^1 \left[p U_2\{Y - C_a, 0\} + (1-p) U\{Y - C_d, 0\}\right] e^{-?t}\, dt > \int_0^{\infty} U\{M_0, V_0\} e^{-?t}\, dt \end{aligned}$$

With probability p, the individual continues after stage 1 and receives the associated lifetime utility after passing another year in school, with probability 1-p the individual enters the labour market after period 1 is over. In both cases, the student first receives the utility from being in school in the first period. This total expected utility from entering school should surpass the utility from going straight to work.

Condition (25) can be rewritten as

(26) $$\begin{aligned} & \int_2^{\infty} \left[U\{M_2, V_2\} - U\{M_0, V_0\}\right] e^{-?t}\, dt > \\ & \int_0^2 \left[U\{M_0, V_0\} - pU\{Y - C_a, 0\} - (1-p) U\{Y - C_d, 0\}\right] e^{-?t} + \\ & + (1-p)\left[\int_2^{\infty} \left(\left[U\{M_2, V_2\} - U\{M_1, V_1\}\right] e^{-?t} dt\right) - \int_1^2 \left(U\{M_1, V_1\} - U\{Y - C_d, 0\}\right) e^{-?t} dt\right] \end{aligned}$$

Now, from condition (23), we know that the last term in square brackets in (26) is negative: it is the return from second stage education for the low ability individuals (discounted one more period, to the moment of the first schooling decision). Comparing condition (22) to condition (26) we see the implication of taking into account the option value of the second stage decision. The threshold for entering the first stage of schooling is lowered. Hence, for some range of parameter values, individuals would not enter school if they would have to sign up for a full two-period education, whereas they will enter school if they can take a new decision at the end of the first school period. They enter school because they can avoid the loss from being low ability-high cost students, once they have learned their ability level.

It is also interesting to compare the student population in case of first stage schooling evaluated by itself with the case of a potential second stage. An individual assessing the first stage of school as an isolated investment will enter school if

(27) $$\int_1^{\infty} U\{M_1,V_1\}e^{-?t}dt+\int_0^{1}E(Y)e^{-?t}dt > \int_0^{\infty} U\{M_0,V_0\}e^{-?t}dt$$

where E(Y) is the expectation of utility while in school and not yet knowing ability (and cost). If the assessment includes the option value of continuing school after the first stage or going to work, the decision rule is given by (25), which can be rewritten as

(28) $$\begin{aligned}&\int_1^{\infty} U\{M_1,V_1\}e^{-?t}dt+\int_0^{1}E(Y)e^{-?t}dt+\\&p\left[\int_2^{\infty}\left(U\{M_2,V_2\}-U\{M_1,V_1\}\right)e^{-?t}dt+\int_1^{2}\left(Y_a-U\{M_1,V_1\}\right)e^{-?t}dt\right]\\&\qquad > \int_0^{\infty}U(M_0,V_0)e^{-?t}dt\end{aligned}$$

where Y_a stands for the utility while in schooling knowing that the ability level is high. The term in square brackets is the net lifetime utility from attending second stage schooling, which we have assumed positive in (24). Thus, compared to single state schooling, it's easier to meet the threshold condition for positive expected benefits: expected benefits are raised by the probability of reaping the additional benefits from being a high ability individual.

If we allow for individual heterogeneity, class composition will be sensitive to the difference in perspective. With first stage schooling assessed on its own benefits, the school will only have students for whom condition (27) holds. With the option of continuing education after the first stage, the school will also attract students for whom (28) holds. This may affect the student mix in terms of risk attitudes as reflected in the utility function. In a more general specification derived from prospect theory, this might also refer to the weights of p. Similar conclusions hold for the comparison between single binding decisions on the total schooling length and the decision in two stages including the option values for the second stage.

Of course we can reduce condition (26) to the mold of more conventional analysis, by reducing utility maximisation to earnings maximisation (ignoring risk). If we write E(C) for expected schooling cost, $pC_a+(1-p)C_d$, the simplification leads to the condition for entering initial schooling in a two-stage education

(29) $$\begin{aligned}&\int_2^{\infty}(w_2-w_0)e^{-?t}dt>\int_o^{2}\{w_o+E(C)\}e^{-?t}dt\ +\\&+(1-p)\left[\int_2^{\infty}(w_2-w_1)e^{-?t}dt-\int_1^{2}(w_1+C_d)e^{-?t}dt\right]\end{aligned}$$

Again, with the last term representing the benefit from continued education for low ability individuals, assumed negative, the threshold for entering is reduced and otherwise non-entering individuals may start school because they have the option to quit at the end of the first stage. In this case it is even straightforward to solve for the equilibrium wage rate w_2 that makes individuals indifferent between starting school or going to work right-away:

(30) $$w_2=\frac{w_0}{p}e^{2?}+\left(e^{?}-1\right)\frac{e^{?}E(C)+pC_a}{p}-\frac{1-p}{p}w_1e^{?}$$

Clearly, with $p = 1$ and $C = 0$, we are back in the Mincer world where every additional year of education yields a return ? . With $p \neq 1$, there is a risk of being low-ability which at given w_1 requires additional compensation in w_2, although compensation in w_1 would be more likely (both w_1 and w_2 should attain equilibrium values). Naturally, non-zero direct costs also push up the equilibrium wage rate. Our more general model also implies equilibrium wage conditions like (30), but it is more difficult to formulate this as an explicit solution.

The analysis has the following implications

- drop-out
  If we take every year (class) as a stage in the schooling process, the analysis directly applies to drop-outs. While drop-outs may have positive ex ante and negative ex post returns, it's quite conceivable that ex post returns are positive as well. Dropping out may simply mean that at the updated information, continued education is not worthwhile

- school design
  It may be quite beneficial to design school structures with an eye on optimal production of information, with well organised decision nodes and curricula designed to maximise information value in the first stage.

- heterogeneous information value
  As noted above, there is a region of parameters values where individuals only enter schooling if they can avoid loss associated with low-ability levels by quitting after the first stage. One of the determinants of this region is the effect of risk on utility, i.e. risk attitude.

5. Implications for empirical work

Tracing the mean-risk frontier for education and skills is an interesting and relevant topic for empirical work. We should be well aware though that the observed frontier may be truncated by optimal supply policies (under rational school design, inefficient combinations will not be offered) and twisted by individuals' choices: differences in abilities and preferences lead to selective observation of fragments of frontiers. As everywhere in models of occupational choice, the hedonic problem has to be faced.

Empirical work along these lines focussing on skills is virtually absent. However, a literature is developing on the link between earnings risk and given school types. Several papers now consider whether the labour market indeed provides compensation for the financial risk of investing in an education (Hartog and Vijverberg, 2002; Diaz Serrano, Hartog and Nielsen, 2004); a significant compensation has been observed for several countries. Christiansen and Nielsen (2002) estimate mean-variance frontiers for 110 educations in Denmark. Existing educations can generally not be mixed and the frontier is simply the set of extreme points (maximum return for given variance, or minimum variance for given return). They find indeed that on average mean returns increase for increasing risk (the regression line relating return to risk slopes upward). The effect is dominated by academic educations, as for manual educations (apprenticeship programmes) there is no significant relationship. One might of course use their results to construct an efficient frontier by allowing mixes of educations and thus suggest new curricula as an optimum combination of existing educations. Christiansen and Nielsen find a rather irregularly shaped frontier and many observations inside the frontier, i.e. educations dominated by educations that have higher return for the same risk. They rationalise these observations by assuming that tastes for non-pecuniary aspects render these educations attractive investments. This is the empirical counterpart of the correction term (17').

Instead of considering educations, as given curricula, one may indeed look at skills, or aptitudes, or whatever it is that schools produce. But this is mostly a vacant field. Consider again the common argument, quoted from Becker in the introduction, that broad, flexible educations are superior to specific, narrow educations. The argument implies that flexible educations, such as liberal arts, are less risky because they allow shifting from declining, "bad draw" positions to prosperous "good draw" positions and thus restrict the downside risk of the investment. This should show up in lower earnings dispersions of broad, flexible educations as compared to the narrow specific educations. Such a relationship is not a priori evident and hence, an interesting topic for empirical analysis. Again, there is not much of a literature on this topic. Dolton and Vignoles (2002) have tested whether a broader education in secondary school in the UK (more than the minimum number of fields) indeed generates higher earnings, but they find no support for this hypothesis. They have not considered the dispersion of earnings however. There is a broader literature focussing on differences in returns to education by field, but without placing in this the context of broader educations providing better safeguards against the vagaries of the labour market and hence without attention for earnings dispersion. More specifically, Anderson and Pomfret (2002) conclude that in some Central Asian republics, the college educated benefited from the transition to a market economy, while secondary and vocational educations did not provide a significant return. Combining with other evidence, they conclude: " We interpret this result as support for the idea that general purpose education becomes particularly valuable in disequilibrium situations". Indeed, dealing with disequilibrium situations as new opportunities may be a way of restricting the risk involved in such transitions.

There is an interesting agenda waiting. One may consider wage dispersions by type of education, after classifying these educations as broad and flexible or narrow and risky. Such analyses should not be too difficult, once one has classified the educations, which may not be easy. We started an analysis along these lines by considering wage loss and duration to a new job for employees who lost their job in a major company failure (Dohmen, 2002; Kriechel, 2003) A key problem we faced is indeed the problem of classification of training programmes and courses as general versus specific, or risky versus flexible. Actually, it is not even clear how we should precisely define and measure a skill. Economic analyses are mostly restricted to an abstract level, where skills are seldom directly measured. Perhaps, it is fruitful to restrict analysis to a few underlying key skills, such as analytical (mathematical), verbal and social skills, a trichotomy that worked well to understand the wage structure in the US (Hartog, 1980). The US Dictionary of Occupational Titles (DOT) specifies job requirements for thousands of jobs. Combining these job requirements with observations on earnings for the job generates a dataset linking means and variances of earnings with job requirements as units of observation. For example, one may utilize the three dimensions of GED: Math, Reading and Languages, each measured at 6 levels, hence giving 236 cells with earnings distributions from which mean and risk can be calculated. We are presently analysing the data along these lines.

References

Anderson, K.H. and R. Pomfret (2002), Relative living standards in new market economies: evidence form Central Asian household surveys, *Journal of Comparative Economics,* 30 (4), 683-708

Becker, G. (1964), *Human capital*, Chicago: University of Chicago Press (third edition, 1993).

Bodie, Z., A. Kane and A. Marcus (1989), *Investments*, Chicago: Irwin

Christiansen, C. and H. Skyt Nielsen (2002), The educational asset market: a finance perspective on human capital investment, working paper, Arhus Business School

Diaz-Serrano, L., J. Hartog and H. Skyt Nielsen (2004), Compensating wage differentials for schooling risk in Denmark, Maynooth (Ireland)/Amsterdam/Aarhus

Dolton, P.J. and A. Vignoles (2002), Is a broader curriculum better? *Economics of Education Review,* 21 (5), 415-430.

Dohmen, T. (2002), Internal labor markets, Ph.D thesis University of Maastricht

Halek, M. and J. G. Eisenhauer (2001), Demography of risk aversion, *Journal of Risk and Insurance*, 68 (1), 1-24

Hartog, J. (1980), Earnings and capability requirements, *Review of Economics and Statistics*, LXII (2), 230-240

Hartog, J. (1981), *Capabilities, allocation and earnings*: Boston: Kluwer

Hartog, J., A. Ferrer-i-Carbonell and N. Jonker (2002), Linking measured risk aversion to individual characteristics, *Kyklos*, 55 (1), 3-26

Hartog, J. and W. Vijverberg (2002), Are wages really compensating for risk aversion and skewness affection? IZA DP 426.

Kriechel, B. (2003), Heterogeneity among displaced workers, Ph.D thesis, University of Maastricht

Levhari, D. and Y. Weiss (1974), The effect of risk on the investment in human capital, *American Economic Review*, 64 (6), 950-963

Rosen, S. (1983), Specialization and human capital, *Journal of Labor Economics*, 1(1), 43-49

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

Ventura, L. and J. G. Eisenhauer (2002), Survey measures of risk aversion and prudence, working paper University of Rome “La Sapienza” and Canisius College, Buffalo NY

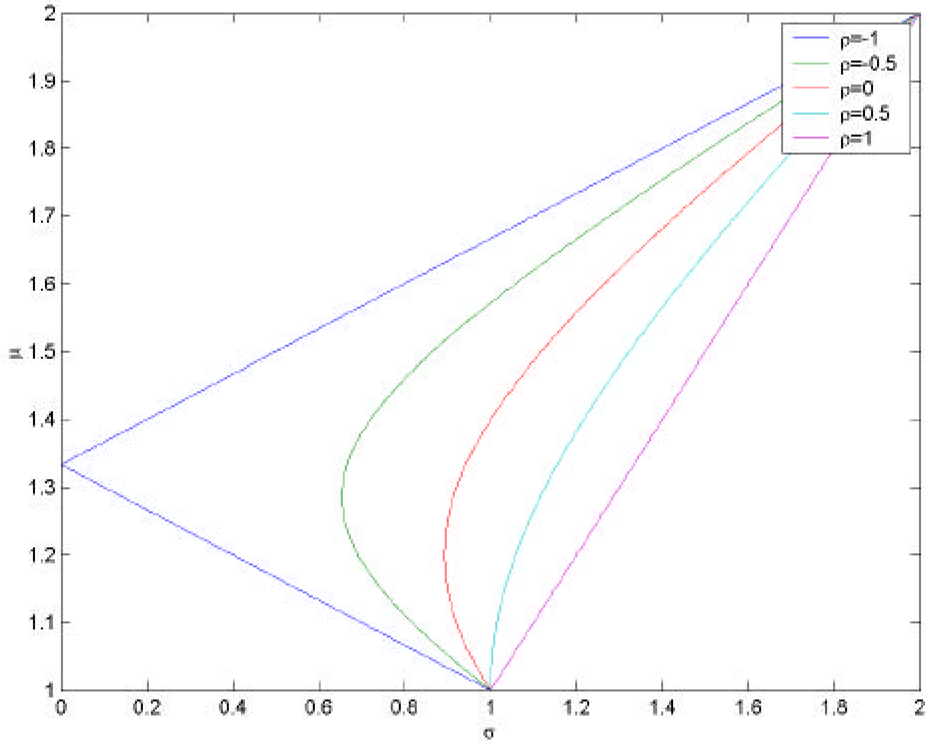

Figure 1. The mean-risk frontier ($\mu_1 = s_1 = 1, \mu_2 = s_2 = 2$) in a linear world.