

Can Risk Aversion Explain Schooling Attainments? Evidence From Italy*

Christian Belzil
Centre de National de Recherche Scientifique,
Institute for the Study of Labor (IZA)
and CIRANO

Marco Leonardi
Universita di Milano and
Institute for the Study of Labor (IZA)

November 22, 2005

Abstract

Using unique Italian panel data, in which individual differences in behavior toward risk are measured from answers to a lottery question, we investigate if (and to what extent) risk aversion can explain differences in schooling attainments. We formulate the schooling decision process as a reduced-form dynamic discrete choice and we analyze how grade transition from one level to the next varies with preference heterogeneity (risk aversion), parental human capital, socioeconomic variables, market skills and persistent unobserved (to the econometrician) heterogeneity. We present evidence that schooling attainments decrease with risk aversion, but despite a statistically significant effect, differences in attitudes toward risk account for a modest portion of the probability of entering higher education. Differences in ability(ies) and in parental human capital are much more important.

JEL Classification: J24.

*We thank David Card and Winfried Koeniger for useful comments.

1 Introduction and Motivation

The connection between individual attitudes toward risk and investment behavior has been widely analyzed in financial economics. This is true both at the theoretical and at the empirical level.¹ Although human capital is undoubtedly the main component of individual assets, the link between risk aversion and human capital accumulation, and in particular schooling, remains largely hypothetical. Most of the work is theoretical and often confined to relatively simple two-period models. In general, the results stress that earnings uncertainty may depress human capital investment.²

Empirical work remains scarce and is rather inconclusive. There is one main reason for this. At the empirical level, determining which asset is more risky is a relatively straightforward econometric question. However, quantifying the marginal risk which characterizes the transition from one level of schooling to the next is a more difficult research agenda. Not surprisingly, economists are currently unable to say if (and to what extent) schooling acquisition is a risky investment although the issue is starting to raise a significant level of interest. Moreover, the degree of education selectivity based on individual differences in risk aversion is completely unknown.

In this paper, we investigate whether risk aversion can explain differences in schooling attainments. We ask three simple questions. Does risk aversion increase or decrease investment in higher education? How does the effect of risk aversion compare with the effects of ability and family human capital? How much of the cross-sectional dispersion in schooling attainments is explained by differences in attitudes toward risk.

In order to answer these questions, we take an approach completely different from what is found in the literature. Using unique Italian panel data, in which an individual specific measure of risk aversion is inferred from an answer to a lottery question, we formulate the schooling decision process as a reduced-form dynamic discrete choice problem (using discrete duration model techniques) and we analyze how grade transition from one level to the next varies with measured risk aversion. In particular, we decompose the probability of entering higher education into four groups of variables; preference heterogeneity (risk aversion), persistent unobserved (to the econometrician) ability heterogeneity, parental human capital (parents' education and occu-

¹See Kocherlakota (1996) for a comprehensive survey.

²This is the case, for instance, in Lehvari and Weiss (1974) and Olson, White and Sheffrin (1979).

pation) and socioeconomic variables (sex, region, age cohort). Using the wage panel component of the data, we can also obtain an estimate of individual market skills (a standard wage fixed effect estimate) and decompose further total heterogeneity into a market skill component and a residual unobserved heterogeneity term.³

Our analysis is based on a sample of Italian individuals. Our methodology therefore relies on the fact that higher education in Italy must be a “reasonably” risky investment. While tuition fees are low in Italy (and typically everywhere in Europe), there is no reason to believe that Italian students face lower psychic costs than do students in other countries.⁴ For the sake of comparison, Italian students face a relatively more incomplete capital market than do US students. Borrowing while in school is practically inexistent in Italy.⁵ The US, on the other hand, has very high tuition rates but also has substantial student loan and fellowship programs. Interestingly, both Italy and the US are characterized by a relatively high level of inequality. Although cross sectional wage dispersion is higher in the US than in Italy, long run (lifetime) inequality is thought to be higher in Italy and, in particular, among the highly educated.⁶ To the extent that the riskiness of the education investment may be at least correlated with the individual’s lifetime inequality, these institutional facts seem to indicate that investing in higher education may be as risky in Italy as in the US.

Aside from its direct contribution to the revived debate on the schooling/risk trade-off, this paper also contributes to the already existing literature on the determinants of schooling attainments. As of now, labor economists have paid a particular attention to the importance of parental human capital and individual abilities (observed or unobserved). This paper adds a new dimension to the analysis of the determinants of schooling; namely the

³However, this is done at the expense of cutting the sample size.

⁴Empirical evidence for the US suggests that differences in psychic costs may be quite important. For instance, the large explanatory power of the individual specific differences in the per-period utility of attending school found in the structural literature is consistent with the existence of strong psychic costs (Keane and Wolpin, 1997, and Belzil and Hansen, 2002). See Heckman, Lochner and Todd (2005) and Belzil (forthcoming) for surveys.

⁵The Italian national statistical office (ISTAT, 2003, Table 1.8) reports that the total number of student loans in Italy in the academic year 1999-2000 was 97.

⁶In a recent paper, Flinn (2002) shows that after taking into account job offer probability while employed and while unemployed and unemployment incidence, lifetime welfare inequality is higher in Italy than in the US. His results are obtained in a search framework with risk-neutral workers.

importance of preference heterogeneity.⁷

In line with most of the theoretical literature, we find that more risk averse individuals tend to get less schooling. However, we find that risk aversion is a far less important determinant of schooling attainment than individual-specific abilities (including measured market skills) and parental human capital (education and occupation). As an illustration, the range of higher education participation probabilities spanned by the 10th-90th percentile of the market skill distribution is ten to twenty times larger than the equivalent range obtained for the risk aversion measure. More specifically, preference heterogeneity account for 10% or less of the cross-sectional dispersion in the probability of entering higher education. Our main findings are quite robust. They are robust to the allowance for alternative measures of risk aversion and market skills. They are also invariant to the allowance for a differentiated effect of risk aversion at different grade levels.

The paper is constructed as follows. In Section 2, we present some background material and review the most important literature. In section 3, we discuss the Bank of Italy Survey of Income and Wealth (SHIW) and provide details about the measure of risk aversion used in our analysis. The econometric model is described in Section 4. Section 5 contains a brief overview of the model specifications. In Section 6, we present evidence that schooling attainments decrease with risk aversion and we compare the effects of risk aversion with the effects of one measure of ability (a wage fixed effect). In the following section, we decompose the total variation in the probability of entering higher education into its main components, preference heterogeneity, ability(ies), parental background, and socio economic attributes. The conclusion is found in Section 8.

2 Background an Relevant Literature

Fundamentally, the marginal risk associated to schooling has two distinct components. One component relates to the accumulation process and is experienced by the individuals at the time schooling decisions are made. The second component relates to post-schooling labor market outcomes and is therefore associated to the (perceived) distribution of random variables

⁷In the structural literature, the term “preference heterogeneity” is often used to refer to differences in taste for schooling and academic abilities (Keane and Wolpin, 1997). In our analysis, these unobserved factors are subsumed in the unobserved heterogeneity term.

which are realized much beyond the time at which schooling is decided.

With respect to the accumulation process, acquiring schooling should be unambiguously viewed as a risky investment. School (and especially college) attendance requires to sacrifice present consumption and to absorb substantial psychic costs in return for future rewards, but successful grade achievement is rarely a certain outcome. For this reason, the probability of losing the investment paid up front cannot be ignored and may act as a strong disincentive.

At the level of labor market outcomes, the argumentation becomes more complicated. In practice, life cycle earnings are affected by random events such as job offers, layoffs, risk sharing agreements between firms and workers (or unions) and many other events. Occupation choices may also affect earnings volatility. The ex-ante probability distribution of those labor market outcomes may depend on schooling but it is far from clear if accumulated schooling contributes to an increase in earnings dispersion or decreases volatility.⁸ On top of this, wages and earnings are typically affected by individual ability heterogeneity as well as by idiosyncratic and business cycle shocks. Separating these components may be particularly difficult (Cunha, Heckman, and Navarro, 2005).

In the long run, labor market productivity and earnings may be affected by structural changes in the economy. Potential technological changes affecting the return to schooling may be viewed as an additional element of risk from the perspective of the student. On the other hand, when schooling is viewed as facilitating adjustment to technological change, this uncertainty may turn out to favor schooling acquisition (i.e.: schooling becomes a form of insurance).⁹

Given this level of complexity, and taking into account both the accumulation process and labor market outcomes, it is difficult to say whether or not individuals perceive schooling acquisition as a truly risky investment. In the earlier literature, a few descriptive analyses of the variability of empirical age/earnings profile have been carried out. However, the notion of variability is usually an “ex post” notion which may have little to do with

⁸For instance, schooling may reduce earnings dispersion by reducing the unemployment incidence or by raising the job offer probabilities (given unemployment) but it may increase wage volatility if more educated workers find jobs in sectors or occupations where wages (or marginal product) is more volatile.

⁹This argument is put forward in Gould, Moav and Weinberg (2001).

“ex ante” risk.¹⁰ Ideally, evaluating the marginal risk would require a statistical analysis of the joint distribution of life cycle wages, unemployment, job offer probabilities and grade completion (or failure) probabilities. In particular, it would also require to disentangle persistent unobserved (from the econometrician perspective) heterogeneity from true dispersion. This would be difficult to achieve and indeed, as of now, such a comprehensive study does not exist.

On top of this, measuring the marginal risk associated to schooling for all relevant labor market outcomes may turn out to be irrelevant if individuals have imperfect information about the law of motion that generates labor market outcomes. If so, individual subjective probabilities may diverge from the Rational Expectation hypothesis and the use of post-schooling panel data on wages and employment outcomes may become irrelevant for the econometrician.¹¹

As it stands now, there is no strong empirical evidence on the effect of education on wage/earnings dispersion, but economists are starting to pay more and more attention to the issue. In a recent paper, Palacios-Huerta (2003) presents an empirical comparison of the properties of risk-adjusted rates of return to schooling within an intertemporal model, using mean-variance spanning techniques.¹² In line with the stream of the literature devoted to the increase in wage inequality, many individuals have analyzed the wage dispersion (basically the variance) within education groups in cross-section data rather than in panel data. The cross-section evidence shows that the variance of wages is higher within the educated group (Lemieux, 2005 and Chay and Lee, 2000). In an attempt to separate individual heterogeneity from ex-ante risk, Belzil and Hansen (2004) estimate a dynamic programming model in which the degree of risk aversion can be inferred from schooling decisions but they assume that the attitude toward risk is represented by a parametric (constant relative risk aversion) utility function. They identify the degree of risk aversion from the degree of heteroskedasticity in the idiosyncratic earn-

¹⁰Mincer (1974) investigates how the variance of earnings differs across schooling levels over the life cycle while Chiswick and Mincer (1972) use age earnings profile to investigate time series changes in income inequality. Kodde (1985) uses the Lehari Weiss model as a background for empirical work and tests predictions from the model from data on subjective estimates (self reported) of future earnings.

¹¹See Geweke and Keane (2001) for an insightful discussion.

¹²Basically, the mean-variance spanning technique amounts to quantifying the effect of introducing a new asset on the mean-variance of another benchmark asset.

ings shock but assume that all persistent unobserved heterogeneity is in the information set of the agent. As panel data on wages, earnings and schooling do not allow them to identify cross-sectional dispersion in risk aversion, they assume homogeneity of preferences and automatically rule out the possibility that differences in schooling are driven by differences in attitudes toward risk. Finally, Cunha, Heckman and Navarro (2005) develop a statistical method which distinguishes between heterogeneity and risk but also allow for a distinction between ex-ante risk and ex-post dispersion. Their method allows the econometrician to infer the set of variables upon which schooling decisions are based, but disregards heterogeneity in risk aversion.¹³

3 Measuring Risk Aversion: The Bank of Italy Survey of Income and Wealth

We use data from the 1995 wave of the Bank of Italy Survey of Income and Wealth (SHIW). The survey collects information on consumption, income and wealth in addition to several household characteristics for a representative sample of 8,135 Italian households. More importantly, the 1995 survey contains a question on household willingness to pay for a lottery which can be used to build a measure of individual risk attitudes.

In the survey, each head of household is asked to report the maximum price he/she is willing to pay to participate to an hypothetical lottery. The question is worded as follows:

“We would now like to ask you a hypothetical question that we would like you to answer as if the situation was a real one. You are offered the opportunity of acquiring a security permitting you, with the same probability, either to gain a net amount of 10 million lire (roughly 5,000 dollars) or to lose all the capital invested. What is the most you are prepared to pay for this security?”¹⁴

The respondent can answer in three possible ways: 1) give the maximum price he/she is willing to pay, which we denote as *bet*; 2) don't know; 3) don't want to participate. Of the 8,135 heads of household, 3,458 answered they

¹³On top of these few papers, a relatively large number of related working papers are being currently circulated. These include Hartog, Van Ophem and Bajdechi (2004), Chen (2003), Harmon, Hogan and Walker (2003) and Davis and Willen (2002).

¹⁴In other words, the expected value of entering the lottery is $0.5 \cdot (10,000,000 - bet)$.

were willing to participate and reported a positive maximum price they were willing to bet (prices equal to zero are not considered a valid response).

The valid responses to the question - *bet* - range from 1,000 lire to 100 million lire and constitute our measure of individual risk aversion. Of the 3,288 heads in our final data set, 3,131 reported a maximum price *bet* less than 10 million lire which implies that they are risk averse individuals, 117 reported *bet* exactly equal to 10 or 5 million lire (i.e. they are risk neutral) and 40 reported *bet* more than 10 million indicating that they are risk lovers. The empirical distribution of *bet* is reported in Table 2.

It should be noted that this measure of risk requires no assumption on the form of the individual utility function and extends to risk-averse, risk-neutral and risk-loving individuals.¹⁵ This lottery question has been used to study the relationship between risk aversion and several household decisions but has never been used to explain schooling attainment.¹⁶

Theoretically, the answer given by the individual may be partly affected by his/her time invariant degree of risk aversion but also partly affected by time varying differences in wealth/income endowment. Guiso and Paiella (2001) show that household income and wealth and individual characteristics have limited explanatory power.¹⁷ Ultimately, they conclude that this measure of risk is a good proxy for the time invariant individual specific component of the attitude toward risk.

In order to verify this claim and obtain a measure of risk aversion which is orthogonal to earnings or wealth, we regress the response to the lottery question, *bet*, on household wealth, household income and a dummy variable *Owner* which indicates whether the head is a home-owner.¹⁸ The measure of household wealth *Hwealth* includes all financial assets held by the household in 1995. The measure of household net income, *Hincome*, includes earnings, pensions and income from real and financial capital. The descriptive statistics of these variables are also in Table 1 in the Appendix. The results of the regression (not shown) indicate that the amount of the bet, *bet*, is positively

¹⁵It should also be noted that, given the answer to the lottery, it is possible to construct a measure of the Arrow-Pratt index of absolute risk aversion. This requires taking a second-order Taylor expansion of the relevant expected utility (Gollier, 2001).

¹⁶Guiso and Paiella (2004), use the question on risk aversion to analyze occupation choice, portfolio selection and migration decisions. Brunello (2002) estimates returns to schooling instrumenting schooling attainment with risk aversion.

¹⁷Interestingly, the main predictor of risk aversion is region of birth.

¹⁸This is also done in Brunello (2002).

related to household income and to household wealth and it is insignificantly related to home-ownership. However, the three variables account for only 2% of the total variance of *bet*. The residual of this regression is variable *betR* in Table 1.

Guiso and Paiella (2001) discuss in details the main advantages of this estimate of absolute risk aversion relative to those already in the literature. They underline that the lottery represents a relatively large risk. In fact, ten million lire corresponds to just over 5,000 dollars and the ratio of the expected gain of the hypothetical lottery to the annual average Italian household consumption is 16 percent. This is considered an advantage since expected utility maximizers may behave as risk neutral individuals with respect to small risks even if they are risk-averse to larger risks. Thus, facing consumers with a relatively large lottery may be a good strategy to elicit risk attitudes.

Apart from the lottery question, we use information on the level of education attained by the head of household, as well as variables such as age, gender, region of birth, parental education and parental occupation. This set of variables is comparable to those which are used in US studies based on the National longitudinal survey (NLS). We select the sample of all heads with a valid answer to the lottery question (3,458) and eliminate those who report a missing value in any of the following variables: education, age, gender, region of birth, education and occupation of the head's father and mother. This selection process leaves us with a final sample of 3,288 heads of household.

The schooling variable takes values for 1 to 6 corresponding to no education, elementary school (typically attained at 11 years of age), junior high school (attained at 14), high school (attained at 18), university degree (attained at 23-24) and post-university degree.

In order to condition on a measure of market ability, we use the residual of a fixed effect regression of log monthly wages on age and age squared. This measure will turn out to be particularly useful when we want to perform a direct comparison of the effects of risk aversion and individual abilities. The SHIW survey has a rotating panel structure and of the 8,135 households interviewed in 1995, 3,645 (45%) belong to the panel section. Of the 3,288 heads selected in the 1995 wave, only 1,822 can be linked to at least one more wave before or after 1995 (The survey panel years are 1989, 1991, 1993, 1995, 1998, 2000, 2002). Therefore a measure of market skills, denoted *msk*, is available for 1,822 heads.

Table 1 in the Appendix shows the descriptive statistics of the sample. In the estimation we use dummy variables derived from the original variables.

There are six dummy variables - *edu1* to *edu6* - for the level of education attained by the individual (no title, elementary school, junior high school, high school, university degree, post-college degree), three dummies - *north*, *centre* and *south* - for the region of birth, one age dummy (*45more* = 1 if age of head more than 45 in 1995) and one sex dummy (*female* = 1).¹⁹ In addition we have one dummy each - *eduf* and *edum* - respectively for the level of education attained by the individual's father and mother (less than high school=0, high school or more=1), and four occupation dummies for blue collar, white collar, self employed and unoccupied for parents' occupation. These variables are denoted *bcf*, *wcf*, *sef*, *uf* for the father and *bcm*, *wcm*, *sem*, *um* for the mother.

4 The Econometric Model

In this section, we present the econometric model. As schooling attainments are reported according to six (ordered) levels, we model schooling decisions with a reduced-form dynamic discrete choice model and we use a hazard function model of grade transition. The grade transition model admits a semi-structural interpretation and may be regarded as an approximation to a sequential dynamic optimization model. While non-parametric identification results are typically known for proportional hazard functions, the estimation of non-proportional hazards with person specific unobserved factors is also quite common when modeling discrete durations.²⁰ Indeed, in recent work, Heckman and Navarro (2005) show that it is also possible to conceive non-parametric (or semi-parametric) identification of reduced-form dynamic discrete choice model under certain conditions.²¹

Our strategy is to model the grade transition function as flexibly as possible. It is important to note that an alternative would be to model schooling as an ordered static discrete choice (say an ordered logit or an ordered probit)

¹⁹The reason we introduce a dummy for the heads of household older than 45 in 1995 is that presumably they started college (if they ever attended it) before 1968. Before 1968 legal restrictions limited the accession to college only to those who had a high school degree in classical or scientific studies, since 1968 accession to college is open to any type of high school degree.

²⁰Discrete (logistic) hazards are used in Baker and Melino (2000) and in Ham and Lalonde (1997).

²¹Heckman and Navarro (2005) also show that similar identification results extend to structural dynamic programming models.

in which the error term may be viewed solely as cross sectional differences in tastes or abilities. However, in such an ordered model, the error term does not allow for randomness in the schooling decisions from the perspective of the agent and should be interpreted solely as randomness from the perspective of the researcher. The ordered probability model specification is usually interpreted as a static (or myopic) model of schooling decisions.²² For these reasons, we favor the hazard specification, although we also present some results obtained with the ordered logit.

We see at least three main advantages to our approach. First, it does not require to specify individual preferences but only requires that the measure of preference heterogeneity is a good proxy for the ordering of the persistent degree of risk aversion across individuals. Indeed, modeling schooling decisions in a discrete duration framework imposes only that choices be made sequentially.

Secondly, it neither requires to model the distribution of labor market outcomes, nor to assume that the distribution of the labor market outcome variables, which are realized in the post-schooling periods, is actually known by the agents at the time of the schooling decisions.

Finally, we do not need to assume that the persistent unobserved (to the econometrician) heterogeneity term(s) affecting labor market outcomes belong(s) to the information set of the agent when schooling decisions were made. Our estimation strategy is therefore consistent with schooling decisions made under imperfect information about individual specific skills.²³

The model allows for different types of factors; measured preference heterogeneity, family characteristics (parents' education and occupation), gender, regional effects, cohort effects and, finally, persistent individual unobserved heterogeneity. In one model specification, we also use an estimate of individual market skills and decompose further total heterogeneity into a market skill component and residual unobserved heterogeneity.

With six (ordered) levels of schooling, we are able to estimate five different hazard rates. The conditional probability of stopping at grade g of individual i (a hazard rate), denoted H_g , is simply

$$H_{g,i} = \Lambda(U_{g,i}) = \frac{\exp(U_{g,i})}{1 + \exp(U_{g,i})} \text{ for } g = 1, 2, \dots, 5 \quad (1)$$

²²This is discussed in Cameron and Heckman (1998).

²³This issue is analyzed formally in Cunha, Heckman and Navarro (2005).

where

$$U_{g,i} = \alpha_{g,i} + \beta'_g X_i \quad (2)$$

The term $\alpha_{g,i}$ represents an individual/grade specific intercept term, X_i is a vector of observable characteristics, and β'_g represents a grade specific vector of parameters measuring the effects of these characteristics. We assume that

$$\alpha_{g,i} = \alpha_g + \theta_i$$

and that θ_i is drawn from an unknown distribution which is approximated by a discrete distribution with K points of support.²⁴ This approach amounts to the estimation of a vector of grade level specific intercept terms for each type. As we include an intercept term in the transition probability, we normalize one support point (namely θ_1) to 0.²⁵

Given the form of the hazard specification (equation 1), it is important to note that the sign of the parameter estimates indicates the direction of the effect of a variable on the exit rate out of school. So a negative estimate will typically imply a positive effect on expected grade completion, and in particular, on the probability of reaching higher education.

We estimate the model by maximum (mixed) likelihood techniques. Defining six different schooling indicators from the lowest schooling level (d_{1i}) to the highest (d_{6i}), the contribution to the likelihood for an individual i who has completed level g , is denoted L_i , and is equal to

$$L_i = \sum_{k=1}^K p_k \cdot [\prod_{s=1}^{g-1} (1 - H_{s,i}(X_i, \theta_k))^s \cdot H_{g,i}(X_i, \theta_k)] \quad (3)$$

where θ_k is the type specific support point and where the type probability, p_k , is specified as $\frac{\exp(p_{0k})}{1 + \exp(p_{0k})}$.

²⁴As typically found in most empirical applications dealing with a univariate duration endogenous variable, it has been found that $K = 2$ is sufficient to characterize unobserved heterogeneity.

²⁵Obviously, the probability of transiting from one grade level (g) to the next ($g + 1$), the grade transition (continuation) probability, is simply

$$1 - H_{g,i} = \frac{1}{1 + \exp(U_{g,i})}$$

We sometimes refer to the grade transition probability as a “continuation” probability. The hazard rate is sometimes referred to as a “termination” probability.

Defining g^h as the number of years of schooling required to attend higher (post-secondary) education and letting S_i as the number of years of schooling completed as of 1995, the probability of attending post-secondary education, denoted $\tilde{P}_i(X_i, \theta_i)$, is given by

$$\tilde{P}_i(X_i, \theta_i) = \Pr(S_i \geq g^h \mid X_i, \theta_i) = \prod_{s=0}^{g^h} (1 - H_{s,i}(X_i, \theta_i)) \quad (4)$$

Given estimates of the type specific population proportions and support points (the p_k 's and the θ_k 's), the individual specific higher education attendance probability, $\tilde{P}_i(X_i, \theta_i)$, can easily be evaluated. To perform the variance decomposition, we choose the individual specific log odds ratio, $G(X_i, \theta_i)$, which is

$$G(X_i, \theta_i) = \left[\ln \left(\frac{\tilde{P}_i(X_i, \theta_i)}{1 - \tilde{P}_i(X_i, \theta_i)} \right) \right] \quad (5)$$

where $G(X_i, \theta_i)$ is treated as an unknown regression function to be estimated. In Section 6, we decompose the total variation in $G(X_i, \theta_i)$ into the different contributions of some specific group of variables; preference heterogeneity, family characteristics (parents' education and occupation), socioeconomic variables (gender, regional effects, cohort effects) and, finally, persistent individual unobserved ability heterogeneity). This approach allows us to obtain a ranking of the various groups of variables and, in particular, to establish if (and to what extent) differences in risk aversion are important.

5 An Overview of the Different Model Specifications

In order to obtain a clear picture of the effect of risk aversion on school attendance, we first estimate a simple version of the grade transition model with 5 grade specific intercept terms and with a common set of parameters assumed to be constant across all grade levels (where $\beta'_g = \beta \forall g$). We refer to it as Model 1. This model was estimated on the full sample first and then re-estimated on a smaller sample which contains information on individual market skills. The results are found in Table 3. In order to illustrate the robustness of the results, this specification was also re-estimated with an alternative measure of risk aversion and as well as discretized measure of market skills (Table 9).

As a second step, we estimated a more flexible version of the model where the regressors have separated effects by grade level. The specification reported herein (in Table 4) allows for the effect of the regressors to change at the level prior to college enrolment. In other words, it allows for the marginal effect of risk aversion heterogeneity to change as an individual who has completed the 4th level is deciding to continue to the 5th level (corresponding to higher education). In the second portion of Table 4, we report estimates of the specification with market skills and, in order to achieve the highest level of flexibility, we allow the risk aversion and the market skill variable to enter non-linearly as well as interactively.²⁶

6 How Does Risk Aversion Affect Grade Progression?

In this section, we concentrate on the discussion of the parameter estimates of the grade transition model. Basically, and in line with the theoretical literature, we present evidence that schooling attainments decrease with risk aversion and we evaluate the robustness of the results to alternative measures of risk aversion and market skills. We report that the effect of risk aversion is not magnified at higher grade levels. Finally, we stress that the range of higher education participation probabilities spanned by the 10th-90th percentile range of the market skill distribution is ten to twenty times larger than the equivalent range for the risk aversion measure.

6.1 Estimates from a Simple Model

The results obtained for the simplest model specification are found in Table 3. The set of parameter estimates and asymptotic standard errors found in the first 2 columns corresponds to the case where we used the largest sample and therefore ignored the market skill measure obtained from a fixed effects wage regression.

The gradual increase in the intercept terms (α_1 to α_5) from level 1 (-3.94) to level 5 (9.30) indicates that the schooling decision process is characterized by an increasing termination rate (given unobserved heterogene-

²⁶The estimations are performed using a FORTRAN program. However, they are easily doable using standard econometric softwares such as SAS or STATA.

ity). The importance of unobserved heterogeneity is readily seen from the support point estimate for type 2 individuals (-4.52) along with the type 1 probability equal to 0.75. This implies that the population is clearly split between a high schooling attainment (low hazard rate) sub-population made of type 2 individuals, and a lower schooling attainment (higher hazard rate) sub-population made of type 1 individuals. The importance of unobserved persistent heterogeneity is a well known feature of most studies based on US data. A large number of studies set in a dynamic framework point out that permanent unobserved heterogeneity, which may represent unobservable factors such as individual specific taste for schooling, academic ability, motivation, or any other unobservable trait which is time-invariant, is indeed the major determinant.²⁷

As documented in many empirical studies, grade termination is lower for those whose parents have achieved higher education. The parameter estimates for father's and mother's education are equal to -2.47 and -1.48 respectively. This positive correlation between individual schooling attainments and parents education is well established in simple correlation analysis (Kane, 1994), in reduced-form dynamic models such as Cameron and Heckman (1998, 2001) as well as in structural dynamic programming models such as, Eckstein and Wolpin (1999) and Belzil and Hansen (2002). Grade continuation is also higher for those who have a parent who worked in a white collar occupation (the omitted category). This is readily seen upon looking at the positive effect (on the hazard rate) of the binary variables for all other occupation types. As expected, we find that individuals living in the North (the most economically developed region of Italy), when compared to those who live in central regions, obtain more schooling. Finally, both females and younger cohorts appear to have lower grade termination rates. However, given the objectives of the paper, these estimates do not raise immediate interest and we do not discuss them in detail.

The parameter that raises most interest is the effect of preference heterogeneity as measured by the individual specific value attached to the lottery. Given unobserved heterogeneity and other measurable characteristics, we find that those individuals who attach a higher value to the lottery (those who are less risk averse) tend to have a lower grade termination rate. The estimate, equal to -0.565, is highly significant, and therefore indicates that more risk

²⁷Cameron and Heckman (1998 and 2001), Eckstein and Wolpin (1999) and Belzil and Hansen (2002 and 2003).

averse individual obtain less schooling. This is in agreement with conventional wisdom. However, we do not know of any comparable results, where the degree of schooling selectivity is directly tied to an observable measure of risk aversion, in the empirical literature.²⁸

It is also interesting to note that our results appear to conflict with those reported in a somewhat related literature that uses smoking behavior as an instrument for schooling, in order to estimate the return to schooling. The first stage regressions often indicate that schooling is inversely related to smoking behavior and it is customary to interpret this finding as evidence that risk averse individuals (those who smoke less) obtain more schooling. However, it should be noted that smoking is an endogenous variable, which is likely to be affected by several factors including intrinsic taste for smoking, parents' background (including education), teenage schooling attainments (performance in school), and other individual specific factors such as risk aversion and discount rates. It is therefore not certain that changes in schooling induced by smoking differences are solely due to risk aversion and, therefore, that standard assumptions (such as monotonicity and homogeneity) made in the IV literature would be valid in this context. Our results, on the other hand, illustrate a marginal effect of risk aversion, holding all other factors constant. They are certainly not incompatible with the hypothesis that young individuals coming from poorer background and less educated families tend to smoke more (given a fixed degree of risk aversion).

As noted earlier, using life cycle wage data, it is possible to obtain an estimate of individual market skills. As market skills are known to be highly correlated with academic skills, the introduction of this variable should reduce the impact of pure unobserved heterogeneity.²⁹ The results found in the last portion of Table 3 indicate again that risk aversion reduces schooling. The parameter estimates goes from -0.565 to -0.278 but remains very significant. Moreover the related marginal effect, which will be illustrated later, remains comparable. At the same time, the strong negative estimate

²⁸For instance, in the theoretical model of Lehvani and Weiss, this relationship would be derived from differentiating the expected utility of staying in school with respect to a measure of concavity of the utility function. In Belzil and Hansen (2004), the effect of risk aversion on schooling is obtained upon differentiating the school attendance probability (involving a closed-form solution to the value function) with respect to a parameter representing absolute (or relative) risk aversion.

²⁹As noted in Section 2, there seems to be virtually no correlation between risk aversion and the individual wage fixed effect.

for the effect of market skill, -3.609, implies that those with higher market skills have lower (higher) termination (continuation) rates. This is consistent with the standard view that market skills and academic skills (or taste for schooling) are strongly and positively correlated. This does not require further discussion.

Despite the introduction of observed market skills, unobserved heterogeneity remains important. Type 1 individuals (corresponding to 34% of the population) now constitute the sub-population of more educated individuals as the support point estimate for type 2 individuals is positive (3.97). Overall, the other estimates appear quite robust. Indeed, the only significant sign reversal observed in the results pertains to the male/female difference (the estimate goes from -0.58 to 1.43). However, this is easily explained by the fact that the wage fixed effect is estimated on a sample which pools male and females and is therefore not orthogonal to gender indicator.

Finally, we have also estimated an ordered logit version of the schooling model. The estimates are presented in Table 8 for both the full sample and the sample with market skills. The dependent variable of the ordered logit is education attainment which takes values 1 to 6 (1 for no formal education to 6 for post-college education). Therefore a positive coefficient indicates that the variable is associated with an increased probability of higher education. Consistent with the results which have been reported earlier for the discrete duration model, more risk averse individuals get less schooling and individuals with high market skills and with high-school educated parents get more schooling. As in the discrete duration model results, the change in the predicted probability of attaining a given level of education is much larger if we go from the tenth (-0.409) to the ninetieth (0.443) percentile of the market skill distribution rather than if we go from the tenth to the ninetieth percentile of the bet distribution (50 and 5000 respectively). For example the predicted probability of attaining at least high school (outcome=4) at the bottom ten decile of the market skill distribution (assuming all other regressors equal zero) is $\frac{\exp(1.476 - 2.64 * 0.409)}{(1 + \exp(1.476 - 2.64 * 0.409))} = 0.59$, at the top ten decile is 0.93. The same probabilities calculated at the bottom and at the top tenth decile of the bet distribution are respectively 0.813 and 0.827. Clearly skill differences have a much larger marginal effect than preferences towards risk.

6.2 Some Robustness Analysis

To evaluate the robustness of the results, we investigate two separate issues. First, we also considered an alternative measure of the monetary value associated to the lottery described above. In order to take into account that individual wealth and income may affect substantially the degree of risk aversion measured by the answer to the lottery, we re-estimated the first model specification of Table 3. If the monetary value is truly affected by wealth, then a measure of risk aversion purged of wealth effects may turn out to be closer to a truly individual specific (time invariant) measure of risk aversion. However, using the residual of the regression of the monetary value of the individual bet on various measures of wealth will change the scaling of the variables.

The estimates, found in Table 9, indicate clearly that the results discussed earlier are indeed robust to the introduction of this alternative measure of risk aversion. The effect of the individual specific monetary value is equal to -0.28 and implies again that risk aversion is negatively correlated with schooling attainment. Most of the remaining parameters have practically not changed.

In order to deal with potential measurement (estimation) error imputed to our measure of market skills, we actually used a binary transformation of the fixed effect estimates. We ran the specification of Table 3 with a binary indicator equal to 1 for those who had a positive value for the wage fixed effect and 0 for those who were below. This transformed variable is less subject to measurement error although it is not as precise as a continuous indicator. The results reported in Table 9, also suggest a fair degree of robustness. The estimate for market skill, -1.74 (0.10), is consistent with a positive correlation between schooling attainment and market skills.

6.3 Does the Effect of Risk Aversion Changes with Grade Level?

We now examine the estimates obtained from the more flexible specifications (Model 2) found in Table 4. It allows for a common set of parameters from level 1 to level 3 and a different set common at level 4 and level 5. This specification enables the effect of risk aversion to change at the grade level where the strategic decision to enter higher education is made. Ultimately, we use the implied marginal effects and the variance decompositions to establish

a ranking between the group of variables in terms of their importance in explaining schooling attainments. As will become clear later, we show that, despite a statistically significant effect, differences in attitudes toward risk are virtually unimportant. Unobserved persistent factors, market skills and family human capital play a substantially larger role.³⁰

The estimates of Table 4 obtained on the full sample suggest that risk aversion decreases grade continuation both at the low and the higher grade levels. The estimates, -0.824 (level 1 to 3) and -0.141 (level 4 and beyond), are both significant at any confidence level. In order to fix ideas, it is useful to compute average grade termination (hazard) rates over the relevant range of the risk aversion heterogeneity variable. To get a clear picture, we report the average hazard rates at various percentiles of the risk aversion distribution; namely the 10th, the 50th and the 90th percentiles. These estimates are found in the left-hand side of Table 5. They illustrate clearly the weak effect of risk aversion on grade termination. As an example, the average probability of terminating at grade level 4 fluctuates between 0.92 for someone endowed with a rather extremely low value for the risk aversion indicator (a bet which is ranked at the 10th percentile) and 0.905 for someone at the 90th percentile. The equivalent ranges are slightly higher for level 3 and level 2, but they remain small. Overall, there is no evidence that the effect of risk aversion is magnified as one approaches the decision to enter higher education.

6.4 How do Differences in Risk Aversion Compare with Differences in Ability?

For a sake of comparison, we perform a similar exercise with the market skill indicator. This will allow us to obtain a relative measure of the importance of preference heterogeneity as opposed to ability heterogeneity. It should be noted that, although the wage fixed effects are imperfect estimates of the true market abilities, they provide us with a scaling measure of the ability variable.

The hazard probabilities are computed from the estimates reported in the second portion of Table 4 and they are found in the right-hand side portion

³⁰We have also estimated a more general model with the effect of the regressors changing at each possible grade level. The basic results remaining the same, and for transparency purposes, we decided to report the results of the specification involving a smaller number of parameters.

of Table 5. There is a huge difference between the relative importance of market skills and risk aversion. The 10th-90th decile difference in hazard probabilities is equal to 0.25 (0.91 - 0.66) at level 4 and 0.45 (0.76 - 0.31) at level 3. Depending on grade level, these differences are 10 to 20 times as large as their corresponding value for the effect of preference heterogeneity.

Schooling decisions appear to be overwhelmingly dominated by skill differences as opposed to differences in attitudes toward risk. At this stage, there is clear evidence that schooling attainments are much more affected by differences in ability than by differences in attitude toward risk.

7 What Fraction of Schooling Attainments is Explained by Risk Aversion?

In order to obtain a more global picture of the importance of risk aversion, we now proceed with a precise variance decomposition. Using the estimates of both Table 3 and Table 4, we compute individual specific probabilities of achieving higher education and transform them into the logarithm of their odd's ratios. These probabilities differ across individuals according to the various groups of variables. Using regression techniques, it is easy to compute how much of the individual differences in odd's ratios are explained by each component separately. In what follows, we show that risk aversion is by far the least important of the group of factors on which we focussed our analysis.

The results are found in Table 6 (for Model 1) and Table 7 (for Model 2). Note that for each of these cases, we have a variance decomposition that ignores the distinction between market skills and other unobserved factors as well as one specification that incorporates it as an observed regressor. We analyze each of these two cases separately.

When differences in market skills are ignored, we find that unobserved persistent differences account for 74% (in Model 1) of the cross-sectional differences in the probability of entering higher education (Table 6) and 79% in Model 2 (Table 7). When the estimate for market skill is introduced, unobserved heterogeneity remains the most important variable and the corresponding values are 40% (Table 6) and 37% (Table 7). However, differences in market skills become the second most important factor (with 28% in Table 6 and 29% in Table 7).

Interestingly, parental human capital accounts for 20% of the explained

differences in virtually all cases. If we group unobserved skills and observed market skills together and label it as “abilities”, parental human capital becomes, by far, the second most important group of factors. Consistent with the very small marginal effects already noticed, the portions of the total variances explained by differences in risk aversion are very small, they vary between 2% (in cases with no market skill estimate) and 12% (with market skill).³¹ The ranking appears quite robust. Aside from the residual socio-economic factors (sex, region and cohort), risk aversion is the least important factor among those on which we focussed our analysis.

8 Concluding Remarks

In this paper, we present evidence that schooling attainments decrease with risk aversion. This result is in line with the early theoretical literature. However, we show that, despite a statistically significant effect, differences in attitudes toward risk are not that important. Unobserved persistent factors, market skills and family human capital play a substantially larger role.

While interesting, these answers deserve some interpretation and also raise one fundamental question; Why is the level of risk associated to schooling, as perceived by individuals?

One possible answer is that despite the intrinsic risk faced by those who decide to enter higher education, workers may have the perception that schooling reduces wage (or earnings) dispersion. In other words, young individuals regard schooling as an insurance and the marginal risk associated to higher grade enrollment is small. If this is true, it would be interesting to see if this is specific to Italy only or, if this is may extend to other countries.

There is another possible answer. It is conceivable that entering higher education may preserve the option value of choosing occupations, sectors or jobs (firms) which are characterized by stable and safe earnings profiles. In other words, the relevant decisions that involve differences in attitudes toward risk are occupation and/or sectoral employment choices. If these choices are made posterior to the decision to enter higher education, schooling decisions, as such, will not reveal selectivity based on differences in risk aversion.

³¹One alternative method would be to compute a regression on all components and evaluate the loss in explanantory power when each group of variable is removed individually. This method leads to an identical ranking.

While we believe that the analysis presented in this paper is interesting in its own right, we recognize that answering these questions would be important. However, it would require a more sophisticated analysis and access to similar data from other countries. This may be an interesting, but challenging, avenue for future research.

References

- [1] Baker, M. and Angelo Melino (2000) "Duration Dependence and Non-parametric Heterogeneity: a Monte Carlo Study" *Journal of Econometrics*, 96(2), 357-393.
- [2] Belzil, Christian (2005) "The Return to Schooling in the Structural Literature: A Survey" forthcoming in *The European Economic Review*.
- [3] Belzil, Christian and Hansen, Jörgen (2004) "Risk Aversion, Education and Earnings Dispersion" *Research in Labor Economics*, volume 23.
- [4] Belzil, C. and Jörgen Hansen (2003) "Structural Estimates of the Intergenerational Education Correlation" *Journal of Applied Econometrics*, 8, 679-696.
- [5] Belzil, C. and Jörgen Hansen (2002) "Unobserved Ability and the Return to Schooling" *Econometrica*, 70(5), 2075-2091.
- [6] Ben-Porath, Yoram (1967) "The Production of Human Capital and the Life Cycle of Earnings" *Journal of Political Economy*, 75(4), 352-365.
- [7] Brunello, G. (2002) "Absolute Risk Aversion and the Returns to Education" *The Economics of Education Review*, 21, 635-640.
- [8] Cameron, S. and James Heckman (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, 106 (2), 262-333.
- [9] Cameron, S. and James Heckman (2001) "The Dynamics of Educational Attainment of Black, Hispanic and White Males" *Journal of Political Economy*, 109(5), 455-499.

- [10] Cameron, S. and Chris Taber (2001) "Estimation of Educational Borrowing Constraints using Returns to Schooling" Working Paper, Northwestern University.
- [11] Chay, K. and David Lee (2000) "Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials", *Journal of Econometrics*, 99, 1-38.
- [12] Chen, Stacey H. (2003) "Estimating Wage Volatilities for College versus High School Careers" Working paper, SUNY Albany.
- [13] Chiswick, B.R. and Jacob Mincer (1972) "Time-series changes in Personal income inequality in the United states from 1939, with Projection 1985" *Journal of Political Economy*, 80(3), S34-S66.
- [14] Cunha, F., J. Heckman and Salvador Navarro (2005) "Separating Uncertainty from Heterogeneity in Life Cycle Earnings" NBER Working Paper 11024.
- [15] Davis, S. and David Willen (2000) "Risky Labor Income and Portfolio Choice" Working Paper, University of Chicago
- [16] Eckstein, Zvi and Kenneth Wolpin (1999) "Why Youths Drop Out of High School: the Impact of Preferences, Opportunities and Abilities" *Econometrica*, 67(6), 1295-1339
- [17] Fersterer, J. and Rudolf Winter-Ebmer (2003) "Smoking, Discount Rates, and Returns to Education" *Economics of Education Review*, 22(6), 561-566.
- [18] Heckman, J., Lochner, L. and Petra Todd (2005) "Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond" NBER Working Paper 11544.
- [19] Heckman, J. and Salvador Navarro (2005) "Dynamic Discrete Choices and Dynamic Treatment Effects" forthcoming in *Journal of Econometrics*.
- [20] Heckman, J. and Ed Vytlacil (1998) "Instrumental Variables Methods for the Correlated Random Coefficient Model" *Journal of Human Resources*, 33 (4), 974-987.

- [21] Flinn, Christopher (2002) "Labour Market Structure and Inequality: A Comparison of Italy and the U.S." *Review of Economic Studies*, 69(3), 611-645.
- [22] Geweke, J. and Michael Keane (1995) "Bayesian Inference for Dynamic Discrete Choice Models without the Need for Dynamic Programming" Working Paper, Federal reserve Bank of Minneapolis.
- [23] Gollier, Christian (2001), *The Economics of Risk and Time*, The MIT Press.
- [24] Gould, E., Moav, O. and Bruce Weinberg (2001) "Precautionary Demand for Education, Inequality and Technological Progress" *Journal of Economic Growth*, 6, 285-315.
- [25] Guiso, L. and Monica Paiella (2001) "Risk Aversion, Wealth and Background Risk" CEPR Discussion Paper 2728.
- [26] Guiso, L. and Monica Paiella (2004) "The Role of Risk Aversion in Predicting Individual Behaviour" CEPR Discussion Paper 4591.
- [27] Ham, J. and Robert Lalonde (1996) "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training" *Econometrica*, 64(1), 175-205.
- [28] Harmon, C., Hogan, V. and Ian Walker (2003) "Dispersion in the Economic Returns to Schooling" *Labour Economics*, 10(2), 205-214.
- [29] ISTAT (2003) "Lo Stato dell'Universita'" Istituto Nazionale di Statistica, Roma.
- [30] Johnson, Thomas (1978) "Time in School: The Case of the Prudent Patron" *American Economic Review*, 68 (5), 862-872.
- [31] Keane, M. P. and Kenneth Wolpin (2001) "The Effect of Parental Transfers and Borrowing Constraints Transfers" *International Economic Review*, 42(4), 1051-1103.
- [32] Keane, M. P. and Kenneth Wolpin (1997) "The Career Decisions of Young Men" *Journal of Political Economy*, 105 (3), 473-522.

- [33] Kocherlakota, Narayana (1996) "The Equity Premium: It's Still a Puzzle" *Journal of Economic Literature*, 24 (1), 42-71.
- [34] Kodde, David (1986) "Uncertainty and the Demand for Education" *The Review of Economics and Statistics*, 68(3) 460-467.
- [35] Hartog, J., van Ophem, H. and Simona Bajdechi (2004) "How Risky is investment in Human Capital" Working Paper, Tinbergen Institute.
- [36] Lehvari, D. and Yoram Weiss (1974) "The Effect of Risk on the Investment in Human Capital" *American Economic Review*, 64(6), 950-963.
- [37] Lemieux, Thomas (2005) "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skills?" *Forthcoming American Economic Review*.
- [38] Mincer, Jakob (1974), *Schooling, Experience and Earnings*, New-York, Columbia University Press.
- [39] Olson, L., White, H. and H.M. Sheffrin (1979) "Optimal Investment in schooling when Incomes are Risky" *Journal of Political Economy*, 87(3), 522-539.
- [40] Palacios-Huerta, Ignacio (2003) "An Empirical analysis of the Properties of Human Capital Returns" *American Economic Review*, 93 (3), 948-964.
- [41] Van den Berg, Gerard (2001) "Duration Models: Specification, Identification and Multiple Durations" in J. J. Heckman and E. Leamer eds. *Handbook of Econometrics*, Volume V, North Holland, Amsterdam.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
bet	3288	2513.083	4798.066	1	100000
msk	1822	0.005	0.379	-2.639	1.421
betR	3288	2.370	4745.169	-11421.100	98108.800
edu1	3288	0.006	0.078	0	1
edu2	3288	0.103	0.304	0	1
edu3	3288	0.320	0.467	0	1
edu4	3288	0.305	0.460	0	1
edu5	3288	0.238	0.426	0	1
edu6	3288	0.028	0.165	0	1
eduf	3288	0.116	0.321	0	1
edum	3288	0.073	0.260	0	1
north	3288	0.388	0.487	0	1
centre	3288	0.168	0.374	0	1
south	3288	0.443	0.497	0	1
female	3288	0.198	0.399	0	1
bcf	3288	0.398	0.490	0	1
wcf	3288	0.207	0.405	0	1
sef	3288	0.361	0.480	0	1
uf	3288	0.033	0.180	0	1
bcm	3288	0.089	0.285	0	1
wcm	3288	0.045	0.209	0	1
sem	3288	0.094	0.291	0	1
um	3288	0.771	0.420	0	1
45more	3288	0.609	0.488	0	1
Hincome	3288	48776.650	36472.330	-68000	771077.100
Hwealth	3288	252549	383045.700	-139000	6785000
Howner	3288	0.634	0.482	0	1

Table 2: The individual specific value attached to the lottery: The distribution of *bet*

Deciles	<i>bet</i> (1,000 liras)
1	50
2	100
3	500
4	1000
5	1000
6	2000
7	3000
8	5000
9	5000

Table 3: Results from a grade transition model: Model 1

	Model 1		Model 1	
	Full Sample		with Market Skill	
	Coeff.	SE	Coeff.	SE
α_1	-3.9430	0.227	-1.9255	0.8054
α_2	-1.1173	0.2003	1.3495	0.8307
α_3	0.8272	0.2293	3.8594	0.9271
α_4	4.9246	0.3363	7.8458	1.0332
α_5	9.3026	0.5414	12.7339	1.2135
bet	-0.5648	0.1021	-0.2779	0.1101
msk	-	-	-3.6087	0.2952
eduf	-2.4693	0.1857	-1.9564	0.2445
edum	-1.482	0.2259	-0.6213	0.2837
bcf	0.7438	0.0564	0.3396	0.0634
bcm	1.3851	0.2238	1.4187	0.3671
sef	0.066	0.0827	-0.0687	0.1241
sem	0.911	0.1989	0.8643	0.3713
uf	0.1319	0.2033	0.3346	0.332
um	0.6952	0.1739	0.9877	0.3127
north	-0.2336	0.0669	-0.2483	0.1861
south	0.2775	0.0661	-0.1562	0.1921
female	-0.5784	0.0882	1.4336	0.1728
45more	0.7439	0.0564	0.3397	0.0631
θ_2	-4.5168	0.2871	3.9748	0.3355
prob(θ_1)	0.7500	0.0093	0.3433	0.0908
Mean log lik.	-1.29017		-1.14068	
Sample size	3288		1822	

Notes: prob(θ_1) and θ_2 (θ_1 is normalized to 0) are the type-specific population proportion and support points.

Table 4: Results from a grade transition model: Model 2

	Model 2		Model 2	
	Full Sample		with Market Skill	
	Coeff.	SE	Coeff.	SE
α_1	-5.4437	0.2109	-7.8536	1.3082
α_2	-2.2365	0.1209	-4.5255	1.289
α_3	0.572	0.1351	-1.5969	1.2597
α_4	5.8253	0.2999	4.2266	4.5356
α_5	9.8162	0.446	8.6274	4.5689
Level 1-3				
bet	-0.8240	0.0773	-0.9493	0.2594
bet**2	-	-	0.0901	0.0347
msk	-	-	1.8756	0.9416
msk**2	-	-	-1.1725	0.2113
msk*bet	-	-	1.3623	0.9207
eduf	-3.2014	0.253	-2.3654	0.3428
edum	-1.7012	0.279	-0.6003	0.4226
bcf	1.0775	0.0661	0.4823	0.0838
bcm	1.8997	0.1705	1.7251	0.4805
sef	0.0396	0.055	-0.0808	0.1371
sem	1.4765	0.1782	1.3369	0.492
uf	0.1405	0.1458	0.4289	0.4127
um	1.1599	0.1278	1.2893	0.4348
north	-0.2245	0.0353	-0.2434	0.1832
south	0.4855	0.0568	-0.052	0.1671
female	-0.7214	0.0825	1.4213	0.2196
45more	1.0778	0.0631	0.4856	0.0821

Table 4: Continued

Level 4-5				
bet	-0.1412	0.0611	-0.3872	0.469
bet**2	-	-	0.1594	0.155
msk	-	-	-0.5916	3.2274
msk**2	-	-	-0.4591	0.5512
msk*arrow	-	-	-0.7761	1.4973
eduf	-2.0924	0.3149	-1.319	0.3547
edum	-1.5189	0.281	-0.8807	0.3917
bcf	0.1051	0.0479	0.0308	0.1054
bcm	1.5377	0.3465	2.3722	0.6882
sef	0.0256	0.0792	0.1422	0.2188
sem	0.5941	0.204	0.4899	0.5182
uf	-0.036	0.242	0.9623	0.6278
um	0.45	0.1351	0.7997	0.3939
north	-0.67	0.118	-0.2492	0.2892
south	-0.5616	0.1131	-0.4495	0.2969
female	-0.1776	0.0799	1.3054	0.3092
45more	0.1053	0.0449	0.0308	0.1054
θ_2	-4.5031	0.2141	3.4135	0.3685
prob (θ_1)	0.4671	0.0936	0.3712	0.0924
Mean log lik.	-1.27404		-1.12342	
Sample	3288		1822	

Notes: prob(θ_1) and θ_2 (θ_1 is normalized to 0) are the type-specific population proportion and support points.

Table 5: The effect of market skill and risk aversion on grade hazard rate by level in Model 2

**Average grade transition probability by grade level
for different values of risk aversion and market skills
(standard errors in parenthesis)**

Percentile	Risk Aversion			Market Skill		
	10th	50th	90th	10th	50th	90th
level 1	0.0444 (0.012)	0.0423 (0.012)	0.0344 (0.012)	0.0528 (0.04)	0.0215 (0.04)	0.0098 (0.04)
level 2	0.3362 (0.04)	0.3300 (0.05)	0.2933 (0.05)	0.3814 (0.05)	0.1965 (0.06)	0.0565 (0.05)
level 3	0.5977 (0.07)	0.5921 (0.08)	0.5676 (0.08)	0.7569 (0.07)	0.5781 (0.08)	0.3069 (0.07)
level 4	0.9193 (0.07)	0.9166 (0.08)	0.9052 (0.08)	0.9154 (0.04)	0.8267 (0.04)	0.6643 (0.04)
level 5	0.9950 (0.02)	0.9948 (0.02)	0.9939 (0.02)	0.9969 (0.02)	0.9901 (0.01)	0.9694 (0.01)

Note: The standard errors are calculated using the delta method.

Table 6: Variance decomposition of the odd's ratio of the conditional probability of entering higher education: Model 1

Variables	Explanatory power	
	Full Sample	with Market Skills
Risk aversion	2%	9%
Ability heterogeneity		
	Unobserved	74%
	Market skills	-
Parental human capital	21%	20%
Other socio/geographic	4%	5%

Note: The explanatory power refers to the R-square of the regression of simulated log odds ratio on each variable (or group of variables) taken individually.

Table 7: Variance decomposition of the odd's ratio of the conditional probability of entering higher education: Model 2

Variables	Explanatory power	
	Full Sample	with Market Skills
Risk aversion	2%	10%
Ability heterogeneity	Unobserved	79%
	Market skills	29%
Parental human capital	19%	22%
Other socio/geographic	2%	2%

Note: The explanatory power refers to the R square of the regression of simulated log odds ratio on each variable (or group of variables) taken individually.

Table 8: Ordered Logit

	Full Sample		with Market Skills	
	Coeff.	SE	Coeff.	SE
α_1	-5.4990	0.2325	-6.4176	0.3611
α_2	-2.7602	0.2072	-3.3031	0.2720
α_3	-1.1823	0.2028	-1.1603	0.2621
α_4	1.1357	0.2055	1.4765	0.2642
α_5	4.4945	0.3014	5.2262	0.3919
bet	4.22E-05	7.88E-06	1.78E-05	9.90E-06
msk	-	-	2.6407	0.1478
eduf	1.1761	0.1399	0.8217	0.1829
edum	1.0531	0.1703	0.5385	0.2193
bcf	-1.2595	0.1050	-1.1340	0.1400
bcm	-0.5544	0.2101	-0.7828	0.2706
sef	-0.8889	0.1038	-0.7734	0.1414
sem	-0.4655	0.2106	-0.3848	0.2801
uf	-1.0758	0.1960	-1.0363	0.2891
um	-0.1465	0.1800	-0.4450	0.2273
north	0.1847	0.0927	0.1958	0.1307
south	-0.2791	0.0920	0.0509	0.1277
female	-0.4698	0.0828	1.1271	0.1412
45more	-1.0225	0.0681	-0.4946	0.0913
Log Likelihood	-4192.1179		-2058.8079	
Sample	3288		1822	

Notes: The dependent variable takes values 1 to 6, 1= no formal education and 6=post-college degree.

Table 9: Robustness Analysis: Results from a grade transition model: Model 1

	Model 1		Model 1	
	Full Sample		with Market Skills	
	Coeff.	SE	Coeff.	SE
α_1	-8.9172	0.4441	-9.5593	0.5963
α_2	-6.0891	0.4403	-6.6179	0.5080
α_3	-4.1271	0.4305	-4.2813	0.4393
α_4	-0.0105	0.3623	-0.4891	0.2513
α_5	4.3576	0.4521	4.0464	0.4411
betR	-0.2826	0.0757	-0.3996	0.0992
msk_d	-	-	-1.7439	0.1086
eduf	-2.4973	0.1869	-2.0164	0.2984
edum	-1.5336	0.2405	-0.9455	0.2215
bcf	1.5039	0.1093	0.2983	0.0530
bcm	1.3802	0.2613	1.5399	0.2329
sef	0.0657	0.1279	-0.0178	0.0887
sem	0.8862	0.2506	0.8768	0.3639
uf	0.1199	0.2157	0.5443	0.3311
um	0.688	0.2559	0.9662	0.1904
north	-0.2442	0.088	-0.3837	0.2096
south	0.3007	0.0867	-0.0931	0.1700
female	-0.6033	0.0923	0.8227	0.1130
45more	0	0.1545	0.3054	0.0524
θ_2	4.5477	0.2758	4.0004	0.4810
prob(θ_1)	0.4623	0.1054	0.4427	0.0923
Mean log lik.	-1.29381		-1.11888	
Sample size	3288		1822	

Notes: In this table we use $betR$, the residuals of the risk variable bet on household income $Hincome$, household wealth, $Hwealth$, and an indicator of home-ownership $Howner$. The market skill variable, msk_d , is a binary variable that splits the sample in those who have a positive wage fixed effect estimate and those who have a negative one.