On the uncertain nature of human capital investments
Mazza, J.

Citation for published version (APA):
Mazza, J. (2012). On the uncertain nature of human capital investments Amsterdam: Universiteit van Amsterdam

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

Introduction

This thesis presents four self-contained essays. Each of the four studies can be read independently, but a common thread runs through them. The most obvious regards the object of study. All the four studies presented here pertain to an often neglected characteristic of educational investments in human capital literature: its unpredictability and how individuals account for and respond to it.

Even if the recognition of the intrinsic uncertainty under which educational choices are taken is certainly not new in economic theory, dating more than forty years back (Becker, 1964; Schultz, 1971), the first early attempts to systematically incorporate uncertainty into a classic human capital theory framework by Weiss (1972), Levhari and Weiss (1974) and King (1974) did not generate adequate attention in the profession, or at least not immediately. It took some decades for those seeds to blossom into what is now a growing body of literature to which the present work aims at contributing.

Depending on the chain of events that an individual is confronted with, educational risk can be defined along two dimensions: ex-ante and ex-post. Ex-ante risks pertain to all those uncontrollable and unforeseeable events that might undermine the successful accomplishment of the elected course of study. No one can predict how well he/she will perform once entered a specific field of study. Nailing or failing depend on abilities, commitment, effort and match between inclinations, abilities and requirements of the educational path undertaken. For some of these elements pre-knowledge can be at best partial, for others impossible.

Ex-post risks stem from occurrences developing after the educational path
has been terminated and the individual approaches the entry into the labor market. Where he will position in the earning distribution of a specific career will not be learned until that person undertakes it. Few elements can help someones prediction about future accomplishment. Even academic success, which could be easily pegged as the most immediate indicator of someones ability, only serves as an imperfect predictor of professional success as the correlation between past performance in school and performance in the labor market is everything but perfect. Furthermore, temporal laps between choice and accomplishment of the selected field of study increase incertitude. The labor market, as every other market, is intrinsically and increasingly, dynamic, its request can vary rather quickly in time. Technological progress can condemn entire professional figures to obsolescence in the worst cases or, more often, seriously undermine their profitability. This specific type of risk had been defined “market risk” (Hartog et al., 2004).

The economically relevant aspect of uncertainty in education investment resides primarily in the possible repercussion that \textit{ex-ante} and \textit{ex-post} risk have for selection of education and, in turn, for the type of skills supplied in the labor market.

Empirical research shows that the dominant sentiment towards risk is aversion, not attraction (Holt and Laury, 2002). This unsurprising result bears some evident consequences for the prototypical utility maximizing agent at the basis of human capital theory. The most evident is the avoidance of those educational paths which are deemed to be too unsafe. This apparently simple reasoning implicitly contains a series of assumptions and consequences opening up vast spaces ahead of economic research. It implies that agents maximize their utility; that educational risk exists and differs across type and/or levels of education; that individuals are aware of the existence of \textit{ex-ante} and \textit{ex-post} uncertainty; that considerations about uncertainty enters their utility function at the time of educational choices and that they enter negatively.

In this work, the utility maximizing agents approach that since the seminal works of Mincer (1958) and Becker (1962) has proved to be both a useful tool and a plausible assumption for modeling investment in human capital will remain unchallenged. Instead, some of the other links of the logical chain of reasoning exhibited above are empirically tested. Chapter 2 elicits, from a sample of Dutch high school students, the level of information about existing labor market conditions in general and future wage variability in particular,
that individuals possess when deciding upon investing in university education. Chapter 3 assesses how, for various samples of individuals coming from different countries and different educational systems, different levels of education contribute to the increase or decrease of risk. In Chapter 4 the importance for concerns about future payoffs uncertainty on choices of type of college education is gauged. To conclude, Chapter 5 resumes the theme already discussed in Chapter 3, but proposes a different, and possibly more robust, methodology for estimation.

Chapter 2 answers a fundamental question for the understanding of investment decisions into education: do individuals have perception of its uncertain nature? This chapter provides empirical evidence on the level and extent of knowledge that Dutch high school students possess at the time immediately preceding the decision on whether to continue or not to tertiary education. The use of high school students has the considerable advantage of targeting on the most relevant group of people, at the decisive moment for their educational investment decision. The definition adopted of level of knowledge includes both awareness of the riskiness entailed in such an investment and the returns that it can generate. The particular structure of the collected data also allows inference on three interrelated aspects: if, and to which amount, a compensation for risk is required in the labor market; whether students possess private information and if they exploit it when forming their expectations; the expected returns that an investment in education should guarantee.

The methodology adopted is partly novel and partly imitative. It is novel in the medium exploited for data gathering as data on individuals’ expected median wages for different schooling scenarios were collected via internet and the participation to the survey was completely voluntary. Internet surveys are still rarely exploited in economics. The reason behind the scarce use of such a potentially powerful tool and in such a freely accessible environment is easily explained by a general skepticism towards non controllable participation in surveys. Unconstrained participation can create several problems to the researcher. The most relevant and potentially undermining for the reliability of the econometric analysis is self-selection of participants. If the type of people willing to participate and finish the survey differ along some relevant dimension from the population of interest at large, the results obtained analyzing this selected group of people can not be extended outside the sample. Tests for the existence of selection on observables are provided and the reliability of gath-
ered data is confronted with laboratory collected ones hoping to clear some doubts on the soundness of this methodology.

The imitative part regards the general structure of the survey and the elicitation of students expectations which heavily relies on a method developed by Dominitz and Manski (1996) who derive, from a small American sample of both high school and college students, the income expectations at different moments of the life cycle conditional on different schooling scenarios. They also extrapolate a subjective earnings distribution, characterizing in this way the riskiness of educational investment. They find positive perceived returns to college education and awareness of the uncertain nature of the investment in their particular sample of students. Their main contribution resides in the methodological part of data collection. The original method has only been implemented in controlled environments, under supervision and with direct feedback on inconsistent probability statements and it has been applied by Schweri et al. (2011) to a sample of Swiss college students.

Chapter 3 is a replication of a recent study by Chen (2008) on a sample of American men. The replication is carried on on the same data as the original work and extended to three additional samples of American females and British and German individuals.

Economists often praise the virtue of replication, but they rarely attempt it. Putting empirical results under careful scrutiny is an important if not essential task per se. Hamermesh (2007) defines pure replication as examining the same question and model using the underlying original data set and scientific replication as the same type of research on different sample and different population. The results presented here underscore the value of both types of replication. Chen (2008) presents estimates for the causal effect of schooling on risk robust to at least two longstanding methodological complications affecting the previous literature on the matter: self-selection into education and the presence of unobserved heterogeneity. Concerns of self-selection arise every time an individual has the freedom of choosing whether or not to participate in any kind of activity such as schooling. Unobserved heterogeneity, instead, indicates all that information used by the individual when choosing among different alternatives hidden to the researcher observing the revealed choice. Both phenomena can severely undermine the reliability of any type of estimates calculated on revealed data. Even though others before her successfully

---

1For a textbook discussion of sample selection and self selection see Cameron and Trivedi (2005)
tackled the issue of selectivity (Abadie, 2002; Abadie et al., 2002; Chen and Khan, 2007) and accounted for unobserved heterogeneity (Carneiro et al., 2003; Cunha et al., 2005; Cunha and Heckman, 2007) when estimating the dispersion in outcomes, Chen (2008) is the first to jointly solve those two puzzles. Therein lies the importance of this work and the motivation for carefully replicating it.

In recent years, applied work has dedicated growing attention the potential dangers that self-selection poses to econometric estimates when the mechanism is unaccounted for (see, for example Willis and Rosen, 1979; Heckman, 1979; Maddala, 1983). The literature on risk and education makes no exception as discussed in Chapter 3 and 5. Chapter 4 tests exactly this crucial issue in the specific context of educational choices by asking if, and to which extent, concerns about uncertainty of future payoffs affects individuals’ schooling decisions.

This unexplored question has potential repercussions on both the design of the appropriate empirical strategies to be followed for drawing reliable conclusions on the casual effect of education on risk and on the formulation of possible policies providing protection against risk by the policy makers. The estimation strategy utilized exploits the model developed in Chapter 5 combining it with other proposed semiparametric estimators (Lee, 1983; Dahl, 2002) and thus preserving most of the valuable attributes of that model.

Chapter 5 is strictly related to Chapter 3 as it answers the same research question, but it does so by forgoing the normality assumption which is central for the validity of the method replicated there. The Heckman (1979) method applied by Chen (2008) is a valid strategy in order to overcome the selectivity issue given that the error terms in the estimated choice and outcome equations are distributed normally. In fact, it has been widely accepted (see, for example, Goldberger, 1983) that the parameter estimates obtained via the classic selection correction model may be significantly biased due to the restrictive parametric distributional assumptions. Because of the importance of the topic, it is important to know if the results of Chen (2008) are driven by the joint normality assumption made on the distribution of the unobservables. The method introduced in Chapter 5 is robust to a possible misspecification of the error term structure producing valid estimates for the effect of schooling on risk in a wider set of cases than its parametrical counterpart.

The last chapter, Chapter 6, summarizes the the main findings and conclusions from the previous four chapters presented in this thesis.
Introduction

So what is new and what is not in my thesis? I surely do not claim full originality of the themes tackled in the present research as I do not claim to have exhaustively treated all the interrelated aspects of risk and education since other frequently researched aspects in this type of literature such as the quantification of the magnitude of risk that education investments entail compared to other type of investments (Palacios-Huerta, 2003; Hartog et al., 2004) or the labor market responses to *ex-ante* and *ex-post* risk (King, 1974; McGoldrick, 1995; Moretti, 2000; Pereira and Martins, 2002; Hartog et al., 2003; Diaz-Serrano and Hartog, 2006) are not discussed here².

What I can safely claim is that every chapter, some more than other, composing this research contain either some new empirical evidence, or some methodological development, or both.

A more detailed, and hopefully objective, assessment of each chapter’s contribution to the science would probably conclude that the most original contributions, concerning both theoretical and empirical advancements, are concentrated in the fourth and fifth chapters of the present work. Chapter 4 tackles a previously neglected research question and thus the empirical findings and related potential policy implication have been rarely, if ever, considered before. If the novel contribution offered in Chapter 4 is empirical, the novel contribution of Chapter 5 is mainly technical. In fact, Chapter 5 presents important improvements to previously applied econometric techniques in the context of the literature of uncertainty and education, improvements that allow to overcome some major difficulties affecting empirical estimations of the impact of schooling on uncertainty.

The first two chapters pay a bigger duty to previous works. In fact, even in the presence of a completely original dataset, it can not be denied that the scope of inquiry and lead lines for the methodology followed in Chapter 2 are common to few previous researches. Having said that, some interesting and novel contributions are still brought about. The most interesting one is probably the *medium* exploited for gathering data. Internet generated dataset are very rare to come across in economic empirical research. The potential gains for developing a sound internet based data gathering procedure are immediately understandable as it would grant quick and easy access to a vast number

²Some of the main results of these two streams of research suggest that educational investments offer superior hedging against risk compared to financial investments (Hartog et al., 2004) and more so at the higher end of the educational distribution (Palacios-Huerta, 2003) and that labor market recognizes and monetarily rewards riskier occupations (King, 1974; McGoldrick, 1995; Moretti, 2000; Pereira and Martins, 2002; Hartog, 2011)
of people and would give the possibility to customize each questionnaire to the specific needs of each particular research. This chapter offers important contribution for the understanding of possible difficulties and the formulation of plausible solutions for similarly designed researches.

The case of Chapter 3 is exceptional as the ultimate scope of replication works is that of repeating precedent analysis and carefully check its main findings, therefore it is not original by construction. Anyhow, as explained before, the importance of replication studies for scientific research must not be underestimated. Checks of published results is paramount for the maintenance of a profession’s internal and external credibility and it is not rare to hear established scientists calling for more not less works of this nature.

In combining both new and old methods and answering to newer and older questions this thesis produces evidence on students awareness of labor market conditions, discussing, in the meantime, the merits and weaknesses of a potentially powerful tool for future research; introduces a new model for casual inference of the effect of schooling in education and answers a previously overlooked, but essential, question on the consequences of uncertainty for schooling choices.