Successful entrepreneurship and human capital

van der Sluis, J.

Successful Entrepreneurship and Human Capital

ISBN 978 90 5170 734 2

Cover design: Crasborn Graphic Designer bno, Valkenburg a.d. Geul

This book is no. 402 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Successful Entrepreneurship and Human Capital

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr J.W. Zwemmer
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op vrijdag 20 april 2007, te 14:00 uur

door

Justin van der Sluis

geboren te Oostdongeradeel

PROMOTIECOMMISSIE

Promotores:
prof. dr J. Hartog
prof. dr C.M. van Praag

Overige leden van de commissie:
prof. dr H. Maassen van den Brink
prof. dr H. Oosterbeek
prof. dr S.C. Parker
prof. dr A.R. Thurik
prof. dr W.P.M. Vijverberg

Faculteit der Economie en Bedrijfskunde

Acknowledgements

Just after having obtained a Master's degree in 'Work and Organizational Psychology', I tried to start up a small business in consultancy together with a good friend. After struggling for a while to find customers and developing a good concept, my attention was drawn to a vacancy for a PhD position "entrepreneurship and education" at the Universiteit van Amsterdam. This seemed to be the ideal opportunity to better develop my academic skills. Moreover, I felt I was better off studying entrepreneurs than being an entrepreneur myself. However, doing a PhD involves risk taking, trying something new, developing new ideas, working independently, selling products and making mistakes. Viewed in this way doing a PhD is not that different from being an entrepreneur.

During the last four years I met a lot of people who all influenced me in one way or the other. My two supervisors Mirjam van Praag and Joop Hartog however influenced me most. Joop's enthusiasm convinced me to start the PhD project. Although he was abroad quite often, when I needed assistance he was always there. Joop's positivism is an example. Mirjam, who was my day-to-day supervisor, introduced me to inspiring people and new ideas. I will always remember the numerous discussions we had, the projects we did together and the fun we had. Mirjam did not only teach me about entrepreneurship but also showed me an entrepreneurial way of thinking. Our work for the Amsterdam Center for Entrepreneurship ensures we will keep our cooperation alive.

The people at Scholar offered me a great research environment. The weekly labor seminar, the discussions in the hallway and the atmosphere pushed me to work harder and make my papers better. In particular I would like to thank Chris and Adam for helping me out when my research got stuck or things got too complicated. But above all I want to thank them and all the other (visiting) Scholars for the lunches we had together. The discussions were great, stimulating, sometimes hilarious and provided enough energy, besides the food, for the rest of the day.

I would also like to thank some people outside Amsterdam. My gratitude goes to Simon Parker, for inviting me at Durham University where I spent two months of my PhD period.

Although North-East England is cold in winter, the environment was warm and friendly. I am also grateful to Wim Vijverberg for his collaboration in the beginning of my PhD project and his very valuable comments in the end.

Finally, I would like to thank my family and friends, in particular my parents Bart and Berna, for providing a listening ear, excusing me at several social happenings and forgiving me for sailing contests lost due to my absence. But above all I would like to thank Marianne for loving and supporting me at all times.

Justin van der Sluis

Amersfoort, March 2007

Contents

Chapter 1

Introduction

1.1 Main objective

The contribution of entrepreneurs to economic performance is significant, a fact which is recognized by both researchers and public policy makers. In most countries, entrepreneurs generate a substantial part of the national income as well as jobs. They also contribute to R&D and innovations (Audretsch and Keilbach, 2004; Van Stel and Storey, 2004). In 2000, the EU Lisbon strategy encouraged the development of higher rates of innovation in Europe. Most European governments emphasized the development of entrepreneurship to reach this goal.[1] Entrepreneurship is also regarded as a good alternative to wage employment for people who need more flexibility in combining work and family obligations than many employers can offer. Moreover, entrepreneurship is often regarded as a route out of unemployment (Fairlie and Meyer, 1996). In addition, entrepreneurship might also be a preferable alternative for minority groups that are discriminated against by employers in the labor market (Moore, 1983).

However, these benefits are generated only if the entrepreneur is successful. Identifying determinants of successful entrepreneurship and using these as instruments to influence the achievement of entrepreneurs can therefore increase welfare. This book focuses on one type of determinant, namely human capital. In particular, the focus lies on measuring the performance effects of formal education.[2] This focus is motivated as follows. First, formal education is a factor that possibly affects entrepreneurship performance and that can be shaped to some extent by public policy. Second, the effect of formal education on employees'

[1]See, for instance, *Working together for growth and jobs: A new start for the Lisbon Strategy*, President Barroso's address to the Spring European Council, 2005, stating that "There are just too many obstacles to becoming an entrepreneur or starting a business, and, therefore, Europe is missing opportunities" (p. 16).

[2]We will also address the performance effects of intelligence.

labor market performance has been measured extensively. The average rate of return to one extra year of education for employees turn out to be large and significant.[3] However, the returns to formal education for entrepreneurs have not yet been measured using the same rigorous methods leading to the quantification of causal effects. Applying similar methods will show if and to what extent the returns to education are different for entrepreneurs and employees. This is interesting since changes in the education system influence both labor market groups. If the benefits of formal education turn out to be different for entrepreneurs and employees, this could have several policy implications.

The aim of this book is to answer the following three research questions:

1. What is the current state of research on the relationship between formal education and entrepreneurship?

2. How high are the returns to formal education for entrepreneurs as compared to employees?

3. What are the returns to (different types of) intelligence, on top of the effect of formal education?

1.2 Outline of the book

Chapter 2 provides a review of empirical studies on the impact of formal education on entrepreneurship selection and performance in industrial countries. It describes the average effects, and explains the variance in magnitude and significance of these effects across almost a hundred studies. Five main conclusions result from this meta-analysis. First, the impact of education on selection into entrepreneurship is insignificant. Second, the effect of education on entrepreneurial performance is positive and significant. Third, based on the measurement of conditional correlations, the returns to a marginal year of schooling are 6.1% for an entrepreneur. Fourth, the correlation between education and earnings is weaker for entrepreneurs than for employees in Europe, but stronger in the USA. Fifth, the returns to schooling for entrepreneurs are higher in the USA than in Europe. A number of suggestions to move the research frontier in this area of inquiry are also presented. One of these suggestions is that entrepreneurship literature on education can benefit from the techniques used to estimate the returns to education for employees. An attempt to do so is made in Chapters 3 and 4.

[3] Throughout this book "the average rate of returns to one extra year of education" will be referred to as "the returns to education".

Chapter 3 presents a first attempt at applying the estimation techniques from employee literature to entrepreneurs. In this study the returns to education for entrepreneurs and employees are estimated and compared based on 19 waves of the NLSY database. While using instrumental variable techniques (IV) and taking account of selectivity into entrepreneurship, the returns to education turn out to be significantly higher for entrepreneurs than for employees (18.3 percent and 9.9 percent, respectively). Various analyses are performed in an attempt to explain the difference. (Indirect) support is found for the argument that the higher returns to education for entrepreneurs are due to fewer (organizational) constraints faced by entrepreneurs when optimizing the utilization of their education.

Chapter 4 confirms the results found in Chapter 3 by using a different set of instruments for education, i.e. changes in compulsory schooling laws in the USA. Again, the returns to education are significantly higher for entrepreneurs than for employees in the USA (26.4 and 13.2 percent respectively).

Chapter 5 focuses on the returns to intelligence, rather than on the returns to formal education. The following three questions are answered empirically for both entrepreneurs and employees; (1) To what extent does an individual's general intelligence level affect earnings? (2) Do different areas of intelligence (e.g. math ability, language ability etc.) affect earnings differently?, and (3) To what extent does the balance in an individual's scores in these areas of intelligence affect an individual's income? The latter question is related to and extends Lazear's Jack-of-all-Trades (JAT) theory pertaining to entrepreneurs.

The results indicate that an individual's level of general intelligence increases both entrepreneurs' and employees' incomes to the same degree. Moreover, entrepreneurs benefit from specific areas of intelligence that differ from those that are valuable for employees. The balance in the various areas of intelligence is another determinant of earning power. The greater the balance, the higher the earnings, but only for entrepreneurs. This finding supports Lazear's JAT theory.

Finally, Chapter 6 presents an overview of the preceding chapters. This chapter also discusses the policy implications of the results in this thesis.

Each of the following Chapters 2 to 5 is based on independent working papers (see the notes at each chapter title referring to the original papers and their authors). Due to this independence, it might happen that some points or clarifications are made twice. I apologize for this occasional overlap across studies.

Chapter 2

Entrepreneurship, selection and performance: A meta-analysis of the role of education[1]

2.1 Introduction

A substantial and technically sophisticated economics literature has developed in the last decade about returns to schooling (e.g., Webbink, 2005; Ashenfelter, Harmon and Oosterbeek, 1999; Card, 1998; Psacharopoulos, 1994). Returns to schooling have been measured for many countries and over a number of years in a way that allows both international comparisons and trend analyses. Innovative methods have been developed and applied within this strand of research to assess whether the measured correlations between schooling and income reflect a causal effect of schooling on earnings (e.g., Webbink, 2005; Ashenfelter et al., 1999). However, almost without exception, returns to schooling measured with these innovative methods refer to the returns employees generate from their years at school.

The objective of the current chapter is to compile, analyze and describe all empirical studies we could locate that measure the effect of formal schooling on entry of and performance among entrepreneurs in industrialized countries.[2] The combined body of evidence will demonstrate whether generally there is an effect of formal education on entrepreneur performance and selection into entrepreneurship. In addition, such a compilation of studies allows measuring the impact of, for instance, the percentage of ethnic minorities in the sample,

[1]This chapter is based on Van der Sluis, Van Praag and Vijverberg (2003).

[2]In Van der Sluis, Van Praag and Vijverberg (2005) we examined the same set of relationships that exist in developing countries. Since the general level of schooling is lower and a substantial portion of the labor force works on the farm, an analysis of the developing country studies is not easily fused with one of industrial country studies.

the size of the sample, the country studied, on the magnitude of the effect of education on entrepreneurship outcomes. The chapter thus develops insights into the outcomes, practices and methodologies used, and leads to a deeper understanding of the problems encountered in this type of entrepreneurship research. It also compares these practices to the studies pertaining to employees' returns to schooling.

Our main instrument to study the relationship between formal schooling and entrepreneurship entry and performance is a meta-analysis, which proceeds in two steps. The first part is descriptive and describes the main effects found in the entrepreneurship literature. The second part is an analysis of variance which uses various characteristics of the studies to explain differences in their measured effects of schooling on entrepreneurship outcomes.

To be able to perform these analyses, all relevant studies from the large number of published and unpublished academic studies have been gathered.[3] Each of the 94 resulting studies measures, among others, the impact of schooling on entrepreneurship entry, performance (such as earnings, profit, survival, and so forth), or both for a specific sample (a given country, time period, gender, occupation, or whatever). This results in a database comprising 299 observations. Many studies report the estimates of several regression models, for instance using various performance measures or various subsamples stratified by gender, race, or country. This explains how 94 studies yield a multiple of observations. The descriptive part of the meta-analysis is based on this dataset and measures the (average) relationship between education and the entrepreneur outcomes mentioned. Besides, it contains information on several characteristics of each of these observations, such as sample size, the type of sample, ability controls used, whether the study is published or not, and, if it is published, the type and level of journal. These factors constitute the potential determinants of cross-study variation in the effect of education.

Several conclusions can be drawn from the meta-analysis. First, the effect of education on entry is neither positive nor negative. Second, entrepreneurship performance, regardless of the performance measure used, is significantly and positively affected by formal schooling. Third, the returns to education, as measured by means of conditional correlations, are estimated at 6.1%.[4] Fourth, the returns to education for entrepreneurs seem to be lower than for employees. However, a compilation of those studies that estimate the returns to education for both entrepreneurs and employees in a comparable fashion shows that entrepreneurs in the USA have slightly higher returns to their education than employees, whereas the reverse is true in Europe. Fifth, the returns to schooling in entrepreneurship are higher in the

[3]Various databases and the Internet have been used to find all relevant studies published after 1980 and all unpublished studies after 1997. The database includes studies up until December 2002.

[4]Based on predominantly USA studies.

United States than in Europe. Moreover, as entrepreneurs, minorities benefit less from their education than others. Female entrepreneurs seem to benefit more from their education than male entrepreneurs. Finally, the set of published studies that have estimated the effect of education on entrepreneurship outcomes does not appear to suffer from a publication bias, unlike the set of published studies of estimates of the returns to education for employees (see Ashenfelter et al., 1999).

The meta-analysis brings significant weaknesses in the schooling-entrepreneurship literature to the surface. The database of compiled studies demonstrates a need for greater uniformity in the definition and measurement of entrepreneurship, schooling, and performance, such that results can be compared numerically across studies. These problems of measurement and definition have prevented the development of a quantitative knowledge base of the (potential) returns to schooling in entrepreneurship. Moreover, a comparison of the quality standards of the methodological approaches indicates that the schooling-entrepreneurship literature lags behind the comparable but neatly defined schooling-employees literature (Webbink, 2005). The schooling-entrepreneurship literature could make significant strides by addressing similar endogenous choices of entrepreneurs with similar techniques.

The chapter proceeds as follows. Section 2.2 briefly summarizes economic theory about the (supposed) relationship between entrepreneurship entry, performance and educational attainment. Section 2.3 describes the data gathering process. Section 2.4 illustrates the lack of standards in the literature that hampers the development of a quantified knowledge base. Section 2.5 summarizes the evidence concerning the relationship between schooling on the one hand and entrepreneurship entry and performance on the other hand. Substantial cross-study variation comes to light, which is precisely the motivation for doing the analysis of variance that is the subject of Section 2.6. Section 2.7 concludes.

2.2 Economic theory[5]

There are several theoretical determinants of entrepreneurship selection and performance that have been empirically tested (see the overview article by Le, 1999). Among them are risk attitude, access to capital, various types of labor market experience, economic conditions, business acumen, family background, psychological traits, and last but not least education.

[5]The meta-analysis includes several empirical studies from other disciplines than economics. However, we do not aim to put forth an overview of theories developed outside the economics discipline. The particular empirical studies reviewed from these broader disciplines including labor relations, sociology and psychology are not rooted in any theoretical framework other than various (sometimes valuable, but quite ad hoc) theoretical conceptualizations.

This section briefly summarizes the theoretical roots of the relationship between education and entrepreneurship.

2.2.1 Formal education as determinant of entrepreneurship selection

Le (1999) argues that there are two different channels (managerial ability and outside options) through which the level of education might influence the propensity to become self-employed. Calvo and Wellisz (1980), inspired by Lucas' 1978 general equilibrium model, explain the impact of educational attainment on the probability of selection into an entrepreneurial position through managerial ability. Education would enhance managerial ability, which in turn increases the probability of entrepreneurship.

The other channel generates an opposite, negative effect on entrepreneurship selection. Higher levels of education may generate better outside options (i.e., more lucrative wage employment under better working conditions) and thus decrease the likelihood of entrepreneurship as the preferred choice. It is yet unclear what the theoretically predicted effect of these offsetting forces might be on the relationship between schooling and entrepreneurship selection.

2.2.2 Formal education as determinant of entrepreneurship performance

According to the Mincerian specification of the determinants of individual earnings, the main factors affecting earnings are schooling and experience. Schooling is acknowledged not only for its productive effect on the quality or quantity of labor supplied, as assumed by Mincer, but it also has value as a signal of productive ability in labor markets without complete information (Riley, 2002; Spence, 1973). This signaling value is also referred to as the screening effect. In signaling, the party with private information –i.e., the employee in the selection and hiring process by employers, and the business owner in the selection process by clients, stakeholders, or business relations– takes the lead in adopting behavior that, upon appropriate interpretation, reveals information about his own type or productivity. Educational attainment has signaling value as long as two self-selection constraints are met: (1) Completing education is either impossible or prohibitively costly for low productivity types, and (2) All high productivity workers indeed self-select into higher education. For these two conditions to hold, it is obviously required that the completion of education be more expensive for low productivity types than for high. With costs aligned in this way, a separating equilibrium exists that renders schooling a credible signal: The high productivity group prefers to complete higher levels of schooling, whereas it is not in the interest of low productivity workers to feign that they are of the high productivity type, even if high educated workers are remunerated accordingly. These conditions are met in the entrepreneurial market as long as entrepreneurial ability and learning ability are sufficiently positively correlated, and stakeholders are less informed about entrepreneurial productivity than prospective entrepreneurs themselves are.

In this context, two hypotheses prevail that have been subject to empirical testing (Riley, 2002; Spence, 1973). The first, the Strong Screening Hypothesis (SSH), states that the return to education would be zero in the absence of screening. In other words, education would not have a productive effect. The second, the Weak Screening Hypothesis (WSH), states that the return to education would be significantly lower in the absence of screening. Under the WSH, human capital productivity effects are not ruled out.

2.2.3 Basic integrated models of choice and performance

Another type of models, known as structural models, simultaneously explains the occupational choice and performance of labor market participants. In these models, the actual division between entrepreneurs and wage labor turns on the distribution of individual characteristics among the utility maximizing population. In Lucas (1978) and Van Praag and Cramer (2001), this characteristic is individual entrepreneurial ability as determined by, for instance, education.[6] In such models, education generates higher levels of (expected) entrepreneurial ability that, in turn, causes higher levels of entrepreneurial performance. Not only may enterprise profit rise, but the entrepreneur may also enjoy success in other dimensions: a larger firm, richer perquisites, higher socioeconomic status. In all, this higher level of expected performance increases the expected utility attached to entrepreneurship and thereby favors this occupational choice. Education now has a positive effect on both the choice of and the performance in entrepreneurship. Of course, if education is also allowed to raise the productivity of the individual as an employee, the effect of education on the choice again becomes ambiguous.

[6]Estimation of the structural model (Van Praag and Cramer, 2001) indeed shows that education plays a significant role in the joint determination of entrepreneurship choice and performance.

2.3 Constructing a database for meta-analysis

The meta-analysis conducted in this chapter assesses whether the vast empirical entrepreneurship literature offers any consistent findings with respect to the impact of educational attainment on performance in and choice of entrepreneurship. Meta-analysis is a quantitative tool that synthesizes previous research findings that share common aspects that can be addressed in a statistical sense. Meta-analytical techniques have largely been developed outside the academic field of economics in the medical and natural sciences. More recently, economists have also started applying meta-analysis, see for instance Phillips (1994), Card and Krueger (1995), Nijkamp and Poot (2004), Longhi, Nijkamp and Poot (2005) and Abreu, De Groot and Florax (2005). Roberts (2005) and Stanley and Jarrell (2005) provide an overview of applications and issues in meta-regression analysis in economics.

A meta-analysis can only be performed based on a sample of studies that all measure a certain type of relationship. As a consequence, our study is limited to much-studied entrepreneurship outcomes such as (i) entry into entrepreneurship, (ii) whether an individual currently is an entrepreneur, and (iii) the various entrepreneurial performance measures (such as earnings, profits, survival and growth). The objective is to assess the mean estimates and the variation across studies in the effect of schooling on the three entrepreneurship outcomes mentioned. Let us assume that each observation in the database to be constructed describes a study of the effect of schooling (S) on entrepreneurship selection or performance (Y). To be more precise, let us define such a relationship as $Y = \gamma X + \beta S + u$, where X is a set of controls. Let b be the estimate of β, and let Z be various features of the way this relationship has been estimated, such as a characterization of the sample used, the type of controls entered, the precise form of the dependent variable, and so forth. With meta-analysis, one describes b across studies in the database, either with simple descriptive statistics or by means of regression analysis. We define this regression model as $b = \alpha Z + \epsilon$. However, some studies report more precise estimates of β than others, and such studies deserve a greater weight in the meta-analysis. The meta-analytical approach by Hedges (1992) is designed to do just that: it weighs each observation of b by the inverse of its standard error.

In order to meet the data requirements for such an analysis, the available empirical literature needs to be categorized and selected according to systematic rules. As Nijkamp and Poot (2004) note, these rules prescribe that the database must have sufficient coverage (i.e., is representative of the population of published and unpublished studies) and precision (i.e., provides high quality information on the issues at hand). Appendix 2.A offers a detailed account of the rules employed in the compilation of the studies, and it also describes features of the resulting database.

In essence, the database contains journal articles, book chapters, and working papers published after 1980 and before December 2002 (when the data gathering was finished), pertaining to entrepreneurship selection and performance in industrial countries. Studies that focused on developing countries or transition economies are omitted in order to preserve homogeneity (see footnote 2). In all, 94 studies are represented in the database, yielding 299 observations describing a quantified relationship between education on the one hand and entrepreneurship selection or performance on the other.[7] Among these, 144 (48 percent) examine performance, 69 (23 percent) investigate entry into entrepreneurship, and 86 (29 percent) specify the dependent variable as "being self-employed." The latter is a stock (rather than flow) variable that is a hybrid of entry (everyone who is self-employed has entered this occupational status) and performance (it generates an overrepresentation of survivors). We therefore keep "stock studies" as a separate category.[8]

Furthermore, whether referring to entry, performance or stock, structural studies must be distinguished from reduced form studies. Several authors have acknowledged that the selection into self-employment is an endogenous choice, dependent on the expected performance or on the utility from income. Failing to account for such selectivity effects may well bias the estimated return to education (or any other variable). Those studies that are labeled as "structural", as opposed to "reduced", attempt to incorporate at least some kind of a deliberate occupational choice (selection correction) on the part of labor force participants. It is worthwhile to compile such studies under a separate heading in order to examine the direction of the selectivity bias.[9] Eleven percent of the observations are structural, constituting 19 percent of the performance observations, 10 percent of the entry observations, and a negligible proportion of stock observations. It should be further noted that none of the studies take account of the potential endogenous nature of the number of years of formal schooling an individual completes.

[7]Many studies actually estimate several specifications of exactly the same relationship on the same (sub)sample, using various sets of controls or estimation techniques. In such cases, we selected the study's (final) set of estimates that best represent the study's objective; other specifications are omitted to create a sample of independent observations, which is a requirement of meta-analytical research.

[8]See Van Praag (2003) for a theoretical analysis of the interrelations between entry, survival and stock of self-employment.

[9]There are actually many choices that could fall under this heading. For example, it is often assumed that the individual is working anyway and that the only choice to be modeled is whether to be self-employed as opposed to working for a wage. However, this choice model could be augmented with the choice whether to work, a choice between working in the public sector or the private sector, a choice to work for a large corporate organization as opposed to a smaller business that offers a similar environment as one's own enterprise, etc. Obviously, there is no study that includes all of these features. The point is that structural studies attempt to remove the bias caused by ignoring one or perhaps several of these choices but that one could easily think of other omitted selectivity factors that still may bias the estimated returns to schooling. Any comparison between reduced form and structural model estimates therefore has obvious limitations.

2.4 A literature in search of standards

The construction of this database revealed a great variation among the studies in facets such as the definition of the primary variables of interest, i.e., entrepreneurial outcomes and education, as well as the type of data that were used, the analytical techniques that researchers employed, and so forth (see Appendix 2.A). Variation in research may represent innovative thinking, or it may reflect a field that is adrift without a rudder. Anyway, the lack of uniform definitions of the key variables in these studies represents the challenge of performing a meta-analysis.

Most entrepreneurship theories refer to untestable notions of the innovative free mind of a resourceful entrepreneur. Fortunately, empirical definitions of entrepreneurship are fairly well comparable across studies, apart from the stock versus entry distinction. Most researchers empirically define entrepreneurs as individuals who are self-employed. Some specifically define the entrepreneur to be the starter of a new firm. Other studies examine entrants, defined as those labor force participants who switch during a specific period from wage employment (and sometimes unemployment) to self-employment.

Unfortunately, the literature has not yet converged upon standardized definitions of performance and educational achievement. Table 2.1 shows the various empirical definitions of self-employment performance that have been used: the majority (52 percent) of performance observations focus on self-employment earnings –annual, monthly, weekly, or hourly– which is then entered in linear or logarithmic form. Another 27 percent examine exit or survival (types of studies that can be easily translated into each other); 7 percent of the observations pertain to hazard models; and 8 percent estimate models of mixed categories of growth (employment growth, asset growth, profit growth). Other categories each comprise less than 5 percent of the 117 performance observations.

Table 2.2 indicates that a lack of uniformity in the measurement of schooling may generate additional problems for a quantified meta-analysis of the relationship between schooling and entrepreneurship. Years of education, the only continuous measure of schooling available, is often used (18, 30 and 41 observations of entry, stock and performance respectively). The dummy "college graduate," though, is the most frequently used measure of education: it appears in 42 entry, 51 stock and 64 performance studies.[10]

[10]Schooling levels in British studies diverge somewhat from others: O-levels, since 1991 called General Certificate of Secondary Education (GSCE), is an exam that all pupils take at age 16. It used to have a rather academic character, but has become slightly broader in 1991. A-level exams are taken by pupils at age 18 who have continued their education after the O-level exam. This exam is traditionally academic and qualifies pupils for a university education.

Table 2.1: Definitions of performance used

	Reduced Form		*Structural Model*	
Performance category	*Number*	*Percentage*	*Number*	*Percentage*
Earnings	61	52.1	22	81.5
Exit	15	12.8	2	7.4
Duration	8	6.8	0	0
Survival	17	14.5	0	0
Growth	9	7.7	2	7.4
Profit	4	3.4	0	0
Perceived odds of success	1	0.8	0	0
Employment created	2	1.7	1	3.7
Total	117	100	27	100

Table 2.2: Definitions of education used

	Reduced Form			*StructuralModel*		
	Entry	*Stock*	*Performance*	*Entry*	*Stock*	*Performance*
Years of education	18	30	41	2	1	14
Dummy: high school dropout	20	11	28	0	0	3
Dummy: high school graduate	26	38	21	2	0	10
Dummy: college dropout	30	7	40	1	0	7
Dummy: college graduate	42	51	64	5	0	7
Dummy: postgraduate	19	6	27	2	0	1
Dummy: O-levels (UK)	3	1	8	1	0	3
Dummy: A-levels (UK)	3	14	8	1	0	3

The fact that there is no agreement on the exact definition of entrepreneurship and on the measurement of entrepreneurial performance and education points to the need to be very careful in the design of the conceptual framework that synthesizes the available evidence in this field of research.

Apart from the lack of agreement on the definitions of the key variables, a comparison of the compiled studies also indicates a lack of common tools and techniques. Studies differ substantially in the selection of control variables that enters the relationship between schooling and entrepreneurship outcomes; this variation may be a cause of the differences in the estimated schooling effects across studies. Moreover, stock and entry studies consider varying sets of activities as alternatives to entrepreneurship: Some studies distinguish the choice for entrepreneurship (including or excluding farmers) from the choice for employed labor, whereas others include the unemployed in the alternative choice set. A small number of studies use a multinomial approach, simultaneously distinguishing several alternative labor market positions such as employment in the public and private sector. Some studies examine entrepreneurship performance in isolation; other studies take account of the endogeneity of the choice for entrepreneurship when estimating the effect of education on entrepreneurship performance. In the regression analysis reported in Section 2.6, the latter studies are included under the heading of "structural models".

2.5 The effect of schooling on entrepreneurial outcomes: Descriptive analysis

The lack of standardization creates variability in the results among the studies that examine the relationship between entrepreneurial outcomes and education. This section examines the general direction of the relationship and exposes the degree of variation in the estimates in a descriptive way; the next section turns to an analysis of variance of the studies in the database and examines causes of the variability in the estimates.

2.5.1 Effect of schooling

The relationship between schooling and entrepreneurship entry/performance is reflected in the various estimated coefficients (i.e., b). Table 2.3 shows the percentages of the estimated effects of schooling on entry, stock and performance respectively, that are significantly positive, insignificantly positive, insignificantly negative, and significantly negative. The table distinguishes eight schooling measures: years of education and seven dummy variables of which the first five and the last (British) two are frequently used in combination.[11]

The table shows that the effect of schooling on entry is mostly insignificant. There are two exceptions: the effect of college dropout shows a significantly positive coefficient in 42 percent of the cases. This might be interpreted as confirmation of the "Bill Gates" effect that dropping out of the regular schooling system might be a common thing to do for a nascent entrepreneur, but an alternative interpretation is equally valid, namely that due to screening in the wage sector, college dropouts are pushed into the alternative route of entrepreneurship. The other exception is the significant positive effect on entry of postgraduate training that appears in as many as 52 percent of studies. This may represent highly educated professionals (doctors, lawyers, financial planners) who set up their independent practice after completion of their study.

[11] The estimated effects of the dummy variables for schooling are calculated in more than 95 percent of the cases relative to a reference category of no or primary schooling.

Table 2.3: Effects of education on entrepreneurship entry and performance

	Entry	Stock	Performance
Years of education			
% negative coefficients ($t < -1.96$)	5.0%	19.4%	1.8%
% insignificant negative coefficients	30.0%	12.9%	10.9%
% insignificant positive coefficients	45.0%	16.1%	20.0%
% positive coefficients ($t > 1.96$)	20.0%	51.6%	67.3%
Total # observations	20	31	55
High school dropout			
% negative coefficients ($t < -1.96$)	10.0%	18.2%	25.8%
% insignificant negative coefficients	60.0%	54.5%	51.6%
% insignificant positive coefficients	10.0%	18.2%	12.9%
% positive coefficients ($t > 1.96$)	20.0%	9.1%	9.7%
Total # observations	20	11	31
High school graduate			
% negative coefficients ($t < -1.96$)	14.3%	13.2%	9.7%
% insignificant negative coefficients	46.4%	34.2%	32.3%
% insignificant positive coefficients	14.3%	13.2%	25.8%
% positive coefficients ($t > 1.96$)	25.0%	39.4%	32.2%
Total # observations	28	38	31
College dropout			
% negative coefficients ($t < -1.96$)	6.5%	42.8%	4.3%
% insignificant negative coefficients	32.3%	0%	55.2%
% insignificant positive coefficients	19.4%	28.6%	12.8%
% positive coefficients ($t > 1.96$)	41.8%	28.6%	27.7%
Total # observations	31	7	47
College graduate			
% negative coefficients ($t < -1.96$)	14.9%	9.8%	0%
% insignificant negative coefficients	23.4%	25.5%	16.9%
% insignificant positive coefficients	34.0%	25.5%	11.3%
% positive coefficients ($t > 1.96$)	27.7%	39.2%	71.8%
Total # observations	47	51	71
Postgraduate			
% negative coefficients ($t < -1.96$)	4.8%	66.7%	0%
% insignificant negative coefficients	23.8%	0%	10.7%
% insignificant positive coefficients	19.1%	33.3%	3.6%
% positive coefficients ($t > 1.96$)	52.4%	0%	85.7%
Total # observations	21	6	28
O-level			
% negative coefficients ($t < -1.96$)	*	*	0%
% insignificant negative coefficients	*	*	36.4%
% insignificant positive coefficients	*	*	36.4%
% positive coefficients ($t > 1.96$)	*	*	27.2%
Total # observations	4	1	11
A-level			
% negative coefficients ($t < -1.96$)	*	7.1%	0%
% insignificant negative coefficients	*	28.6%	9.1%
% insignificant positive coefficients	*	57.2%	63.6%
% positive coefficients ($t > 1.96$)	*	7.1%	27.3%
Total # observations	4	14	11

* Statistics are not reported, because the number of observations is five or fewer.

Performance in the entrepreneurial sector has a clear significantly positive relationship with schooling: the higher the type of schooling (i.e., college graduate and postgraduate) and the more years of education, the higher the chances of better performance.[12] Some researchers, especially those adhering to the strong screening hypothesis, would be surprised by the positive relationship (see the next subsection). The positive relationship is not found with regard to the British school system that works with A-levels and O-levels, perhaps as a consequence of the curriculum offered.

While Table 2.3 summarizes evidence on whether education matters, an equally important question is how large the impact of education really was estimated to be. Such a quantitative assessment requires subsets of estimates of b that are comparable in that they derive from empirical models that use the same combination of entrepreneurial outcome (entry, stock, or specific performance measure) and educational attainment. To ensure a degree of reliability, we also require these subsets to contain at least 30 studies. At this point, the lack of uniformity in measurement and definition becomes critical and, as should already be clear from the discussion above, forces us to discard more than a few studies from further analysis. For example, there are not enough studies that examine entry on the basis of years of schooling to warrant the computation of an average effect.[13]

Table 2.4 shows the assembled subsets of studies. The second column, labeled N(1), lists the number of studies included in the relevant subset. Columns three and four indicate the education and outcome variables that form the basis of the subset. The subsets are divided into two groups. The first group allows a numerical comparison of estimation results.[14] The second group, shown in Table 2.4, only permits an ordinal analysis of the estimation results, as will be explained in the next section.

[12] Note that performance consists of various different performance measures. For each of these performance measures we find a positive effect of education.

[13] A further requirement is that the estimation method must indeed be comparable. Estimates of logit, probit, and linear probability models are readily transformed to comparable magnitudes. Performance studies are not straightforwardly comparable. For example, some studies examine earnings in logarithmic form whereas other studies specify a linear variable. Earnings may refer to hourly, weekly, monthly, or annual values. They may reflect before or after tax values, and they are expressed in the local currency. Even so, all of these estimation results might be made comparable by expressing the education effect in elasticity form or by standardizing the parameter estimates with the standard deviation of earnings, but that requires descriptive statistics on earnings and education values that are often not reported in the studies. In fact, no more than 30 studies in the entire database revealed the standard deviation and the mean of the performance measure.

[14] Subset IV only consists of 21 observations. Here, the minimum-30-study rule is violated because this subset offers the best comparison to the established returns-to-schooling estimates of the employees literature.

Table 2.4: Subsets of studies suitable for meta analysis

						Weighted	Sign of effect		
Subset	N(1)	Education	Entrepreneurial outcome	N(2)	Average b (s.d.)	Average b (s.d.)*	neg.	insig.	pos.
A. Suitable for numerical analysis									
I	40	College grad	Entry	33	-.004 (.299)	-.029 (.155)	18%	67%	15%
II	37	High school grad	Stock	31	-.031 (.132)	-.026 (.107)	16%	42%	42%
III	50	College grad	Stock	43	.018 (.144)	.0195 (.018)	12%	47%	42%
IV	21	Years	Performance: log earnings	21	.061 (.013)	.054 (.028)	0%	14%	86%
B. Suitable for ordinal analysis									
V	37	Years	Performance: earnings	34			0%	9%	91%
VI	43	College grad	Performance: earnings	43			0%	16%	84%
VII	31	Years	Stock	27			11%	33%	56%
VIII	55	Years	Performance: all combined	52			2%	31%	67%
IX	31	High school grad	Performance: all combined	31			19%	58%	23%
X	47	College dropout	Performance: all combined	47			15%	68%	17%
XI	71	College grad	Performance: all combined	69			6%	29%	65%

N(1) indicates the number of studies before duplicate sources are dropped
N(2) indicates the number of studies after duplicate sources have been dropped (see text)
*Average computed by method of Hedges (1992), which takes the quality of the studies into account (by weighting by their standard error)

Table 2.4 complements Table 2.3: it shows the percentages of positive, negative, and insignificant effects for the defined subsets. For subsets I to IV, which allow quantitative analyses, the averages and standard deviations of the estimated effect of schooling are given as well.

The table shows that the effect of education on entry is ambiguous. In contrast, the effect of schooling on performance, broadly measured, is unambiguously positive: the marginal effect of an extra year of education on performance is significantly positive in 67% of the studies (Subset VIII). The estimated effects are especially pronounced when studying the effect of education on the more specialized performance measure of "earnings": between 84 and 91 percent of the studies find a significantly positive relationship, whereas none of them reveal a significantly negative relationship. The main result from the numerical analysis is that the average return to schooling in subset IV (which, as it happens, mostly includes USA studies (81%)) is 6.1 percent.

Using the method of Hedges (1992) we recalculated the average return to education. The Hedges method takes the precision (standard deviation) of each study's estimate into account by weighing. Table 2.4 shows that the returns to education are somewhat lower when using the Hedges correction, namely 5.4%.

Altogether, the returns to education for entrepreneurs resulting from this comparative analysis are somewhat lower than the 7 to 9 percent commonly found in the wage sector (see Ashenfelter et al., 1999).

2.5.2 Returns to schooling for entrepreneurs in comparison to employees

A more direct comparison of the returns to education for entrepreneurs and employees can be obtained from a separate analysis of some studies in the meta-analysis database. Twenty studies have measured the returns to education for entrepreneurs and employees in a comparable fashion (See Table 2.5).

Table 2.5: Comparing the effect of education on the performance of employees and entrepreneurs

Author	Dependent Variable	Education	Sample	Country	Entr > Empl *	Screen**
Brown & Sessions 1998	Log hourly earnings	Dummies	Male	Italy	No	No
De Wit & Van Winden 1989	Log net earnings per hour	Years	Both	NL	No	No
Alba-Ramirez 1994	Log net monthly earnings	Dummies	FTM	Spain	Equal	No
Brown 1998	Log hourly earnings	Dummies	Male	UK	No	Yes
Rees & Shah 1986	Log annual earnings	Years	Male	UK	No	No
Dolton & Makepeace 1990	Log salary	Dummies	Both	UK	No	No
Taylor 1996	Log gross hourly earnings	Dummies	Male	UK	No	No
Lombard 2001	Log hourly wage	Dummies	MW	USA	Yes	No
Tucker 1985	Log annual labor income	Years	Both	USA	Yes	Yes
Tucker 1987	Wage	Dummies	Male	USA	Equal	Yes
Evans & Leighton 1990	Log earnings	Years	Male	USA	Yes	No
Fredland & Little 1981	Earnings	Years	Male	USA	Yes/Equal	Yes
Lofstrom 2002	Weekly wage	Years	HPIS	USA	Yes	No
Fairly & Meyer 1996	Log fulltime earnings	Dummies	HPIS	USA	Yes	No
Robinson & Sexton 1994	Annual earnings	Years	Both	USA	Yes	No
Macpherson 1988	Hourly wage	Years	MW	USA	Yes	No
Borjas & Bronars 1989	Log weekly income	Dummies	HPIS	USA	Yes	No
Gill 1988	Log hourly wage	Years	Both	USA	Yes	No
Moore 1983	Annual earnings	Years	Both	USA	Yes	No
Simpson & Sproule 1998	Log pre-tax annual earnings	Dummies	Both	Canada	Yes	No

*This column denotes whether the education coefficient is larger for entrepreneurs than for employees. **This column denotes whether the screening hypothesis is tested explicitly. NL=Netherlands, HPIS=High Percentage Immigrants in Sample, FTM=Full Time Male, MW=Married Women.

Less than half of these twenty studies focuses on the screening function of education. One of the ways in which the (strong or weak version of the) screening hypothesis is tested empirically is to compare the returns to education for employees to the returns for entrepreneurs, where the latter group is considered as an unscreened control group. Almost all screening studies reject the strong screening hypothesis: i.e. these studies find positive returns to education for entrepreneurs. However, the evidence related to the weak screening hypothesis is mixed. Studies based on USA data reject the weak screening hypothesis, (Robinson and Sexton, 1994; Evans and Leighton, 1990; Tucker, 1987, 1985; Fredland and Little, 1981), implying that the returns to education are not higher for employees than they are for entrepreneurs in the United States. Studies using European data (UK, Italy, and The Netherlands) support the weak screening hypothesis (Brown and Sessions, 1999, 1998; De Wit and Van Winden, 1989; Rees and Shah, 1986). The latter result implies that the returns to education are (slightly) lower for entrepreneurs than for employees in Europe.

The majority of the twenty studies that compare returns to education for entrepreneurs with those returns for employees use the comparison to highlight differences in labor market participation and success factors between minorities and non-minorities and/or between females and males (e.g., Lofstrom, 2002; Lombard, 2001; Fairlie and Meyer, 1996; Borjas and Bronars, 1989; Gill, 1988; MacPherson, 1988; Moore, 1983). The results from these (exclusively USA) studies are consistent with the results obtained in the screening literature: the estimated returns to education for entrepreneurs are at least as high as –and usually (slightly) higher than– for employees.

The summary of these studies fails to support a conclusion that the returns to education are higher for employees than they are for entrepreneurs, as one would gather from comparing the descriptive statistics of Table 2.4 with the meta-analysis of Ashenfelter et al. (1999). Studies pertaining to Europe may indicate that the returns to education are slightly lower for entrepreneurs than for employees, but the opposite result is found for the studies that pertain to the United States. The statistics of Table 2.4 and Ashenfelter et al. (1999) each summarize a larger body of literature, but the pairs of estimates in the 20 studies used here are methodologically more homogeneous by design and therefore have substantive value.

2.6 The effect of schooling on entrepreneurial outcomes: Regression analysis

2.6.1 The regression approach

The descriptive statistics in the previous section provided meaningful insights into the impact of education on entrepreneurship selection and performance. However, equally important is to find out which characteristics of the studies influence the descriptive results found. To this end we estimate the model $b = \alpha Z + \epsilon$ as described in Section 2.3. In these regressions, we control for the quality of the estimate that each study contributes by means of the Hedges (1992) approach.

Several notes are in order to ensure that the estimation results from this model are meaningful. First, the models must be estimated across a subset of comparable studies: the result of the meta-analysis will be more meaningful when there is less variation across studies in the measurement of schooling attainment and entrepreneurial outcome (including the specific methodologies used). However, the advantage of homogeneous definitions has to be traded off with the size of the resulting sample for each of the regressions. Table 2.4, Panel A, shows the subsets that are used for (separate) regression analyses.

Second, a subset is suitable for regression analysis only if it meets a particular statistical requirement, namely independence of observations. In the meta-analytical regression context, a regression equation contained in a study constitutes an "observation", and a properly constructed subset of observations makes up a "sample." Thus, at issue is whether the observations included in the sample offer independent measurements of the impact of education. A detailed examination of the studies reveals the potential for a violation of the independence assumption across observations that use comparable samples from the same (publicly available) dataset.[15] There are 16 studies,[16] involving a total of 37 observations, that use the same data source and (roughly) the same subsample to examine a particular entrepreneurial outcome variable. We have randomly kept one observation of each of the 16

[15]In many cases, multiple observations of the same type (for instance, performance) come from one publication that uses a single dataset. As was indicated, to preserve independence, only one of the estimated models is selected, in particular the one that best shows the estimated relationship. But if one publication presents separate estimates for, e.g., men and women, both are included in the database, since independence is not in danger. It must be acknowledged that two estimates pertaining to different subsamples from one study (or from two studies by the same author) might still be correlated statistically, from the perspective of a meta-analysis, because these two estimates come from the same source (author(s)). Our use of the term "independence of observations' pertains to the statistical independence of the samples that generate the estimated education effects.

[16]Available upon request.

problematic combinations and deleted the other 21 observations. This deletion disqualifies sub-sample VII in Table 2.4 for a separate analysis since its size now becomes smaller than 30. Moreover, subset IV is not utilized due to its small size. Nine subsets remain: three allowing a numerical analysis of the schooling coefficients found, and six requiring an ordinal analysis of the direction of effects of schooling on entry or performance. The column in Table 2.4, labeled N(2), shows the number of independent studies in each subset that will be used in the meta-analytical regression models. Subsets I to III each consist of studies that use exactly the same methodology for establishing the effect of the exactly identical educational measure on the uniquely defined phenomena studied (entry in subset I and stock in subsets II and III). Since the coefficients for education established in these narrowly defined samples are numerically comparable to each other and vary as a continuous variable over a whole range of values, the effects of the potential determinants of variation can be estimated by OLS.

Third, note that the remaining six subsets (V - XI, omitting VII) are not sufficiently homogeneous and therefore need to be processed differently. Thus, the estimates of b cannot be made comparable and estimation of a regression model $b = \alpha Z + \epsilon$ is meaningless. However, there is a way to pool such small subsets in meaningful ways. Namely, t-statistics of b reflect the sign and significance of the estimated relationship, where it does not matter so much whether the dependent variable is measured in linear or logarithmic form. Better yet, one may pool across all forms of performance measures, as long as the parameter estimates and t-statistics are recorded in such a way that the hypothesized effect of education points in the same direction. For performance measures for which "the more, the better" does not hold (i.e., exit from self-employment and the hazard out of self-employment, where "more" is worse), the sign of the t-statistic simply needs to be reversed. On the basis of this recoded t-statistic, we define the variable t^* as taking on the value 0 for observations that have established a significantly negative effect, 1 for studies that find an insignificantly negative effect, 2 for studies that show an insignificantly positive effect, and 3 for significantly positive effects. Finally, we regress t^* on characteristics of the studies by means of an ordered probit model: $t^* = \alpha Z + \epsilon$. A positive estimate of the parameter α indicates that an increase in characteristic Z makes it more likely (or less unlikely) that education raises the entrepreneurial outcome under consideration.

2.6.2 Sources of variation in b

What variables should be included in Z? Since subset sizes are not particularly large, meta-analytical models must be specified parsimoniously. Subset sizes between 30 and 71 do not permit rich models, and the estimated effects require a cautious interpretation.

Actually, theory provides little guidance in generating hypotheses about the determinants of the returns to schooling for entrepreneurs, nor about the determinants of the relationship between entry and schooling. This analysis therefore has an explorative character and provides some first answers to the following kinds of questions. Is the return to schooling or the effect of schooling on entry, for example, higher in the USA than elsewhere, higher for men than for women, or higher for whites than for non-whites? Is the return to schooling diminishing or increasing over time? Does the performance measure selected affect the return to schooling result? Moreover, is there any distinction in the effect of education across the various schooling levels? And does the relationship vary by the scientific level or by the type of journal (e.g. economics, management, etc.) in which the study is published? Is there something such as a publication or reporting bias in the sense that there is an over-representation of significant results? Does the elimination of self-selection bias by estimating a structural model instead of a reduced form model affect the relationship between schooling and performance or schooling and entry? With this in mind, let us turn to a description of the explanatory variables of the meta-analytical models, i.e., the components of the Z vector.

Sample characteristics of observations

The first group of control dummy variables is meant to control for three sample characteristics: whether the sample is taken from the USA or from some other country,[17] the percentage of females in the sample, and the percentage of non-white (including immigrants) individuals.[18] These explanatory variables have been dropped from the analyses whenever the cross-sectional variation within a subset was too small, for instance when all but one study were from the USA.

Time

A variable is included that indicates the (earliest) year from which the observations in the sample have been drawn. In specifications not included here, we also tested whether there is

[17]Multi-country studies that are shown in Table 2.A-1 always split their samples into various single country observations. These are the observations used.

[18]Sometimes, when the sample consisted of a mix of males and females and/or of a mix of native inhabitants and immigrants, and when the study did not supply the relevant statistics, OECD statistics of the relevant country are inserted. To assess whether the effect of this inserted variable is different from the effect of the individual study-specific numbers, each regression model includes a dummy variable that is 0 for "actual sample statistic" and 1 for "inserted OECD country value". The dummy was not significant and is therefore not reported in the table of estimates reported below.

any relationship between the dependent variables of the analyses and the year of publication of each of the studies. This was not the case.

Characteristics of journals

The number of unpublished observations in each of the subsets is insufficient to warrant the inclusion of a separate "publication" dummy. Preliminary work did not indicate any effect of the branch of journal in which a study is published (see Appendix 2.A). These dummies were omitted. The impact factor of the journal is included in the equations as an indicator of the quality of the study itself. If all better quality studies report higher (or lower) schooling effects, then there is reason to believe that the "true" effect is indeed higher (or lower) than what a simple average (such as Table 2.4) suggests.[19]

Control variables

Table 2.B-1 in Appendix 2.B shows the most commonly used control variables and their effects. We investigated whether the inclusion of these variables in the observed studies affected the effect of schooling on entry, stock, and performance: quod non. Similarly, the effect of specific schooling dummies (for instance college graduate) is not affected by the inclusion of other schooling dummies such as college dropout or postgraduate schooling. Indicators of the presence of these control variables are therefore omitted from the final regression model.

Publication/Reporting bias

As in Ashenfelter et al. (1999), Roberts (2005), and Stanley (2005), the observed universe of published results may have emerged solely because they were statistically significant. It might be the case that studies remain unpublished because they failed to provide a statistically significant rejection of the null-hypothesis of no effect. It is also possible that authors drop insignificant determinants from their models, and that therefore a statistically insignificant education effect leads to the omission of all education variables. If so, those studies would not appear in the database. The method of Hedges (1992) is suitable for testing and controlling

[19]The impact factor is based on the objective ranking, the Social Sciences Citation Index (SSCI), annually published by the Institute for Scientific Information (ISI). This index gives for thousands of journals the number of citations each year. This leads to an impact factor for each journal, using the following formula: impact factor year X = (Cites in year X to articles published in year X-1 and in year X-2)/(Number of articles published in year X-1 and in year X-2). Unpublished studies or studies included in journals that do not (yet) have an impact factor are assigned an impact factor of zero. The small number of unpublished studies, which of course do not have an impact factor yet, does not warrant a separate dummy.

for this so called publication or reporting bias of the schooling measures. Controlling for the quality of the studies, in terms of the measured standard errors, this method tests if studies with a p-value higher than 0.05 or between 0.01 and 0.05 have an equal probability of being observed as studies with p-values lower than 0.01. Studies with $p < 0.01$ are assumed to be observed with a probability equal to one. If the probability of being observed for the $p > 0.05$ and $0.01 < p < 0.05$ studies is not equal to one there is evidence of publication bias.

Unlike in the first three subsets in which the dependent variable is the schooling coefficient itself, the Hedges (1992) method cannot and does not need to be used in the latter six subsets. The dependent variables in these subsets are based on t-values, which prevents the application of the Hedges method.[20] The focus on t-values inherently controls for the quality of the studies in terms of the measured standard errors. However, it is not possible to check for the presence of publication bias in these subsets.

Performance measures used

The last four subsets include all types of performance measures. These equations control for the types of performance measures used in order to assess whether the impact of schooling differs between these measures. After some preliminary attempts, it proved to be sufficient to differentiate among these variables with a dummy variable that indicates earnings-related performance measures.[21]

Estimation methods and types of data

The method of estimation may influence the magnitude of the education effect. Thus, the estimation method is indicated through a dummy variable indicating whether a structural probit has been estimated (in subset I) or whether a structural earnings model has been used (in subsets V and VI). Since the practice of structural model estimation is almost uniquely present in entry and earnings studies, this control variable is omitted in all other subsets.

Another control dummy for whether a panel dataset has been used or not appeared to have no effect in any of the subsets, not even the subset referring to entry studies where the absence of panel data must imply the (often thought to be inferior) use of cross-sectional data retrieved from questionnaires that ask ex-post about individuals' labor market histories.

[20] The Hedges method weighs the coefficients by their standard deviation, as seen in Table 2.4.

[21] The rest of the performance measures are for the most part related to survival. This makes the inclusion of another performance dummy impossible as the reference group would then be too small.

2.6.3 Results from the meta-analysis

Table 2.6 reports the estimates resulting from the meta-analytical models that examine the determinants of the variation in the estimated impact of educational attainment on the entrepreneurial selection and performance outcomes.

Schooling and entry (subset I)

The effect of college graduation on the probability of selection into an entrepreneurial position has several observable significant determinants. The effect is higher in the United States than elsewhere, implying either an educational environment that is more conducive for entrepreneurship development or better business conditions that attract more highly educated individuals. The education effect is the same for males and females, and there are no racial differences either. The effect of education on entry increases over time indicating that the stimulus generated by schooling has been rising over time.

The first equation in Table 2.6 furthermore shows that estimating a structural probit model gives rise to a lower estimate of the effect of education on entry. Thus, part of the overall education effect on entry works through income differentials (typically measured as wage earnings minus entrepreneur earnings). When discussing the policy implications of the effect of education on entrepreneurship, it is important to realize that it is the direct effect (corrected for income), not the overall effect, that is the most important. This direct effect is estimated with bias when one does not account for expected income differences. Application of the Hedges method indicates that there is no publication bias among the set of studies measuring the effect of education on entry.

Schooling and stock (subsets II and III)

The size of the effect of schooling (high school and college graduation in the second and third subset respectively) on the probability of being an entrepreneur varies systematically with several of the included study characteristics. The most interesting is the USA dummy: The schooling effect in the USA is smaller than elsewhere, although the difference is only marginally significant for college graduates and not significant for high school graduates. While the USA business climate seems to encourage the educated to start a business (see subset I), it does not seem to be able to retain the same people when they are entrepreneur.

Table 2.6 furthermore shows that the high school effect is more positive for females. For both high school and college graduate effects, the impact factor shows up significantly positive, suggesting that more rigorous publication outlets report somewhat more positive

Table 2.6: Determinants of the effect of schooling on entrepreneurship

Subsample:	*I*	*II*	*III*	*V*	*VI*	*VIII*	*IX*	*X*	*XI*
Type of education variable	College grad.	High school grad.	College grad.	Years of schooling	College grad.	Years of schooling	High school grad.	College dropout	College grad.
Type of entrepreneurship variable	Entry	Stock	Stock	Earnings	Earnings	Perform.	Perform.	Perform.	Perform.
USA	0.22**	-0.14*	-0.05		0.35	0.55	1.58*	1.18*	1.53**
	(3.71)	(1.92)	(.85)		(0.29)	(0.10)	(1.67)	(1.92)	(3.32)
% Female	0.10	0.12**	0.04		-0.34	4.71**	-0.28	1.08	-0.74
	(1.64)	(2.3)	(0.65)		(0.21)	(2.40)	(0.31)	(1.21)	(1.20)
% Non-white	-0.01	0.01	0.01		-0.29**	-0.16**	0.20**	-0.06	-0.08*
	(1.06)	(0.9)	(0.71)		(3.12)	(2.39)	(2.59)	(0.83)	(1.9)
Year of earliest	0.01**	-0.01**	0.01	0.08*	-0.10	-0.03	0.13**	0.05	0.04
observation	(2.69)	(3.64)	(0.61)	(1.87)	(1.33)	(1.38)	(2.55)	(1.28)	(1.24)
Journal impact	0.09	0.12**	0.08**	-0.45			-1.25**	0.83	0.31
factor‡	(1.45)	(2.86)	(2.30)	(1.27)			(-2.00)	(1.42)	(0.80)
Performance measured						2.28**	1.74**	-0.15	0.77*
as earnings						(3.83)	(3.16)	(0.25)	(1.93)
Structural model	-0.19**			-0.15	-0.17				
	(2.00)			(1.27)	(0.14)				
Publication bias found†	no	no	no						
Sample size	33	31	43	34	43	52	31	47	69
Adj(Pseudo)R2				0.21	0.45	0.38	0.24	0.09	0.3
Meta-analysis regression method	ML (Hedges)	ML (Hedges)	ML (Hedges)	OrdProb	OrdProb	OrdProb	OrdProb	OrdProb	OrdProb

†Hedges method (1992). ‡Unpublished articles and books are assigned an impact factor of zero. The stars in the table indicate the significance levels of the coefficients; *implies a p-value lower than 0.10, **implies a p-value below 0.05. T-values are shown in brackets.

schooling effects. Even so, we find no evidence of publication bias.

Schooling and performance (subsets V-IX)

Table 2.6 shows that the probability of positive and significant performance-related returns to education is generally higher in the USA than elsewhere. This might be due to a schooling system that is more oriented towards teaching entrepreneurship skills and knowledge, or to better business conditions that benefit the more highly educated more in the USA than elsewhere.

The link between education and performance seems to be stronger for women than for men: The larger the percentage of females in the sample, the larger is the probability that the effect of years of schooling on performance (Subset VIII: all measures combined) is significantly positive.[22] However, regression results with the other performance subsets render no significant difference between males and females. Non-whites or minorities benefit less from their education given the selection into an entrepreneurial position: The higher the percentage minorities in the sample, the lower the likelihood that the effect of education on business performance is positive.[23] This latter effect might be caused by the fact that the group of non-whites often includes immigrants who have pursued (part of) their education abroad with a probably smaller directly productive effect in the country of immigration.

There is some evidence of an increasing trend in the return to schooling over time. Apparently, education has become more worthwhile over the years to achieve superior performance as an entrepreneur. This is in line with the dramatic shifts in industry and occupation compositions and technological advances over the last two decades, which obviously favor educated aspirants for entrepreneurship.

When the performance measure is earnings, rather than survival, duration, or growth, the effect of education is substantially stronger: The effect measured in subsets VIII and IX is statistically significant, and subset XI shows a positive effect as well, though only marginally significant. Apparently, education is more influential with respect to earnings. The effect of the impact factor is mixed.

[22] See Table 2.1 for the different performance measures that are combined.

[23] There is one exception: The relationship between high school graduation and performance.

2.7 Conclusion, evaluation and suggestions for further research

This chapter's objective has been to give an overview of the literature on the relationship between entrepreneurship selection and performance on the one hand and formal schooling on the other hand. We quantify the existing studies and summarize them by the application of meta-analysis. In this manner we examine on the basis of the heterogeneous evidence whether and to what extent schooling is related to entrepreneurship outcomes and what factors determine the variation of this relationship across studies.

In all, the analysis leads to the following conclusions. There is no evidence of a systematic relationship between an individual's schooling level and the probability of selection into entrepreneurship. This does not contradict economic theory, which points out that education has two opposite effects on entrepreneurship entry.

Moreover, the relationship between schooling and performance is significant and positive, in line with economic theory. To formulate more precisely, the higher the schooling level or the more years of education have been pursued, the higher are the chances that performance is good: earnings are higher, growth is more likely, survival chances are better. For the specific relationship between years of schooling and (log annual dollar) earnings as an entrepreneur in the United States, the average return to a year of schooling pursued is 6.1 percent.[24] This return seems to be smaller than the usual comparable estimates for employees that are between seven and nine percent.

However, there are 20 studies in the database that estimated the returns for both entrepreneurs and employees in comparable ways and may therefore be used to examine the difference in the returns to education for entrepreneurs and for employees more closely. This yields the third conclusion: It is not clear from these studies that the returns to education are uniformly higher for employees than for entrepreneurs. More specifically, all studies pertaining to Europe indicate that the returns to education are slightly lower for entrepreneurs than for employees. However, the opposite result is found for the studies that pertain to the United States.

Meta-analytical regression models investigate whether the estimated return to education is affected by characteristics of the studies such as the nature of the sample, the estimation method, the type of publication outlet, and the usage of control variables. The evidence indicates that the effect on entry is higher in the United States than elsewhere. Moreover,

[24] The average of 6.1% is made up of a small number of studies that analyze an other country than the USA i.e., 19% of the studies do not analyze the USA.

the likelihood of a positive and significant effect of education on earnings is also higher in the USA than elsewhere. The relationship between education and performance is more likely to be positive for females than for males. For non-whites or immigrants this relationship is less likely to be significantly positive than for whites. Furthermore, in line with expectations, the effect of schooling on entry and performance has increased over the past decades. Moreover, the chance is higher that the effect of educational attainment is significant and positive for earnings-related performance measures than for other entrepreneurial performance measures such as survival, duration or growth. Estimation methods used do matter in the instance of entrepreneurship selection: Structural probit estimation that account for potential earnings differences between entrepreneurship and employment decreases the estimated effect of education on entrepreneurship entry.

We do not find any indication that publication/reporting bias exists in studies of the relationship between schooling and entrepreneurship outcomes, whereas Ashenfelter et al. (1999) concluded that the literature on the relationship between schooling and employee incomes may be overstating the effect of education.

2.7.1 Evaluation of the "State of the Art" and suggestions for further research

Implementing a meta-analysis forces one to closely compare essential features of the many studies in the education-entrepreneurship literature and thus to outline the state of the art of research into the effect of education on selection into entrepreneurship and performance as an entrepreneur. Though much research effort has been directed towards the issue, we have found many "black holes", issues that either have not been addressed at all or were not addressed in a satisfactory manner. We benchmark against the common practice in the returns to schooling in employment literature.

A first drawback is the lack of homogeneity in definitions used of both schooling and performance. Only 35 percent of the studies uses "years of schooling" as the measure of educational attainment. Most researchers utilize a set of dummy variables for specific levels of schooling that is specific to the educational system in a given country. This may have some value: (i) education effects may be nonlinear, and (ii) the dummy variables may highlight specific curricular features of the educational system. Nevertheless, such differentiation in the measurement of schooling renders a comparison across studies nearly impossible. It is therefore advisable to supplement such detailed dummy specifications with more conventional years of schooling models.

The same level of differentiation is found with respect to the measurement of performance: various (by themselves useful) definitions of performance are in fashion. When a less traditional measure of entrepreneurial performance is examined, it would be useful to also study the more standard performance variables. For example, when one finds that enterprise survival responds more to schooling than earnings in a given sample, one obtains a richer understanding of the notion of performance than when only the survival results are reported.

A second issue that should receive more attention in the literature is the role that ability and other often unobserved factors might play in determining entrepreneurial selection and performance. It is quite plausible that the "effect" of schooling that is typically estimated as being a (conditional) correlation is not completely causal: Ability and other factors might increase performance and also lead to more schooling, thus leading to a spurious positive effect of schooling on performance. A deeper theoretical concern is that schooling itself is endogenous to one's performance in the labor market. Although future earnings are not the only reason to pursue an education, the prospect of earning higher incomes entices many students to stay in school longer. In the established returns-to-schooling literature that focuses on wage employment, this issue is well recognized, and whenever the data permit, researchers attempt to correct for the ability bias and the endogeneity of schooling by including measures of innate ability into the specifications, by using instrumental variables of one type or another, by running controlled experiments or by studying twins. This is not at all the case in the entrepreneurship counterpart of this literature. Ability controls are hardly used, and none of the studies made reference to the endogeneity of schooling. A major challenge in entrepreneurship research will be to perform these types of analyses and to find appropriate ways to endogenize the decision to pursue schooling. Webbink (2005) provides an excellent overview of the latest state of affairs in the employee literature concerning the research methods used for studying the effectiveness of (various forms of) education.

A third observed issue is that only 12 percent of the studies that we surveyed corrected for selection biases; the vast majority paid no attention to selectivity issues. Omission of such a correction should at least be acknowledged in the type of recommendations these studies put forth. The results from the meta-regression analysis indicate that failure to account for selectivity leads to overestimation of the effect of education on entrepreneurship entry.

There are other lacunae in our understanding of the role of human capital in entrepreneurship as well. Little is known about the occurrence and effect of specific training. Furthermore, the literature demonstrates a lack of attention paid to the effect of various types of schooling on entrepreneurship entry and performance. It may well be true that the type of curriculum is more important than the level of schooling pursued, but this issue is considered in only a small number of studies pertaining to different countries. It is quite conceivable that both curriculum and level of schooling impact entrepreneurial outcomes through productivity (human capital theory) and sorting (screening or signaling theory; see Riley (2002) and Wolpin (1977)). To disentangle this represents another challenge.

Notable in this respect is that Lazear (2004) has recently initiated a new strand of research based on his "Jack-of-all-Trades" theory. The idea of the theory is that entrepreneurs, in their capacity as "Jacks-of-all-Trades", may require a broad mix of skills, possibly obtained through formal education, rather than a certain level of education. Thus, besides the level of formal education, the choice of courses or tracks at school could be an important determinant of entrepreneurship outcomes. Lazear (2004), Wagner (2003), and Silva (2007) perform empirical tests that support the idea that entrepreneurs are indeed "Jacks-of-all-Trades". However, on the basis of a panel dataset of graduate students as well as a sample of the general Italian working population, Silva (2007) concludes that the causality of the results established by Lazear and Wagner is the other way around: Entrepreneurs may indeed be generalists, but changes in the breadth of their knowledge across different fields, caused by education or work experience, does not increase the propensity of becoming an entrepreneur.[25]

Moreover, little is known about the (trends in) differences across industries of the (productive and screening) effects of education on performance. Does success as an entrepreneur nowadays require more education than in the past due to the evolution of new more advanced technologies? Did both the emergence of high tech industries and the downsizing in the traditional corporate sector open up new opportunities for highly educated individuals who formerly used to be absorbed into the large-scale industrial corporate sector? Answers to these questions are not yet known.

Finally, entrepreneurship might be described as the process of bringing inputs, technologies, and output markets together. The role of education in this process is insufficiently addressed. More specifically, to what extent and under what conditions are human and financial capital complements to or substitutes of each other in the entrepreneurial process? For example, there is some agreement about the fact that entrepreneurs face a capital constraint that either prevents them from selecting into an entrepreneurial position at all, or else negatively affects their investment opportunities during the start-up phase of their venture. It is quite conceivable that banks and other capital suppliers use education as a means to screen prospective entrepreneurs about whom little other information is observed. In these ways, education will also have an indirect effect on performance and selection into entrepreneurship. The recent study by Parker and Van Praag (2006) finds evidence in support of this.

[25] Chapter 5 explores this issue in more detail.

The accumulated evidence in this literature is by no means conclusive yet. Many challenging complexities remain. A few of these complexities are dealt with in this book. In particular, Chapters 3 and 4 look at the possibility that the "effect" of schooling, as estimated in the studies summarized in this chapter, is not completely causal. Chapter 5 focuses on the returns to intelligence and the "Jack-of-all-Trades" theory, rather than on the returns to formal schooling.

References meta-analysis industrialized countries[26]

Alba-Ramirez, A. (1994), 'Self-employment in the midst of unemployment: the case of Spain and the United States', *Applied Economics* **26**, 189–204.

Bates, T. (1985), 'Entrepreneur human capital endowments and minority business viability', *The Journal of Human Resources* **20**(4), 540–554.

Bates, T. (1990), 'Entrepreneur human capital inputs and small business longevity', *The Review of Economics and Statistics* **72**(4), 551–559.

Bates, T. (1995), 'Self-employment entry across industry groups', *Journal of Business Venturing* **10**, 143–156.

Bates, T. (1999), 'Exiting self-employment: An analysis of Asian immigrant-owned small business', *Small Business Management* **13**, 171–183.

Bates, T. and C. Dunham (1993), 'Asian-american success in self-employment', *Economic Development Quarterly* **7**(2), 199–214.

Bernhardt, I. (1994), 'Comparative advantages in self-employment and paid work', *Canadian Journal of Economics* **27**(2), 273–289.

Blanchflower, D. and B. Meyer (1994), 'A longitudinal analysis of the young self-employed in Australia and the United States', *Small Business Economics* **6**(1), 1–19.

Blanchflower, D. (2000), 'Self-employment in OECD countries', *Labour Economics* **7**, 471–505.

Blanchflower, D., A. Oswald and A. Stutzer (2001), 'Latent entrepreneurship across nations', *European Economic Review* **45**, 680–691.

Blumberg, B. and G. Pfann (1999), 'Social capital and the self-employment decision', *Business Investment Research Center Working Paper* .

Boden, R. (1996), 'Gender and self-employment selection: An empirical assessment', *Journal of Socio-Economics* **25**(6), 671–682.

[26]This reference list includes all studies included in the meta-analysis as available at the time of research (2003). The studies referred to in the text are listed in the main bibliography at the end of this book.

Boden, R. (1999*a*), 'Flexible working hours, family responsibilities, and female self-employment', *American Journal of Economics and Sociology* **58**(1), 71–83.

Boden, R. (1999*b*), 'Gender inequality in wage earnings and female self-employment selection', *Journal of Socio-Economics* **28**, 351–364.

Boden, R. and A. Nucci (2000), 'On the survival prospects of men's and women's new business ventures', *Journal of Business Venturing* **15**, 347–362.

Borjas, G. and S. Bronars (1989), 'Consumer discrimination and self-employment', *Journal of Political Economy* **97**(3), 581–605.

Borjas, G. (1986), 'The self-employment experience of immigrants', *Journal of Human Resources* **21**(4), 485–506.

Bosma, N., M. Van Praag, R. Thurik and G. De Wit(2002), 'The value of human and social capital investments for the business performance of startups', *Small Business Economics* **23**(3), 227–236.

Boyd, R. (1991), 'A contextual analysis of black self-employment in large metropolitan areas, 1970-1980', *Social Forces* **70**(2), 409–429.

Brown, S. and J. Sessions (1998), 'Education, employment status and earnings: A comparative test of the strong screening hypothesis', *Scottisch Journal of Political Economy* **45**(5), 586–591.

Brown, S. and J. Sessions (1999), 'Education and employment status: A test of the strong screening hypothesis in Italy', *Economics of Education Review* **18**(4), 397–404.

Bruce, D. (1999), 'Do husbands matter? Married women entering self-employment', *Small Business Economics* **13**, 317–329.

Bruce, D. (2000), 'Effects of the United States tax system on transitions into self-employment', *Labour Economics* **7**, 545–574.

Bruce, D. (2002), 'Taxes and entrepreneurial endurance: Evidence from the self-employed', *National Tax Journal* **55**(1), 5–24.

Bruce, D., D. Holtz-Eakin and J. Quinn (2002), 'Self-employment and labor market transitions at older ages', *Working paper, Syracuse university* .

Bruederl, J. and P. Preisendörfer (1998), 'Network support and the success of newly founded businesses', *Small Business Economics* **10**(3), 213–225.

Bruederl, J., P. Preisendörfer and R. Ziegler (1992), 'Survival chances of newly founded business organizations', *American Sociological Review* **57**(2), 227–242.

Caputo, R. and A. Dolinsky (2001), ‘Women’s choice to pursue self-employment: The role of financial and human capital of household members’, *Journal of Small Business Management* **36**(3), 8–17.

Carrasco, R. (1999), ‘Transitions to and from self-employment in Spain: An empirical analysis’, *Oxford Bulletin of Economics and Statistics* **61**(3), 315–341.

Carroll, G. and E. Mosakowski (1987), ‘The career dynamics of self-employment’, *Administrative Science Quarterly* **32**(4), 570–589.

Clark, K. and S. Drinkwater (1998), ‘Ethnicity and self-employment in Britain’, *Oxford Bulletin of Economics and Statistics* **60**(3), 383–407.

Clark, K., S. Drinkwater and D. Leslie (1998), ‘Ethnicity and self-employment in Britain 1973-1995’, *Applied Economics Letters* **5**(10), 631–634.

Cooper, A., C. Woo and W. Dunkelberg (1988), ‘Entrepreneurs’ perceived chances for success’, *Journal of Business Venturing* **3**, 97–108.

Cooper, A., J. Gimeno-Gascon and C. Woo (1994), ‘Initial human and financial capital as predictors of new venture performance’, *Journal of Business Venturing* **9**(5), 371–395.

Cooper, A., T. Folta, J. Gimeno-Gascon and C. Woo (1992), ‘Entrepreneurs’ exit decisions: The role of the threshold expectations’, *Academy of Management Proceedings* pp. 75–79.

Cowling, M. (2000), ‘Are entrepreneurs different across countries?’, *Applied Economics Letters* **7**, 785–789.

Cramer, J., J. Hartog, N. Jonker and M. Van Praag (2002), ‘Low risk aversion encourages the choice for entrepreneurship: An empirical test of a truism’, *Journal of Economic Behavior and Organization* **48**, 29–36.

De Wit, G. and F. Van Winden (1989), ‘An empirical analysis of self-employment in the Netherlands’, *Small Business Economics* **1**(4), 263–272.

Dolton, P. and G. Makepeace (1990), ‘Self-employment among graduates’, *Bulletin of Economic Research* **42**(1), 35–53.

Dunn, T. and D. Holtz-Eakin (2000), ‘Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links’, *Journal of Labor Economics* **2**(18), 282–305.

Evans, D. and L. Leighton (1989), ‘Some empirical aspects about entrepreneurship’, *American Economic Review* **79**(3), 519–535.

Evans, D. and L. Leighton (1990), ‘Small business formation by unemployed an employed workers’, *Small Business Economics* **2**(4), 319–330.

Evans, D. and B. Jovanovic (1989), 'An estimated model of entrepreneurial choice under liquidity constraints', *Journal of Political Economy* **97**(4), 808–827.

Evans, M. (1989), 'Immigrant entrepreneurship: Effects of ethnic market size and isolated labor pool', *American Sociological Review* **54**(6), 950–962.

Fairlie, R. and B. Meyer (1996), 'Ethnic and racial self-employment differences and possible explanations', *Journal of Human Resources* **31**(4), 757–793.

Fairlie, R. (1999), 'The absence of the African-American owned business: An analysis of the dynamics of self-employment', *Journal of Labor Economics* **17**(1), 80–108.

Fairlie, R. (2002), 'Drug dealing and legitimate self-employment', *Journal of Labor Economics* **20**(3), 538–567.

Flota, C. and M. Mora (2001), 'The earnings of self-employed Mexican Americans along the U.S.-Mexico border', *Annals of Regional Science* **35**, 483–499.

Fredland, J. and R. Little (1981), 'Self-employed workers: Returns to education and training', *Economics of Education Review* **1**(3), 315–337.

Fuchs, V. (1982), 'Self-employment and labor force participation of older males', *Journal of Human Resources* **17**(3), 339–357.

Gentry, W. and R. Hubbard (2000), 'Entrepreneurship and household saving', *NBER Working Paper 7894* .

Gill, A. (1988), 'Choice of employment status and the wages of employees and the self-employed: Some further evidence', *Journal of Applied Econometrics* **3**(3), 229–234.

Gimeno, J., T. Folta, A. Cooper and C. Woo (1997), 'Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms', *Administrative Science Quarterly* **42**, 750–783.

Guiso, L., P. Sapienza and L. Zingales (2002), 'Does local financial development matter', *Discussion Paper CEPR 3307* .

Hamilton, B. (2000), 'Does entrepreneurship pay? An empirical analysis of the returns to self-employment', *Journal of Political Economy* **108**(3), 604–631.

Hammarstedt, M. (2001), 'Immigrant self-employment in Sweden: Its variation and some possible determinants', *Entrepreneurship & Regional Development* **13**, 147–161.

Honjo, Y. (2000), 'Business failure of new software firms', *Applied Economic Letters* **7**, 575–579.

Hout, M. and H. Rosen (2000), 'Self-employment, family background, and race', *Journal of Human Resources* **35**(4), 670–692.

Hundley, G. (2001), 'Why women earn less than men in self-employment', *Journal of Labor Research* **22**(4), 817–829.

Johansson, E. (2000), 'Self-employment and liquidity constraints: Evidence from Finland', *Scandinavian Journal of Economics* **102**(1), 123–134.

Kidd, M. (1993), 'Immigrant wage differentials and the role of self-employment in Australia', *Australian Economics Papers* **32**(60), 92–115.

Kuhn, P. and H. Scheutze (2001), 'Self-employment dynamics and self-employment trends: A study of Canadian men and women, 1982-1998', *Canadian Journal of Economics* **34**(3), 760–784.

Laferrère, A. (2001), 'Self-employment and intergenerational transfers', *International Journal of Sociology* **31**(1), 3–26.

Lentz, B. and D. Laband (1990), 'Entrepreneurial success and occupational inheritance among proprietors', *Canadian Journal of Economics* **23**(3), 563–579.

Lin, Z., G. Picot and J. Compton (2000), 'The entry and exit dynamics of self-employment in Canada', *Small Business Economics* **15**, 105–125.

Lofstrom, M. (2000), 'Self-employment and earnings among high-skilled immigrants in the United States', *IZA Discussion Paper no. 175* .

Lofstrom, M. (2002), 'Labor market assimilation and the self-employment decision of immigrant entrepreneurs', *Journal of Population Economics* **15**(1), 83–114.

Lombard, K. (2001), 'Female self-employment and demand for flexible, nonstandard work schedules', *Economic Inquiry* **39**(2), 214–237.

Long, J. (1982), 'The income tax and self-employment', *National Tax Journal* **35**(1), 31–42.

MacPherson, D. (1988), 'Self-employment and married women', *Economic Letters* **28**(3), 281–284.

Maxim, P. (1992), 'Immigrants, visible minorities, and self-employment', *Demography* **29**(2), 181–198.

Meager, N. and P. Bates (2001), 'The self-employed and lifetime incomes', *International Journal of Sociology* **31**(1), 27–58.

Mehta, S. and Cooper, A. (2000), 'Optimism as a predictor of new firm performance', *FER 2000 XXXII* .

Moore, C. and R. Mueller (2002), 'The transition from paid to self-employment in Canada: The importance of push factors', *Applied Economics* **34**, 791–801.

Moore, R. (1982), 'Employer discrimination: Evidence from self-employed workers', *Review of Economics and Statistics* **65**(3), 496–501.

Rees, H. and A. Shah (1986), 'An empirical analysis of self-employment in the U.K.', *Journal of Applied Econometrics* **1**(1), 95–108.

Robinson, P. and E. Sexton (1994), 'The effect of education and experience on self-employment success', *Journal of Business Venturing* **9**(2), 141–156.

Sanders, J. and V. Nee (1996), 'Immigrant self-employment: The family as social capital and the value of human capital', *American Sociological Review* **61**(2), 231–249.

Schuetze, H. (2000), 'Taxes, economic conditions and recent trends in male self-employment: A Canada-US comparison', *Labour Economics* **7**, 507–544.

Simpson, W. and R. Sproule (1998), 'Econometric analysis of Canadian self-employment using SLID', *Income and Labour Dynamics Working Paper Series* .

Storey, D. and P. Wynarczyk (1996), 'The survival and non survival of micro firms in the UK', *Review of Industrial Organization* **11**, 211–229.

Taylor, M. (1996), 'Earnings, independence, or unemployment: Why become self-employed?', *Oxford Bulletin of Economics and Statistics* **58**(2), 253–266.

Taylor, M. (1999), 'Survival of the fittest? An analysis of self-employment duration in Britain', *Economic Journal* **109**(454), c140–c155.

Taylor, M. (2001), 'Self-employment and windfall gains in Britain: Evidence from panel data', *Economica* **68**, 539–565.

Tucker, I. (1985), 'Use of the decompensation technique to test the educational screening hypothesis', *Economics of Education Review* **4**(4), 321–326.

Tucker, I. (1987), 'The impact of consumer credentialism on employee and entrepreneur returns to higher education', *Economics of Education Review* **6**(1), 35–40.

Tucker, I. (1988), 'Entrepreneurs and public-sector employees: The role of achievement motivation and risk in occupational choice', *Journal of Economic Education* **19**(3), 259–268.

Tucker, I. (1990), 'Employer seniority discrimination', *Economics Letters* **32**, 85–89.

Uusitalo, R. (2001), 'Homo entreprenaurus?', *Applied Economics* **33**, 1631–1638.

Van Praag, M. and J. Cramer (2001), 'The roots of entrepreneurship and labour demand: Individual ability and low risk aversion', *Economica* **269**, 45–62.

Van Praag, M. and H. Van Ophem (1995), 'Determinants of willingness and opportunity to start as an entrepreneur', *Kyklos* **48**(4), 513–540.

Wagner, J. and R. Sternberg (2002), 'The role of the regional milieu for the decision to start a new firm: Empirical evidence for Germany', *Discussion Paper 494* .

Appendix 2.A: Construction of the database

This Appendix provides detailed information on the construction of the database used for the meta analysis in this study and describes some of the features of the data.

Database: Precision, coverage, rules, and sources

We aim for a fairly complete coverage of empirical studies that estimate a quantified relationship between entrepreneurship (entry and/or performance) and education. Several restrictions are imposed on studies for inclusion in the database. Only English language primary studies that are written for an academic audience are included. Studies pertaining to all industrialized countries are included, whereas studies pertaining to countries in a transition are excluded. Studies that pertain to limited parts of the population of entrepreneurs, as defined for instance by gender, age or race, are included in the sample and are marked as such.

The first avenue of search is the Internet. Therefore, only journal articles and book chapters that have been published since 1980 are considered, due to the lack of studies in several key virtual databases prior to that year. ECONLIT and the Web of Science are used as the primary sources for published journal articles.

To establish the relationship between the scientific weight of a journal (as measured by its impact factor) on the one hand and the schooling-entrepreneurship relationship on the other hand, it is required to include journals of all scientific weights. Moreover, to establish whether the schooling-entrepreneurship relationship varies between journals belonging to different academic fields – the field of entrepreneurship research is well known for its multidisciplinary character –, various academic fields are covered: economics, management, small business and entrepreneurship, human resources and labor relations, and other fields such as sociology and psychology.

Besides journal articles and book chapters, the database also includes unpublished papers. The motivation for this is twofold: It is the only way of including the most recent research output, and it enlarges the sample. Working papers are considered from 1997 to December 2002, the date at which we completed the construction of the database. In order to prevent double counting and to preserve independence of observations, working papers have been omitted that appeared more recently as publications (sometimes with a different title or authorship). The primary (virtual) search engines for working papers are the SSRN (Social Science Research Network), WOPEC (WOrking Papers in EConomics), Nep-ent (an e-mail alerting service on recent working papers in entrepreneurship research), and working papers series of well-known research institutes such as NBER, CEPR and IZA, as well as Frontiers of Entrepreneurship Research (a published selection of papers presented at the annual Babson-Kaufmann conference on entrepreneurship that is cited frequently).

The second avenue of search for both published and unpublished papers is a scan through the references of each sampled paper. Furthermore, the Web of Science enables searches of citations. Thus, after an article has been published, it is straightforward to find all other articles (in the journals covered) that refer to the studies that were already captured in the sample. Together, these two citation search methods yield a substantial number of additional studies that must be inserted into the database.

Table 2.A-1 shows the resulting selection of studies. It consists of 94 studies that included at least one valid observation on either the quantified relationship between schooling and entry/selection (a transition to entrepreneurship) or between schooling and performance (earnings, duration, etc.). As shown in the first summary row of the table, altogether, the 94 studies yielded 299 observations. Among these, 144 (48 percent) examine performance, 69 (23 percent) investigate entry into entrepreneurship, and 86 (29 percent) specify the dependent variable as "being self-employed". The latter is a stock (rather than flow) variable that is a hybrid of entry (everyone who is currently self-employed has entered this occupational status) and performance (it generates an overrepresentation of survivors). Therefore, "stock studies" are a separate category (Van Praag, 2003).

Table 2.A-1: Sources of the sample

A=Entry B=Stock C=Performance	Number of observations							
	Reduced Form			*Structural*				
Study	*A*	*B*	*C*	*A*	*B*	*C*	*Total*	*Country*
Summary								
All Studies	62	85	117	7	1	27	299	All countries
Individual Studies								
1. Alba-Ramirez, 1994	4	0	1	0	0	0	5	Spain, USA
2. Bates, 1985	0	0	1	0	0	0	1	USA
3. Bates, 1990	0	0	1	0	0	0	1	USA
4. Bates, 1995	6	0	0	0	0	0	6	USA
5. Bates, 1999	0	0	1	0	0	0	1	USA
6. Bates; Dunham, 1993	1	0	0	0	0	0	1	USA
7. Bernhardt, 1994	0	1	1	0	0	0	2	Canada
8. Blanchflower, 2000	0	16	0	0	0	0	16	*
9. Blanchflower; Meyer, 1994	2	0	2	0	0	0	4	USA,Australia
10.Blanchflower; Oswald; Stutzer,01	0	3	0	0	0	0	3	**
11. Blumberg; Pfann, 1999	1	0	0	0	0	0	1	Netherlands
12. Boden, 1996	2	0	0	0	0	0	2	USA

Continued on next page ...

Table 2.A-1 continued ...

A=Entry B=Stock C=Performance	Number of observations							
	Reduced Form			Structural				
Study	*A*	*B*	*C*	*A*	*B*	*C*	*Total*	*Country*
13. Boden, 1999	0	2	0	0	0	0	2	USA
14. Boden, 1999	1	0	0	0	0	0	1	USA
15. Boden; Nucci, 2000	0	0	4	0	0	0	4	USA
16. Borjas, 1986	0	6	0	0	0	0	6	USA
17. Borjas; Bronars, 1989	0	4	4	0	0	0	8	USA
18. Bosma et al, 2002	0	0	3	0	0	0	3	Netherlands
19. Boyd, 1991	0	2	0	0	0	0	2	USA
20. Brown; Sessions, 1998	0	1	0	0	0	1	2	UK
21. Brown; Sessions, 1999	0	1	0	0	0	1	2	Italy
22. Bruce, 1999	0	0	13	0	0	0	13	USA
23. Bruce, 2000	1	0	0	0	0	0	1	USA
24. Bruce, 2002	2	0	12	1	0	0	15	USA
25. Bruce; Holtz-Eakin; Quinn, 2000	2	0	2	0	0	0	4	USA
26. Bruderl; Preisendorfer, 1992	0	0	1	0	0	0	1	Germany
27. Bruderl; Preisendorfer, 1998	0	0	3	0	0	0	3	Germany
28. Caputo; Dolinsky, 2001	0	1	0	0	0	0	1	USA
29. Carrasco, 1999	1	0	0	0	0	0	1	Spain
30. Carroll; Mosakowski, 1987	2	0	0	0	0	0	2	Germany
31. Clark; Drinkwater; Leslie, 1998	0	2	0	0	0	2	4	UK
32. Clark; Drinkwater, 1998	3	0	0	0	0	3	6	USA
33. Cooper; Folta; Gimeno; Woo, 1992	0	0	2	0	0	0	2	USA
34. Cooper; Gimeno; Woo, 1994	0	0	2	0	0	0	2	USA
35. Cooper; Woo; Dunkelberg, 1988	0	0	1	0	0	0	1	USA
36. Cowling, 2000	0	13	0	0	0	0	13	***
37. Cramer; Hartog; Van Praag, 2002	1	0	0	0	0	0	1	Netherlands
38. De Wit; Van Winden, 1989	1	0	0	1	0	1	3	Netherlands
39. Dolton; Makepeace, 1990	1	0	0	1	0	1	3	UK
40. Dunn; Hotz-Eakin, 2000	1	0	0	0	0	0	1	USA
41. Evans, 1989	0	1	0	0	0	0	1	Australia
42. Evans: Jovanovic, 1989	1	0	2	0	0	0	3	USA
43. Evans; Leighton, 1989	4	1	1	0	0	0	6	USA
44. Evans; Leighton, 1990	4	1	1	0	0	0	6	USA
45. Fairlie, 1999	2	0	2	0	0	0	4	USA
46. Fairlie, 2002	1	1	0	0	0	0	2	USA
47. Fairlie; Meyer, 1996	0	2	1	0	0	1	4	USA
48. Flota; Mora, 2001	0	0	2	0	0	0	2	USA
49. Fredland; Little, 1981	0	0	1	0	0	0	1	USA
50. Fuchs, 1982	1	0	0	0	0	0	1	USA
51. Gentry; Hubbard, 2000	1	0	2	0	0	0	3	USA
52. Gill, 1988	0	0	0	0	1	1	2	USA
53. Gimeno; Folta; Cooper; Woo, 1997	0	0	2	0	0	1	3	USA
54. Guiso; Sapienza; Zingales, 2002	0	1	0	0	0	0	1	Italy
55. Hamilton, 2000	0	0	3	0	0	0	3	USA
56. Hammarstedt, 2001	0	1	0	0	0	0	1	Sweden
57. Honjo, 2000	0	0	1	0	0	0	1	Japan

Continued on next page ...

Table 2.A-1 continued ...

A=Entry B=Stock C=Performance	Number of observations							
	Reduced Form			Structural				
Study	*A*	*B*	*C*	*A*	*B*	*C*	*Total*	*Country*
58. Hout; Rosen, 1999	0	1	0	0	0	0	1	USA
59. Hundley, 2001	0	0	4	0	0	0	4	USA
60. Johansson, 2000	1	0	0	0	0	0	1	Finland
61. Kangasharju; Pekkala, 2002	0	2	4	0	0	4	10	Finland
62. Kidd, 1993	0	1	1	0	0	0	2	Australia
63. Kuhn; Scheutze, 2001	4	0	2	0	0	0	6	Canada
64. Laferrere, 2001	0	1	0	0	0	0	1	France
65. Lentz; Laband, 1990	0	0	1	0	0	0	1	USA
66. Lin; Picot; Compton, 2000	1	0	1	0	0	0	2	Canada
67. Lofstrom, 2000	0	1	1	0	0	1	3	USA
68. Lofstrom, 2002	0	1	1	0	0	1	3	USA
69. Lombard, 2001	0	1	1	0	0	0	2	USA
70. Long, 1982	0	2	2	0	0	0	4	USA
71. Macpherson, 1988	0	1	0	0	0	1	2	USA
72. Maxim, 1992	0	1	0	0	0	2	3	Canada
73. Meager; Bates, 2001	1	0	1	0	0	0	2	UK
74. Mehta; Cooper, 2000	0	0	2	0	0	0	2	USA
75. Moore, 1983	0	0	4	0	0	0	4	USA
76. Moore; Mueller, 2002	3	0	0	0	0	0	3	Canada
77. Rees; Shah, 1986	1	0	0	1	0	1	3	UK
78. Robinson; Sexton, 1994	0	3	3	0	0	0	6	USA
79. Sanders; Nee 1996	0	2	0	0	0	0	2	USA
80. Scheutze, 2000	0	1	0	0	0	0	1	USA, Canada
81. Simpson; Sproule, 1998	0	2	0	2	0	2	6	Canada
82. Storey; Wynarczyk, 1996	0	0	3	0	0	0	3	UK
83. Taylor, 1996	1	0	0	1	0	1	3	UK
84. Taylor, 1999	0	0	6	0	0	0	6	UK
85. Taylor, 2001	1	0	1	0	0	1	3	UK
86. Tucker, 1985	0	0	1	0	0	0	1	USA
87. Tucker, 1987	0	0	2	0	0	0	2	USA
88. Tucker, 1988	0	1	0	0	0	0	1	USA
89. Tucker, 1990	0	2	0	0	0	0	2	USA
90. Uusitalo, 2001	1	2	0	0	0	0	3	Finland
91. Van Praag, 2002	0	0	3	0	0	0	3	Netherlands
92. Van Praag; Cramer, 2001	0	0	1	0	0	1	2	Netherlands
93. Van Praag; van Ophem, 1995	1	0	0	0	0	0	1	USA
94. Wagner; Sternberg, 2002	1	0	0	0	0	0	1	Germany

*USA, France, Belgium, Netherlands, Germany, Italy, Luxembourg, Denmark, Ireland, UK, Greece, Spain, Portugal, Norway, Austria, Sweden, Finland, Canada; **Bulgaria, Canada, Czech Republic, Denmark, East Germany, France, United Kingdom, Hungary, Israel-Arabs, Israel-Jews, Italy, Japan, New Zealand, Norway, Poland, Portugal, Russia, Slovenia, Sweden, Switzerland, United States of America; *** Portugal, United Kingdom, Austria, Sweden, Finland, Belgium, Denmark, Germany, Greece, Spain, France, Ireland, Italy

Furthermore, for reasons indicated in Section 2.3, structural (entry, performance and stock) studies are distinguished from reduced form studies into the same relationship. Those studies labeled as "structural" attempt to incorporate at least some kind of a deliberate occupational choice of labor force participants, and it is worthwhile to compile such studies under a separate heading in order to examine the direction of the selectivity bias. Eleven percent of the observations is structural: 19% (27/(27+117)) of the performance observations, 10% (7/(7+62)) of the entry observations and a negligible proportion of stock observations.

Characteristics of the data

Underneath the first summary row, Table 2.A-1 shows the number of valid observations per study per research relationship of interest (either entry or performance). The table also shows to which country or country group each study pertains. The geographical distribution of studies into entrepreneurship entry and performance is summarized in Table 2.A-2. The USA dominates by far, contributing 56 percent of all observations. The U.K. is the country of study for 10 percent of the observations, continental Europe for 23 percent, and Australia and Canada for nine percent. Some studies pertain to various countries.

Table 2.A-2: Countries studied

Observation	USA	UK	Australia Canada	Continental Europe	All / other
Entry	36	5	9	12	0
Stock	37	5	7	32	4
Performance	83	11	6	16	1
Structural	11	10	6	8	0
Total	167	31	28	68	5

Most studies (92 percent) include the entire working population (age group 20-65) in their samples (not tabulated). Six percent of the studies are limited to young labor force participants/entrants and two percent to older workers. As for race, 80 percent of the studies use no clause and pertain to all races (with or without control variables to allow for different intercepts). Twelve percent of the observations are confined to white labor force participants. Eight percent of observations reflect the determinants of entrepreneurship entry and success of ethnic and minority groups, as well as of immigrants from certain geographical areas, as ethnicity is believed to affect these determinants at large (see section 2.5). The same holds true for gender. Ten percent of the studies are limited to females, 40 percent pertain to males, and half of the studies use both while allowing for differences through an intercept shifter.

Figure 2.A-1 shows the distribution of sample sizes across studies by type of study. The median sample size for entry studies is 9540, for stock studies 13900, for performance studies 610, and for structural studies 4760 for entry and 1545 for performance. Obviously, the median sample sizes in performance studies are small relative to the same type of studies of employees. The distribution differs somewhat over performance measures, though. The median sample size for earnings studies is 350 only, for exit studies 2615, for duration 455, and for survival (defined as the opposite of exit) 1055.

Table 2.A-3 characterizes the origin of the studies in the sample. 93.6 percent of the studies found have been published by December 2002. Panel A distinguishes several types of journals in which these studies have been published: general economics journals (44%), labor economics and education journals (23%), small business and entrepreneurship journals (15%), management journals (2%), working papers (6%), and other journals (10%). Structural studies are over-represented in the category of economics journals, whereas stock studies are over-represented in labor economics journals.

Besides type, journals are also categorized in terms of their (2001) SSCI impact factor that we use as a measure of scientific weight or quality. Panel B shows that the majority of studies has been published in journals of rather low impact. Somewhat less than twenty percent of studies have been published in journals with an impact factor higher than 1.0. Contrary to expectations, structural studies are underrepresented in journals of higher impact and somewhat overrepresented in journals without an impact factor.

Panel C finally shows the distribution over publication years of each type of study: Half of the studies have been published in 1998 or later. The table clearly shows a revived interest in the topic. No clear distinct trends pop up for particular types of studies, except that structural studies were relatively popular in the mid 1990s.

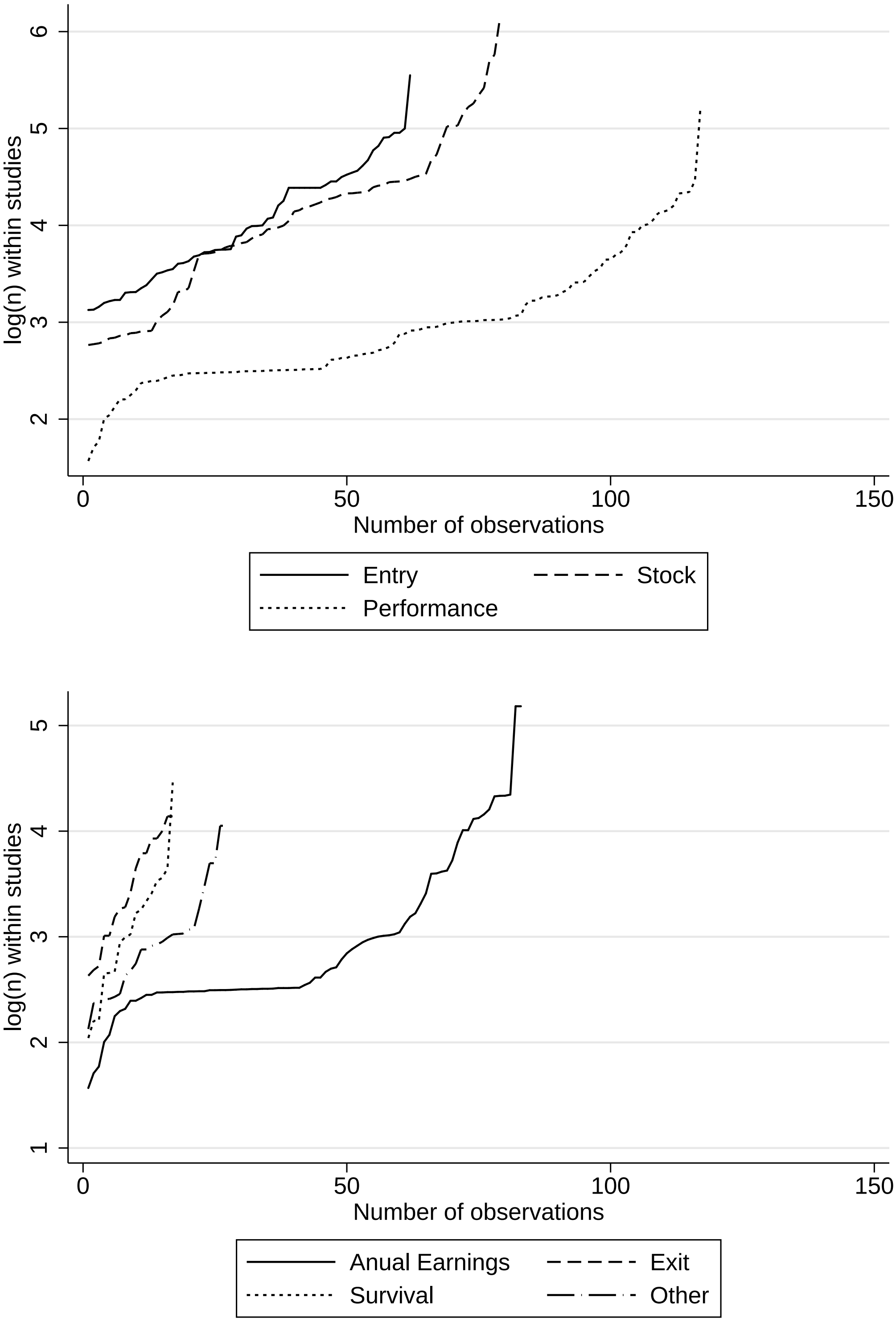


Figure 2.A-1: Distribution of sample sizes over observations

Table 2.A-3: Characteristics of studies

Panel A: Journal categories of (published) studies

#	Ec*		Lab*		SB/E*		Mng*		WP*		Oth*	
Entry	29	46.7%	6	9.7%	15	24.2%	2	3.2%	5	8.1%	5	8.1%
Stock	35	41.1%	31	36.5%	6	7.1%	0	0.0%	2	2.4%	11	12.9%
Performance	49	42.0%	24	20.5%	21	17.9%	5	4.3%	10	8.5%	8	6.8%
Structural	19	54.3%	7	20.0%	2	5.7%	0	0.0%	1	2.9%	6	17.1%
Total	132	44.3%	68	22.7%	44	14.7%	7	2.3%	18	6.0%	30	10.0%

Panel B: Impact factors of published studies

#	>1.5		1.0-1.5		0.5-1.0		0-0.5		No	
Entry	7	12.3%	1	1.7%	19	33.3%	23	40.4%	7	12.3%
Stock	8	9.8%	10	12.2%	10	12.2%	32	39.0%	22	26.8%
Performance	13	12.1%	13	12.1%	33	30.9%	32	29.9%	16	15.0%
Structural	1	3.3%	1	3.3%	9	30.1%	12	40.0%	7	23.3%
Total	29	10.4%	25	8.9%	74	26.4%	100	35.7%	52	18.6%

Panel C: Publication year of published studies

#	1980-1989		1990-1993		1994-1997		1998-1999		2000-2002	
Entry	15	26.3%	7	12.3%	13	22.8%	10	17.5%	12	21.1%
Stock	19	23.2%	8	9.8%	9	11.0%	33	40.1%	13	15.9%
Performance	22	20.6%	13	12.1%	9	8.4%	30	28.0%	33	30.9%
Structural	8	23.5%	3	8.8%	14	41.2%	2	5.9%	7	20.6%
Total	64	22.9%	31	11.1%	45	16.1%	75	26.7%	65	23.2%

*Ec=Economics; Lab=Labor economics; SB&E=Small Business and Entrepreneurship; Mng=Management; Wp=Working papers; Oth=Other journals.

Methodologies and types of data used

There is a substantial variation in terms of the estimation methods used (OLS, duration, or probit), the type of data used (panel or cross-section), and the processes that are treated as endogenous in the structural models. Since probit, logit and linear probability models can be easily translated into each other, these are treated as a single method (referred to as "probit"), and the relevant coefficients are translated accordingly.

Almost half of the entry studies use panel data and "probit." Models in the other half are mainly estimated by the same methodology on cross-sectional data (37 percent) or by multinomial logit models using panel data (10 percent). The percentage of entry studies that are based on cross-sectional data is rather high: These data must include some information on the career histories of the individuals included in the sample such that the researchers can assess who has ever started as an entrepreneur and when that was the case, so that time-varying variables can be recoded to that moment. The multinomial logit models examine the flows between various labor market states such as self-employment, wage employment, unemployment, and in some studies also, non-employment. These models are used to study, for example, the decision by unemployed workers to become self-employed or to find a new wage job. The estimation methods used for stock studies are not very surprising: 95 percent use probit models or variants, where the division between cross-sectional and panel data usage is two thirds to one third. Estimation methods for performance models vary across the performance measures in an expected way.

Structural models usually combine entry and performance (earnings) equations; commonly, such studies use a structural probit and an OLS-equation. The latter includes an inverse Mill's ratio, derived from the reduced form probit equation of entry, such that consistent estimates of the wage-employment and self-employment income equations result. The expected income difference for these labor market states is then finally included in the structural probit models. Most studies find significant coefficients for both the inverse Mill's ratios and the expected income difference, pointing at (positive) selection effects and incentives generated by expected income differences for occupational choices.

As mentioned earlier, only 19 percent of the performance observations and even fewer of the entry and stock observations estimate structural models. Structural models were relatively popular in the 1990s; it seems surprising that, with the threats of endogeneity bias in standard OLS analysis, the structural model has not become a standard approach. This must be due to difficulties in finding identifying instruments for entry (see Chapter 3).

Appendix 2.B: The impact of control variables

Section 2.6 discussed the effects of the inclusion of specific control variables in the sampled observations on the observed schooling-entrepreneurship relationships. It turned out that there was no such effect. This appendix discusses the *direct* effects of these control variables on entrepreneurship entry, stock and performance. Table 2.B-1 shows the most commonly used control variables (column 2), the number of observations that uses them (column 3) and the signs of the resulting estimates (columns 4 to 7). Controls are tabulated whenever the estimated effect of the control is reported in more than 5 studies. The exception to this self-imposed rule is the control variable of ability, which is important but rarely used (in contrast with the comparable literature on wage employees where approximately 20 percent of studies use ability controls (Ashenfelter et al., 1999)).

As in the regular returns-to-schooling literature, the Mincerian variables age and (general or self-employment) experience appear very frequently, as do marital status, gender, and race.[27] The distinction between general and self-employment experience, which is somewhat akin to specific human capital, is made in some studies but could clearly receive more attention. Health status appears in some studies. The use of parental background variables is usually limited to father's self-employment, if it is included at all.

The last four columns of the table summarize the impact of these control variables on entrepreneurial outcomes. The general picture is described as follows. Males are more likely to become self-employed and to succeed as such. Age is commonly modeled in a non-linear fashion, as "age squared" has frequently been included. Therefore, for those studies we have also calculated the marginal effect of age at age 30. The relationship of age with each outcome variable tends to have an inverted U-shape, but the marginal effect at age 30 is ambiguous with respect to performance and stock and positive only with respect to entry. People with a minority ethnic background are somewhat less inclined to start as entrepreneurs, and they perform slightly poorer than those not labeled as an ethnic minority. People with a handicap or who are in bad health do not have a significantly different probability to become self-employed, though success is harder to achieve.

[27] Table 2.B-1 in fact understates the concerns in the literature with gender and race, since a number of the studies analyze the entrepreneurial outcomes among specific gender and race subsamples.

Table 2.B-1: Effect of control variables on entrepreneurial outcomes

entrepreneurial outcome	variable	# studies	% negative coefficients ($t < -1.96$)	% insignif. negative coefficients	% insignif. positive coefficients coefficients	% positive coefficients ($t > 1.96$)
Entry	Male	28	7.1%	0.0%	14.3%	78.6%
Stock	Male	42	2.4%	7.1%	16.7%	73.8%
Performance	Male	41	4.9%	17.1%	14.6%	63.4%
Entry	Age	44	4.5%	13.6%	11.4%	70.5%
Stock	Age	66	3.0%	9.1%	9.1%	78.8%
Performance	Age	76	7.9%	46.1%	15.8%	30.3%
Entry	Age squared	36	63.9%	13.9%	13.9%	8.3%
Stock	Age squared	41	58.5%	17.1%	17.1%	7.3%
Performance	Age squared	64	25.0%	57.8%	12.5%	4.7%
Entry	Age effect at 30	36	27.8%	12.1%	10.1%	50.0%
Stock	Age effect at 30	41	36.6%	13.4%	13.4%	36.6%
Performance	Age effect at 30	63	17.5%	50.9%	17.3%	14.3%
Entry	Married	51	11.8%	19.6%	27.5%	41.2%
Stock	Married	49	8.2%	22.4%	20.4%	49.0%
Performance	Married	53	7.5%	32.1%	32.1%	28.3%
Entry	Ethnic minority	19	36.8%	26.3%	26.3%	10.5%
Stock	Ethnic minority	12	83.3%	8.3%	8.3%	0.0%
Performance	Ethnic minority	36	22.2%	44.4%	13.9%	19.4%
Entry	Disabled	12	16.7%	50.0%	25.0%	8.3%
Stock	Disabled	22	4.5%	31.8%	31.8%	31.8%
Performance	Disabled	25	52.0%	16.0%	28.0%	4.0%
Entry	Self-employed father	10	10.0%	0.0%	10.0%	80.0%
Stock	Self-employed father	4				
Performance	Self-employed father	17	11.8%	41.2%	35.3%	11.8%
Entry	Self-Employment experience	5				
Stock	Self-Employment experience	0				
Performance	Self-Employment experience	25	16.0%	36.0%	36.0%	12.0%
Entry	Employment experience	18	11.1%	5.6%	38.9%	44.4%
Stock	Employment experience	14	0.0%	0.0%	7.1%	92.9%
Performance	Employment experience	48	8.3%	10.4%	35.4%	45.8%
Entry	Ability	4	0.0%	50.0%	50.0%	0.0%
Stock	Ability	0				
Performance	Ability	4	0.0%	0.0%	50.0%	50.0%

In line with common wisdom, people with a self-employed father are more likely to become an entrepreneur. However, they are not more likely to be successful, as is sometimes thought and seldom found: This informal type of experience is hardly productive. Parental self-employment may therefore be less an indicator of informal experience and more a measure of a certain preference for risk and working conditions such as, for instance, freedom of operation, decision making, working hours, and type of obligations. Prior self-employment experience appears to have no impact on performance. In the five studies on entry that used this variable, it always increases the probability to enter; in this way, it seems to measure the same kinds of underlying factors as parental self-employment.[28] General work experience, on the contrary, increases both the probability of entry and success.

[28]Not tabulated because it does not meet our self-imposed "more than five studies" rule.

Chapter 3

Why are the returns to education higher for entrepreneurs than for employees?[1]

3.1 Introduction

In Chapter 1 we discussed the importance of entrepreneurship for society. Entrepreneurs are considered the engine of the economy, responsible for sustained levels of competition, the creation of jobs, and innovation. These benefits, which accrue to society at large, justify public expenditure to develop and stimulate entrepreneurship. In this thesis the focus lies on the beneficial effects of formal education on entrepreneur performance. The previous chapter already gave us an indication of the benefits of education for entrepreneur success. However that chapter also noted that the "effect" of schooling that is typically estimated based on (conditional) correlations is not causal: Ability and other factors might increase performance and also lead to more schooling, thus leading to a spurious positive effect of schooling on performance. If we want to make policy recommendations we have to figure out the magnitude of the returns to education. The aim of this chapter is to measure the magnitude of the causal effect of formal education on entrepreneur performance relative to the magnitude of this effect for employees.

The contribution of this chapter to the literature is discussed in the next section. To this end, we first discuss the empirical literature on the relationship between education and entrepreneurship outcomes, and compare this to the literature pertaining to the education-income relationship for employees. Section 3 describes the sample (USA NLSY 1979) and

[1] This chapter is based on Van der Sluis, Van Praag and Van Witteloostuijn (2005). For a non-technical and qualitative summary see Van der Sluis and Van Praag (2007).

the identification strategy applied. We use a random-effects IV-approach, taking account of the endogeneity of both schooling and self-selection into entrepreneurial positions in income equations. In Section 4, we present the estimation results from the education and selection equations, as well as from the income equation. The latter reveals significantly higher returns to education for entrepreneurs than for employees: The coefficient of the interaction of 'education level' and 'entrepreneur' (both instrumented) is significantly positive.

Section 5 is devoted to finding an explanation for the result that entrepreneurs obtain higher returns from their education than employees. It turns out that the rather large and significant difference cannot be attributed to, for instance, risk premiums or returns to capital included in the entrepreneurial income variable. However, we find (indirect) support for the following explanation of the larger return to education for entrepreneurs: Entrepreneurs can better control and manage the optimal employment of their investment in education than their employed counterparts in the labor market. This leads to conclusions and policy recommendations that follow from the estimation results under quite broad assumptions. These are discussed in the concluding Section 6.

3.2 Empirical literature

Almost a hundred empirical studies have measured the relationship between schooling and entrepreneurship outcomes. In Chapter 2 we provided an overview and performed a meta-analysis to assess whether there are any consistent findings with respect to the relationship between education on entry and performance in(to) entrepreneurship. Four outcomes are relevant to the current chapter.

First, the relationship between education and selection into an entrepreneurial position is mostly insignificant – i.e., in 75 percent of the cases. However, the relationship between schooling and performance is unambiguously positive and significant in 67 percent of the observed studies.

Second, the meta-analysis gives insight into the level of the returns to education for entrepreneurs in terms of income. The return to a marginal year of schooling is 6.1 percent, on average. This insight, though, is based on the measurement of conditional correlations on the basis of a rather small sub-sample of USA studies using similar measures of education and earnings.

Third, the meta-analysis identifies approximately twenty studies that have actually measured the relationship between education and earnings for both entrepreneurs and employees in a comparable fashion. The measured returns to education turn out to be similar for entrepreneurs and employees. More specifically, in Europe the returns to education seem to be slightly lower for entrepreneurs than for employees, and in the USA the opposite result is found.

The fourth conclusion from the meta-analysis is that previous studies have not yet employed estimation strategies that account for the endogenous nature of schooling in performance equations and unobserved individual characteristics that may drive the result, possibly leading to inconsistent estimates. Many of these studies measure the relationship between education and entrepreneurship outcomes as a by-product while focusing on different issues. Therefore, they have not pursued to apply empirical strategies that measure the effect of education consistently, such as the instrumental variables approach or twin studies.

Amongst labor economists, who study the returns to education for employees, taking account of the endogenous nature of schooling and of unobserved heterogeneity has become more and more common practice (Ashenfelter et al., 1999). The first strategy used to cope with unobserved ability is trying to make the unobservable observable. Various proxies of intelligence and test scores have been included in income equations. The effects of adding such controls on the estimated returns to education have been ambiguous (see Ashenfelter et al., 1999, Table 3).[2] Inclusion of ability proxies in the income function does not completely shield the estimated returns against ability bias due to an imperfect correlation between such proxies and ability. Nor does it control for endogeneity since ability is not necessarily perfectly correlated with the optimization behavior of individuals. Additional approaches are thus used to estimate the returns to education for employees.

The second identification strategy uses a sample of monozygotic twins (for instance, Ashenfelter and Krueger, 1994; Bonjour, Cherkas, Haskel, Hawkes and Spector, 2003). Identification comes from those twins who differ in their schooling and earnings outcomes, assuming that all unobserved factors are approximately equal. Drawbacks of this identification technique are that most twin studies rely on small samples and voluntary participation, are very sensitive to measurement error, and do not really cope with the endogeneity of the schooling decision (Bound and Solon, 1999). The usual finding is that treating education as an exogenous variable leads to downwards biased estimates of the returns to education (Ashenfelter et al., 1999).

The third identification strategy used is the instrumental variable (IV) approach. Instruments are identified that explain a substantial proportion of the variance of the endogenous variable, education in this case, but are unrelated to the dependent variable – i.e., income. Thus, the instrumented endogenous variable is not related to the error term anymore. This

[2] Theory predicts that omitting ability in the wage equation causes OLS-estimates to be upward biased (Griliches, 1977; Harmon and Walker, 1995; Ashenfelter et al., 1999).

method strongly hinges on the quality and validity of the identifying instruments used. Like using twins, the IV-strategy leads to higher estimates of the returns to education of employees than when treating education as an exogenous variable. This is not only the case when parental background variables are used as identifying instruments (Blackburn and Neumark, 1993), but also when changes in compulsory schooling laws are introduced as such (Angrist and Krueger, 1991; Oreopoulos, 2003).

Since the meta-analysis and prior to this study, two studies have been performed that use the IV-methodology to measure the returns to education for entrepreneurs, i.e. Van der Sluis and Van Praag (2004) and Parker and Van Praag (2006). In the current chapter, we re-evaluate the returns to education for entrepreneurs relative to employees, without some of the drawbacks that characterized the earlier attempts. Like Van der Sluis and Van Praag (2004), and unlike Parker and Van Praag (2006), we measure the returns to education for entrepreneurs as well as employees. Unlike Van der Sluis and Van Praag (2004), the current chapter measures the returns to education for both groups within one framework (income equation) such that the (significance of the) difference in the returns to education across the two groups can be compared (by including interactions: see below). Like Van der Sluis and Van Praag (2004), but unlike Parker and Van Praag (2006), the data enable estimating income equations without survival bias for entrepreneurs by using the panel structure of the data. Moreover, unlike Van der Sluis and Van Praag (2004), the returns to education are estimated while taking account of self-selection into entrepreneurial positions based on unobserved characteristics. Furthermore, our set of identifying instruments does not include parental education levels. In this respect, too, this chapter differs from both previous studies. Last but not least, in the current chapter, we are able to find an explanation for the robust finding that the returns to education are higher for entrepreneurs than for employees. These differences indicate the contribution of the current chapter.

We shall benchmark our results against previous applications of the estimation strategies described, both for entrepreneurs and employees.[3] This may lead to some insight in the quality of our identification strategy and choice of instruments. The data and empirical methodology used are presented in the following section.

[3]See Ashenfelter et al. (1999) for an overview of the returns to education for employees using various estimation methods.

3.3 Data

3.3.1 Data description

We estimate the effect of education on incomes for both entrepreneurs and employees on a sample drawn from the National Longitudinal Survey of Youth (NLSY) in the USA. The nationally representative part of the NLSY consists of 6,111 individuals aged between 14 and 22 years in 1979.[4] They have been interviewed annually up to 1994, and since then on a bi-annual basis. Our analysis is based on 19 waves, where the first interviews were held in 1979, and the last in the year 2000. Within each observed year, the sample includes all persons who are entrepreneurs or employees (defined below), while excluding students and people who are unemployed or otherwise not working. The resulting sample size per year includes, on average, 2,646 entrepreneurs/employees. On average, each individual is included in the sample in 12.8 waves. Before turning to the descriptive statistics, we first define the three endogenous variables empirically – i.e., entrepreneur/employee, education and income – and mention three appealing features of the dataset.

An entrepreneur is defined as a person whose main occupation in the labor market is on a self-employed basis or who is the owner-director of an incorporated business. Farmers are excluded from the sample.[5] Furthermore, we exclude "hobby" entrepreneurs from the sample by using a lower boundary of 300 hours per year worked as an entrepreneur. An employee is defined as a person whose main occupation is a salaried job. The education level of both groups is measured in years of completed schooling, with a maximum of 20.

Hourly income is constructed as the average earnings (for entrepreneurs, the average income withdrawn from their firm) over a year divided by the number of hours worked in that year (see Fairlie (2005*b*) for an evaluation of the income variable in the NLSY for entrepreneurship research).

An important feature of the sample is that it includes both entrepreneurs and employees, and it records individuals' switches between these states over time. All entrepreneurship spells, also short ones, are recorded.[6] Therefore, the sub-sample of entrepreneurs does not suffer from survival bias – i.e., the returns to education will not pertain to surviving entrepreneurs only. Moreover, incomes and all other relevant variables are measured in a comparable way for both groups such that the returns to education for employees and en-

[4]The original NLSY sample consists of 12,686 individuals. From this sample we excluded the supplementary military sample and the supplementary minority sample.

[5]Their economics are very different from other occupations. From 1979 to 2000, we left out 299 farmers. Most studies drop farmers or study them separately.

[6]Spells shorter than half a year are not observed.

trepreneurs can be estimated in one equation.[7]

Another appealing feature of the NLSY is that it contains the Armed Services Vocational Aptitude Battery (ASVAB), an IQ-like test score that is used as a proxy for general ability. ASVAB (administered in 1979-1980)[8] includes ten components: (1) General science, (2) Arithmetic reasoning, (3) Word knowledge, (4) Paragraph comprehension, (5) Numerical operations, (6) Coding speed, (7) Auto and shop information, (8) Mathematics knowledge, (9) Mechanical comprehension, and (10) Electronic information. Respondents were of different ages and had different levels of education when the test was administered. We remove the age and education effects from the ASVAB by regressing each normalized test score on a set of age and education dummies (Blackburn and Neumark, 1993; Fairlie, 2005*a*). The individuals' residuals are used as the new test scores. The general ability proxy that shall be used is extracted from these new test scores by means of factor analysis. Hence, the value of each individual's general ability is a weighted average of the ten scores on the ASVAB components, where the factor loadings resulting from the factor analysis, are used as weight.

Another quality of the NLSY is the presence of detailed family and individual background variables. Some of the family background characteristics qualify as identifying instruments as they are possibly good predictors of the educational level of the respondent or the respondent's choice for entrepreneurship, while otherwise independent of their future earnings. Although administered in 1979-1980, these variables are most of the time recollections of household characteristics at the age of 14 (e.g., the presence of a library card in the household). We will discuss these variables in detail in the next subsection.

Table 3.1 shows the means and standard deviations of all the variables that are directly or indirectly used in the analyses. The values in Table 3.1 represent the averages of the specific variable over the period 1979-2000, where each year-sample includes only entrepreneurs and employees. We highlight the statistics of the three (endogenous) variables that are of main interest. First, the average percentage of entrepreneurs in the labor force is six. We observe at least one spell of entrepreneurship in the period 1979-2000 for twenty four percent of the sample. Moreover, those individuals who have been entrepreneurs in the observed period, have been so for 3.3 years, on average.

---

[7]Recent evidence suggests that entrepreneurs underreport their income more than employees do (Lyssiotou, Pashardes and Stengos, 2004). This might have implications for our estimation results. Moreover, the incomes of entrepreneurs might include, besides labor income, the income from capital invested in the business (Fairlie, 2005*b*). This might confound the comparability of entrepreneurs' and employees' incomes. We shall address and test the presence of these potential problems, amongst others, in Section 5.

[8]The early adminstration of the ASVAB allows us to treat this variable as exogenous in all equations.

Table 3.1: Summary of descriptive statistics

Variable	Total sample Mean	SD	N			
% Entrepreneurs	0.06	0.24	50268			
% Ever entrepreneur	0.24	0.43	50268			
	If person has ever been Entrepreneur					
	Mean	**SD**	**N**			
Duration entrepreneur spell(in years)	3.28	3.05	12063			
	Employees			**Entrepreneur**		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Hourly earnings (in $)	10.47	15.56	47195	14.52	29.49	3073
Education (in years)	13.06	2.34	47195	13.11	2.43	3073
Control Variables						
General ability (ASVAB)	5.12	1.11	47195	5.27	1.07	3073
Male (dummy)	0.51	0.5	47195	0.64	0.48	3073
Married (dummy)	0.51	0.5	47195	0.64	0.48	3073
Not healthy (dummy)	0.02	0.15	47195	0.03	0.18	3073
Live outside city (dummy)	0.23	0.42	47195	0.23	0.42	3073
Live in South (of USA, dummy)	0.31	0.46	47195	0.26	0.44	3073
Hispanic (dummy)	0.04	0.2	47195	0.03	0.16	3073
Black (dummy)	0.09	0.29	47195	0.04	0.19	3073
Age (in years)	28.21	5.7	47195	30.19	5.39	3073
(Internality of) locus-of-control	6.26	1.98	47195	6.54	1.95	3073
Education mother (in years)	11.7	2.43	47195	12.1	2.27	3073
Education father (in years)	11.82	3.35	47195	12.23	3.28	3073
Total business value (in $)*				191411	484410	1004
Total asset value (in $)**	88578	231261	25831	134864	253607	2032
Instruments						
Magazines (in hh at age 14, dummy)	0.69	0.46	47195	0.77	0.42	3073
Library card (in hh at age 14, dummy)	0.76	0.43	47195	0.78	0.41	3073
Stepparent (in hh at age 14, dummy)	0.06	0.23	47195	0.05	0.21	3073
# siblings	3.2	2.1	47195	3.11	1.96	3073
(Adhering to) strict religion (dummy)***	0.19	0.39	46232	0.16	0.37	3046
Father entrepreneur (likelihood)***	-1.08	8.79	46232	-0.25	4.24	3046

*Available from 1985 onwards excluding 1991. **Available from 1988 onwards excluding 1991. ***Available from 1981 onwards. The text in Subsection 3.3.2 together with footnote 19 explain how this likelihood is calculated (and how negative values may result).

Second, we notice that both the mean and the standard deviation of the distribution of hourly incomes are higher for entrepreneurs than for employees.[9] This common observation is (partly) explained by the absence of a 'minimum wage' and preformed salary scales

[9]The medians of the hourly income distributions of entrepreneurs and employees show the same pattern as the means, being 9.6 and 8.1, respectively (not tabulated).

for entrepreneurs. Third, the average level of education that individuals complete in the USA is thirteen years (slightly above high school level), being equal for entrepreneurs and employees.[10]

3.3.2 Empirical methodology

Our aim is to estimate the returns to education for entrepreneurs and employees as consistently as possible. To this end we estimate the income equation under (3.1) by means of a random effects (RE) model. The RE-model uses the observations that are distributed over years and individuals in an appropriate manner such that yearly observations for one individual are not treated as independent.[11]

$$W_{it} = \beta S_{it} + \gamma E_{it} + \delta SE_{it} + \eta X'_{it} + \zeta EX'_{it} + c_i + u_{it} \tag{3.1}$$

In equation (3.1), W_{it} represents the log hourly earnings for individual i in year t, S_{it} the number of years of formal schooling obtained at t, E_{it} is a dummy indicating whether person i is an entrepreneur in year t, and SE_{it} is an interaction of the dummy E_{it} and the variable S_{it}, such that its coefficient, δ, indicates the magnitude of the difference in returns to education between entrepreneurs and employees. Furthermore, X'_{it} is a vector including the control variables of Table 3.1, as well as dummies controlling for cohort effects, age effects[12] and macroeconomic shocks using the method developed by Deaton (2000). This method transforms the year dummies, age dummies and birth year dummies such that the year effects add to zero, and are orthogonal to a time trend.[13] The vector EX'_{it} is the vector of interactions of all components of the vector X'_{it} and the dummy E_{it}. Finally, c_i is an unobserved individual-specific effect, and u_{it} a white noise error term.

However, as was pointed out already, the variables S_{it} – the number of years of schooling – and E_{it} – being an entrepreneur – are likely to be endogenous in the income equation. Therefore, two first-stage equations are estimated, such that the predicted values resulting from these equations can be used as the instrumented values of education and entrepreneurship

[10]A discussion of some of the control and instrumental variables tabulated is provided when discussing their usage.

[11]A fixed-effects model to deal with unobserved heterogeneity cannot be used for this purpose, because formal education is (almost) time-invariant for labor market participants. The only source of variation in the number of years of schooling per individual over time is that some respondents worked for a while before finishing their formal education.

[12]We use age instead of experience in the earnings equation, as in Harmon and Walker (1995). Experience is a negative function of education, and is therefore endogenous in equation (3.1).

[13]These transformed dummies are included in all regression models, but their coefficients will not be shown in the tables reporting the estimation results.

in equation (3.1).

$$S_{it} = \gamma F_i' + \lambda E_{it} + \varpi EF_{it}' + \kappa X_{it}' + \varsigma EX_{it}' + \mu_i + \epsilon_{it} \tag{3.2}$$

$$E_{it} = \eta G_i' + \theta X_{it}' + \omega S_{it} + \vartheta_i + \varepsilon_{it} \tag{3.3}$$

Equations (3.2) and (3.3) represent the first-stage equations. In (3.2), schooling is denoted by S_{it}, and its set of identifying instruments by F_i'. Entrepreneurship status is denoted by E_{it}. The vector EF_{it}' is the vector of interactions of all identifying instruments of the vector F_{it}' and the dummy E_{it}. X_{it}' is a vector of control variables and EX_{it}' is the vector of interactions of all components of the vector X_{it}' and the dummy E_{it}. In (3.3), entrepreneurship status is denoted by E_{it}, and its instruments by G_i'. X_{it}' is a vector of control variables as in (3.2) and S_{it} again denotes schooling. Finally, μ_i and ϑ_i represent the unobserved individual-specific effects, and ϵ_{it} and ε_{it} the white noise error terms in the respective equations.

Variables can be used as identifying instruments if they pass the criteria for quality and validity. The quality criterion comes down to requiring a sufficient correlation between the set of identifying instruments and the endogenous regressor, S_{it} and E_{it} in this case. Instruments are valid if they affect performance via the education equation only. A set of instruments therefore passes the (Sargan) validity/over-identification test if it is not correlated with the error term in the earnings equation.

A set of four identifying instruments for education is extracted from the NLSY data: (1) "Magazines present in the household at age 14", (2) "Library card present in the household at age 14", (3) "The presence of a stepparent in the household", (4) "Number of siblings in the household". All these instruments are expected to have a significant effect on the number of years of education attained. The descriptive statistics of these variables can be found in Table 3.1.[14]

Having magazines and/or a library card in the household signifies access to reading/studying material and might inspire the child to learn more, which in turn influences the total amount of education that can be obtained. In contrast, we expect that the presence of a stepparent reduces the level of education: A stepparent in the household increases the probability that there has been turmoil (divorce or death of a parent) in the child's learning environment, thereby influencing the child's educational attainment negatively. The number of siblings is expected to have a negative effect on the amount of education obtained (Black, Devereux and Salvanes, 2005; Parker and Van Praag, 2006). The lower average amount of (both financial

[14] Blackburn and Neumark (1993) have used an IV-approach to estimate the returns to education (for employees) based on the NLSY data. They use a broad set of identifying instruments, including the set of instruments we use and the education levels of the respondents' parents.

and non-financial) resources transferred from the parents to children with more siblings has a negative impact on the education level of the child.[15]

There are two sorts of critique on the aforementioned instruments. First, family background variables may, besides influencing education, have a direct impact on the labor market performance of the respondent. In order to minimize this direct impact, which would turn the instrument invalid, the education levels of the parents are used as control variables in all equations - rather than as additional identifying instruments, as is common when using this type of instruments. Moreover, controlling for indicators of ability decreases the likelihood that the identifying instruments do in fact measure the unobserved (inherited) talents of the respondent. Second, the variable 'number of siblings', which is expected to affect the education level partly through the parents' possibility of spending resources on the child's education, might be invalid too: The availability of (inherited) resources could also have a direct effect on the child's ability to invest in a new business, thereby diluting any capital constraints and thus increasing business earnings (see Parker and Van Praag, 2006). We address this critique in Section 5, where we find that the estimation results are invariant to the inclusion of a direct measure of assets in the earnings equation.[16] Notwithstanding these critiques, the set of identifying instruments for education passes the tests of quality and validity, as will be shown in Section 4.

In addition to the instruments for schooling, an instrument is required for selection into entrepreneurship. Correcting for this kind of selectivity has proven to be difficult for it requires an exclusion restriction that affects the entrepreneurship decision but not earnings (Rees and Shah, 1986; Gill, 1988; MacPherson, 1988; De Wit and Van Winden, 1989; Taylor, 1996; Clark and Drinkwater, 1998).[17] We take two different and imperfect routes. However, if both lead to the same result, we feel more confident about our findings.

The first route and instrument we propose is the entrepreneurship status of the father, as in Taylor (1996) and De Wit and Van Winden (1989). Several studies have demonstrated

[15] Recent findings from Rodgers, Cleveland, Van den Oord and Rowe (2000) indicate that the previously imagined relationship between family size and child-IQ is non-existing. This strengthens the view that family size is a valid instrument.

[16] The measure of assets is excluded from the basic set of regressions since it is available for fewer years , and would therefore limit the sample size (see Table 3.1).

[17] Instruments used include the number of children of the respondent (as in Rees and Shah (1986); MacPherson (1988); De Wit and Van Winden (1989); Clark and Drinkwater (1998)), income from dividends, rents or interest (as in Gill (1988)), a self-employed father (as in De Wit and Van Winden (1989); Taylor (1996)), or whether the individual owns or rents the house (s)he occupies (as in Clark and Drinkwater (1998)). Most authors acknowledge the imperfections of their instruments, whereas some others only use them implicitly as instrument. However, so far, instruments affecting entrepreneurship choices but not outcomes, based on tax or other reforms that vary over time and/or geographically, have not yet been identified by researchers, including ourselves.

that persons who have entrepreneurial fathers have a higher probability of becoming an entrepreneur themselves (Taylor, 1996, 1999, 2001; Laferrère, 2001; Fairlie and Robb, 2007). In addition, having an entrepreneur as father is not associated with better performance (Taylor, 2001; Fairlie and Robb, 2007). The entrepreneurship status of the father might therefore be a good instrument. However, since the NLSY does not contain information on the parents' entrepreneurship status, a proxy is constructed in the following way. First, twenty different profession groups are distinguished in the sample, both for the respondents and their fathers.[18] Second, the (respondents') sample proportion of entrepreneurs is calculated for each of these 20 profession groups. For some professions, such as professional services, this proportion is higher than for other professions (such as teaching). The sample proportion is used as a proxy for the probability that a person with a certain profession is an entrepreneur, which is denoted by C_p.[19] Third, C_p is allocated to each father's profession p.[20] C_p for each respondent's father is treated as the likelihood that the father is/was an entrepreneur. This is used as identifying instrument. In addition, and to increase the fit, a variable is included in the entrepreneurship equation that interacts the respondent's age with C_p: The data show that the impact of C_p on the respondent's entrepreneurship status E_{it} increases over time. This variable can be treated as a second identifying instrument.

The second route we propose uses the religion of the individual as instrument for the entrepreneurship selection equation. Two contradicting theoretical arguments motivate a link between entrepreneurship and religion (Drakopoulou Dodd and Seaman, 1998). First, religious people would be over-represented within the group of entrepreneurs since both religion and entrepreneurship develop core values as thrift, hard work and independence. Second, people adhering to (stricter) religions would be under-represented within the group of entrepreneurs since entrepreneurship is time consuming and would leave little time for religious activities. Empirical evidence reveals a negative relation between adhering to strict religions and entrepreneurship (Van Praag and Van Ophem, 1995; Drakopoulou Dodd and Seaman, 1998), which supports the latter argument. Moreover, empirical evidence shows

[18]Since a person's occupation can vary over time, the variable identifies the person's occupation with the longest tenure. In this way we hope to capture the person's "true profession".

[19]More precisely, $C_p = \frac{A_p - A}{A}$, where A_p is the fraction of entrepreneurs in a profession p and A is the fraction of entrepreneurs in the total population. This leads to a positive score for professions that 'generate' a higher proportion of entrepreneurs and a negative score for those professions that are less prone to be performed as an entrepreneur. These numbers are then attributed to the parents of the respondents based on their professions.

[20]The real percentage of entrepreneurs in profession p at the time the father was working might be different from the calculated value of C_p that pertains to statistics based on the next generation. This will be problematic if the relative percentages of entrepreneurs in each of the twenty professions changed. We assume this not to be the case.

that the relationship between intensely adhering to a religion and an entrepreneur's earnings is insignificant, implying that the intensity of one's religious affiliation would be a valid instrument (Drakopoulou Dodd and Seaman, 1998, Table 4). Since the data are not informative about the religious intensity of respondents, but only about their adherence to a specific religion, the empirical proxy for religion that we use is a dummy variable indicating whether respondents 'adhere to one of the stricter religions', i.e., Lutheran or Methodist (see Van Praag and Van Ophem (1995)). This dummy variable is treated as identifying instrument.

3.4 Estimation results

In this section, we discuss the main estimation results from applying the discussed empirical methodology to the panel dataset. As a benchmark, we estimate an earnings equation as in equation (3.1) by means of random effects (RE), where an individual's education level and entrepreneurship status are treated as exogenous and ability controls are excluded. The first two columns of Table 3.2 show the results. The RE-estimated returns to education are 6.9 percent for entrepreneurs and 6.0 percent for employees. The returns are thus somewhat higher for entrepreneurs than for employees, and this difference is marginally significant. This is in accordance with previous studies using OLS estimation on USA data (Fredland and Little, 1981; Tucker, 1985, 1987; Evans and Leighton, 1990; Robinson and Sexton, 1994). Before discussing the remaining estimates, we shall first improve the model by including a proxy (ASVAB) for general ability (next set of two columns in Table 3.2) and then by using IV (the last two columns of the table).

Table 3.2: (Second stage) earnings equations (RE)

	Benchmark RE		Ability control		IV-Education	
Variable	**Coeff.**	**(SE.)**	**Coeff.**	**(SE.)**	**Coeff.**	**(SE.)**
Education	0.060**	(0.002)	0.055**	(0.002)	0.099**	(0.012)
E*Education	0.009†	(0.005)	0.009†	(0.005)	0.084**	(0.025)
General Ability			0.064**	(0.005)	0.042**	(0.008)
E*General Ability			0.008	(0.010)	-0.020	(0.013)
E	-0.754†	(0.445)	-0.723	(0.445)	-0.095	(0.500)
Male	0.237**	(0.010)	0.205**	(0.011)	0.224**	(0.011)
E*Male	0.408**	(0.020)	0.405**	(0.020)	0.439**	(0.023)
Married	0.060**	(0.005)	0.059**	(0.005)	0.061**	(0.005)
E*Married	-0.066**	(0.019)	-0.067**	(0.019)	-0.082**	(0.020)
Not Healthy	-0.053**	(0.013)	-0.053**	(0.013)	-0.053**	(0.014)
E*Not Healthy	0.005	(0.047)	0.001	(0.046)	0.022	(0.048)
Live outside city	-0.081**	(0.007)	-0.083**	(0.007)	-0.084**	(0.007)
E*Live outside city	0.002	(0.010)	0.003	(0.010)	-0.011	(0.011)
Live in South	-0.062**	(0.012)	-0.056**	(0.012)	-0.058**	(0.010)
E*Live in South	0.104**	(0.021)	0.105**	(0.021)	0.105**	(0.022)
Hispanic	0.057*	(0.027)	0.080**	(0.027)	0.036	(0.027)
E*Hispanic	-0.053	(0.060)	-0.049	(0.060)	-0.157*	(0.069)
Black	-0.112**	(0.018)	-0.030	(0.019)	-0.057**	(0.019)
E*Black	0.009	(0.045)	0.017	(0.047)	-0.027	(0.049)
Locus of control	0.018**	(0.003)	0.014**	(0.003)	0.007*	(0.003)
E*Locus of control	0.018**	(0.005)	0.017**	(0.005)	0.006	(0.006)
Mother education	0.008**	(0.003)	0.006*	(0.003)	-0.004	(0.004)
E*Mother education	0.003	(0.005)	0.002	(0.005)	-0.019*	(0.009)
Father education	0.011**	(0.002)	0.009**	(0.002)	0.002	(0.003)
E*Father education	0.017**	(0.004)	0.016**	(0.004)	0.004	(0.005)
Intercept	0.589**	(0.121)	0.649**	(0.121)	0.884**	(0.138)
N	50268		50268		50268	
R^2 Within	0.45		0.45		0.44	
R^2 Between	0.46		0.47		0.46	
R^2 Overall	0.44		0.45		0.43	

Significance levels : † : 10% $*$: 5% $**$: 1%
E denotes Entrepreneur.

Including the ability proxy into the earnings equation leads to a decrease of the estimated returns to education for both entrepreneurs (from 6.9 to 6.4 percent) and employees (from 6.0 to 5.5 percent). The difference between the returns to education for entrepreneurs and employees remains the same and marginally significant. The other coefficients (not yet discussed) do not change either. The drop in returns supports theory (see footnote 2).

The next step is to instrument education with the discussed set of family background variables and apply IV-estimation.[21] The results of estimating the first-stage equation (3.2) are presented in the first two columns of Table 3.3. All family background variables are significant and about 32 percent of the variation in education is explained.

To assess the credibility of the results that will be obtained by using the selected identifying instruments, we check whether the proposed set of identifying instruments is (i) of sufficient quality, (ii) valid, and (iii) whether instrumentation is relevant at all. The results from the tests of validity and relevance do critically depend on the choice of regressors to be used in the second-stage earnings equation (see below). To test the first criterion, we performed a Chi-square test supporting the quality of the set of identifying instruments ($\chi^2_{(df8)} = 216.78$, and $p = 0.000$).

In non-panel IV-estimation the Sargan-test (1988) is used to test the second criterion – i.e., instrument validity and over-identification. Since the Sargan test is not available for RE-models, we follow a different route. As the aim of the Sargan-test is to test whether the set of identifying instruments is uncorrelated with the error term in the earnings equation, we regress the residuals of the RE-IV-regression on our instruments and the control variables. All identifying instruments turn out to be insignificant. Moreover, the overall R^2 is close to zero (R^2=0.0001). Hence, the use of this particular set of family background variables as identifying instruments is valid (given the complete set of independent variables used in both the first and second-stage equations).[22]

Third, we perform a Hausman test (1978) to analyze whether it is relevant to use IV in the first place. If not, implying that education is exogenous, RE-estimates would not be biased due to endogeneity. We find that instrumentation of the education variable is necessary – i.e., education is endogenous ($\chi^2_{(df1)} = 14.31$, and $p = 0.0002$).

[21] A two-step method generates unbiased coefficients, the standard errors, calculated using $\widehat{S_{it}}$ instead of S_{it}, are however not unbiased. To get the correct standard errors we use the 'XTIVREG package' provided by STATA. An additional advantage of using this package is that it uses Feasible Generalized Least Squares (FGLS) to estimate the first stage equation. Using FGLS makes sure the correct error structure is estimated even in the case of small or no variation over time in the dependent variable (schooling in our case).

[22] If we 'improperly' compute the OLS Sargan-test ($\chi^2_{(df7)} = 5.027$, and $p = 0.657$), the same result is found. The set of identifying instrument is also tested to be valid when capital-constraint related variables, such as residence value, stock value, value of assets over 500 dollar, value of inheritances and total savings, were included in the earnings equation. This renders additional support for the validity of the identifying instrument 'number of siblings' that could perhaps affect earnings through its effect on capital constraints.

Table 3.3: First stage equations of education and selection into entrepreneurship

Variable	**Education** **Coeff.**	**(SE.)**	**Selection (father)** **Coeff.**	**(SE.)**	**Selection (religion)** **Coeff.**	**(SE.)**
Education			0.058	(0.071)	0.074	(0.061)
General Ability	0.455**	(0.029)	-0.131**	(0.045)	-0.127**	(0.040)
E*General Ability	0.040**	(0.012)				
E	0.322	(0.541)				
Male	-0.413**	(0.057)	0.481**	(0.068)	0.470**	(0.061)
E*Male	-0.008	(0.025)				
Married	0.022**	(0.006)	0.263**	(0.038)	0.257**	(0.035)
E*Married	0.036	(0.023)				
Not Healthy	-0.014	(0.016)	0.156†	(0.090)	0.195*	(0.081)
E*Not Healthy	0.054	(0.057)				
Live outside city	-0.056**	(0.010)	0.236**	(0.052)	0.220**	(0.048)
E*Live outside city	0.029*	(0.012)				
Live in South	0.084	(0.062)	-0.016	(0.081)	-0.066	(0.061)
E*Live in South	0.004	(0.027)				
Hispanic	1.149**	(0.144)	-0.379*	(0.180)	-0.426**	(0.150)
E*Hispanic	0.118	(0.076)				
Black	0.740**	(0.103)	-0.641**	(0.143)	-0.692**	(0.131)
E*Black	0.040	(0.058)				
Locus of control	0.133**	(0.014)	0.016	(0.022)	0.005	(0.018)
E*Locus of control	-0.006	(0.006)				
Mother education	0.189**	(0.014)	0.031	(0.022)	0.016	(0.019)
E*Mother education	-0.015*	(0.007)				
Father Education	0.140**	(0.010)	-0.014	(0.016)	-0.009	(0.015)
E*Father Education	0.004	(0.005)				
Age			0.168**	(0.031)	0.181**	(0.028)
Age squared			-0.002**	(0.000)	-0.002**	(0.000)
Year			0.006	(0.017)	0.007	(0.013)
Family background variables used as instruments for education						
Magazines	0.599**	(0.065)				
E*Magazines	0.027	(0.028)				
Library	0.325**	(0.068)				
E*Library	-0.001	(0.030)				
# Siblings	-0.078**	(0.014)				
E*# Siblings	-0.012*	(0.006)				
Stepparent	-0.761**	(0.118)				
E*Stepparent	0.204**	(0.051)				
Family background variables used as instruments for entrepreneurship						
Father entrepreneur			0.271*	(0.135)		
Father entrepreneur*Age			0.001*	(0.000)		
Strictly religious					-0.213**	(0.066)
Intercept	5.381**	(0.268)	315.898**	(84.654)	-19.682	(25.120)
N	50268		42425		50148	
R^2 Within	0.14		n.a.		n.a.	
R^2 Between	0.35		n.a.		n.a.	
R^2 Overall	0.32		n.a.		n.a.	

Significance levels : † : 10% $*$: 5% $**$: 1%
E denotes Entrepreneur.

The last two columns of Table 3.2 show the second-stage IV-results estimated with 2SLS. Applying IV results in significantly higher estimates of the returns to education. For employees, the returns jump from 5.5 percent to 9.9 percent. This increased estimate of the returns to education is consistent with previous research, using various sets of identifying instruments (Ashenfelter et al., 1999). More specifically, in a comparable fashion, Blackburn and Neumark (1993) use the NLSY to estimate the returns to education. They also find that the returns to education for employees are 10 percent.

A novel finding is the greater jump in the returns to education for entrepreneurs from 6.4 percent to 18.3 percent. This leads to the remarkable result that the returns to education for entrepreneurs are a significant 85 percent higher than the comparable returns for employees.[23]

Our next step is to correct for the endogenous selection into entrepreneurship. As discussed above, we constructed two separate instrument sets to address this problem – i.e., (1) 'likelihood father entrepreneur' and its interaction with age, and (2) 'adhering to a strict religion'. The first-stage (RE-probit) results from these attempts are shown in the middle and right columns of Table 3.3. The likelihood that an individual's father is an entrepreneur significantly increases the probability that an individual is observed to be an entrepreneur; and this effect is significantly stronger, the older the individual is. Individuals adhering to stricter religions are significantly less likely to be entrepreneurs.

Again, we first assess the empirical suitability of the instruments before we continue. When it comes to the quality of the instruments, the Chi-square test results show that 'father entrepreneur' (combined with its interaction with age) is a weak instrument ($\chi^2_{(df2)} = 4.03$, and $p = 0.045$), and that 'religion' is of sufficient quality ($\chi^2_{(df1)} = 10.38$, and $p = 0.001$). In order to assess the validity of the first set of instruments (likelihood father entrepreneur, and its interaction with age), we proceed in the same fashion as for education.[24] The residuals of the RE-IV-equation explaining earnings are regressed on the set of instruments and the control variables. Both instruments have an insignificant relation with the residuals, and the R^2 of the regression is 0.0002.[25] Unfortunately, an RE-probit does not produce residuals necessary to compute the Hausman-test required for testing the relevance of instrumenting the entrepreneurship variable. In all, we conclude that the sets of instruments used for endogenizing the selection equation are far from perfect, and we shall therefore treat the second-stage results with caution.

[23]Using RE, the returns were 15 percent higher for entrepreneurs than for employees.

[24]Since religion exactly identifies entrepreneurship, an over-identification test is not possible.

[25]As before, we also computed the OLS Sargan-test ($\chi^2_{(df1)} = 8.485$, and $p = 0.0004$). Although this test cannot be easily transferred to a RE-setting, the statistics indicate that the instruments are not valid.

Table 3.4: Earnings equations where both education and entrepreneurship are instrumented

Variable	IV-education-entrepreneur(father) Coeff.	(SE.)	IV-education-entrepreneur(religion) Coeff.	(SE.)
Education	0.103**	(0.016)	0.110**	(0.015)
E*education	0.080**	(0.005)	0.077**	(0.004)
E	-0.005	(0.070)	0.128†	(0.066)
General ability	0.037**	(0.014)	0.047**	(0.012)
Male	0.244**	(0.037)	0.193**	(0.033)
Married	0.038*	(0.019)	0.009	(0.018)
Not Healthy	-0.039*	(0.018)	-0.071**	(0.018)
Live outside city	-0.076**	(0.018)	-0.104**	(0.016)
Live in South	-0.063**	(0.014)	-0.047**	(0.013)
Hispanic	-0.018	(0.045)	0.040	(0.041)
Black	-0.069	(0.052)	-0.002	(0.050)
Locus of control	0.008*	(0.004)	0.005	(0.004)
Mother education	-0.006	(0.005)	-0.011*	(0.005)
Father education	0.004	(0.003)	0.001	(0.003)
Intercept	0.845**	(0.139)	1.847**	(0.267)
N	42425		50148	
R^2 Within	0.45		0.44	
R^2 Between	0.40		0.40	
R^2 Overall	0.42		0.42	

Significance levels : † : 10% $*$: 5% $**$: 1%
E denotes Entrepreneur.

Table 3.4 shows the second-stage results when entrepreneurial status is instrumented.[26] The first set of columns reveals the results of using the likelihood of having an entrepreneurial father (and its interaction with age) as instruments; the next set of columns indicates the results of using religion as identifying instruments for entrepreneurship. The difference in returns to education between employees and entrepreneurs is of similar size in both cases in comparison to the case where only education was treated as endogenous and entrepreneurship as exogenous. The returns to education for entrepreneurs are again around 18 percent, while the returns to education for employees are around 10-11 percent. This suggests that, on average, employees would benefit from higher returns to education as entrepreneurs. The current entrepreneur status, rather than whether people are of the "entrepreneurial type", seems to explain the difference in the returns to education. This conclusion is also supported by an additional analysis. Re-estimating the earnings equation on a sample of individuals

[26]In this equation, the only variable interacted with the respondent's (instrumented) entrepreneurship status E_{it} is education. Omitting all the other interactions with E_{it} does not lead to any biases since the instrumented E_{it} is completely exogenous.

who have been observed as both entrepreneurs and employees in the sample period rendered the same difference in returns to education between entrepreneurs and employees. If "entrepreneurial types" could simply do more with their education, we should observe a smaller difference in returns to education between entrepreneurs and employees in this particular sub-sample.

Based on the fact that the instruments used for entrepreneurial status are not too convincing and given the finding that correcting for selectivity leaves the key results unchanged, the chapter is continued with the estimates that have been generated by instrumenting education only.

Before trying to understand why entrepreneurs benefit more from their education than employees, by checking the robustness of the result against various alternative explanations, we first discuss the other coefficients reported in Table 3.2. The discussion is based on the results reported in the last two columns of the table.

Table 3.2 shows that an increase of one standard deviation in general ability increases one's earnings by approximately four percent, irrespective of whether one is an entrepreneur or an employee. Males earn significantly higher incomes than females, confirming previous findings for both segments of the labor market. The gender effect differs largely across labor market segments. Male wage employees earn 22 percent more than their female counterparts. The comparable difference between male and female entrepreneurs is 66 percent. This large gender effect for entrepreneurs vis-à-vis employees is consistent with previous studies (Moore, 1983; Tucker, 1987; De Wit and Van Winden, 1989; Dolton and Makepeace, 1990; Robinson and Sexton, 1994).

Interestingly, the correlation between being married and income is positive for employees, but not for entrepreneurs: The income of married employees is 6.1 percent higher than the income of single employees, whereas this difference is an insignificant -2.1 percent for entrepreneurs. Previous findings support this result (Moore, 1983; Tucker, 1987; Gill, 1988; Dolton and Makepeace, 1990; Evans and Leighton, 1990). People with health limitations earn 5.3 percent less, and this difference applies both to entrepreneurs and employees. People living outside cities earn 8.4 percent less than others, irrespective of their occupational status, i.e. entrepreneur or employee. Notably, living in the South leads to lower earnings for employees and higher earnings for employees. Blacks earn 5.7 percent lower incomes than whites, where the effect is of the same order of magnitude for entrepreneurs and employees.[27]

[27] Support for a difference in the effect of race on earnings between the groups in previous studies is not clear (see Fairlie and Meyer, 1996; Moore, 1983; Fredland and Little, 1981; Rees and Shah, 1986; Evans and Leighton, 1990; Dolton and Makepeace, 1990). However, the comparability of these studies with ours is only limited because we distinguish blacks as an ethnic group explicitly, unlike the others. In so doing, we try to take out the often mixed effect of other ethnicities.

Internality of locus of control, which is an indicator of the individual's perception that (s)he is in control of the environment,[28] and parental education all have small or insignificant effects on earnings for both entrepreneurs and employees.

3.5 Why are entrepreneurs' returns to education higher?

This section is devoted to finding an explanation for the result that the estimated return to education is significantly higher for entrepreneurs than for employees. We check the validity of six possible explanations. The benchmark results are the estimated returns to education in the last columns of Table 3.2.

3.5.1 Risk premium

The first check relates to the question as to whether the difference in returns to education between entrepreneurs and employees can be attributed to a *risk premium* required by higher educated entrepreneurs. More highly educated individuals would perhaps require higher risk premia for being an entrepreneur if higher educated individuals experience more additional income risk as an entrepreneur compared to an employee vis-à-vis lower educated individuals. This check proceeds in three steps. First, by regressing the individual (time) variances of the residuals of the income equations presented above on *entrepreneurs'* education levels and control variables, we find no significant education effect. Hence, the variance of entrepreneurial income, our indicator of risk, is not higher for more highly educated individuals, all else equal. Second, estimating the same equation for employees reveals a significant positive coefficient for education. Third, the variance in earnings is lower for employees than for entrepreneurs, at all possible education levels. These three observations together imply that entrepreneurs are exposed to more income risk than employees are, but that the difference is a decreasing rather than an increasing function of education. Thus, we conclude that the higher returns to education for entrepreneurs are not a kind of risk premium.

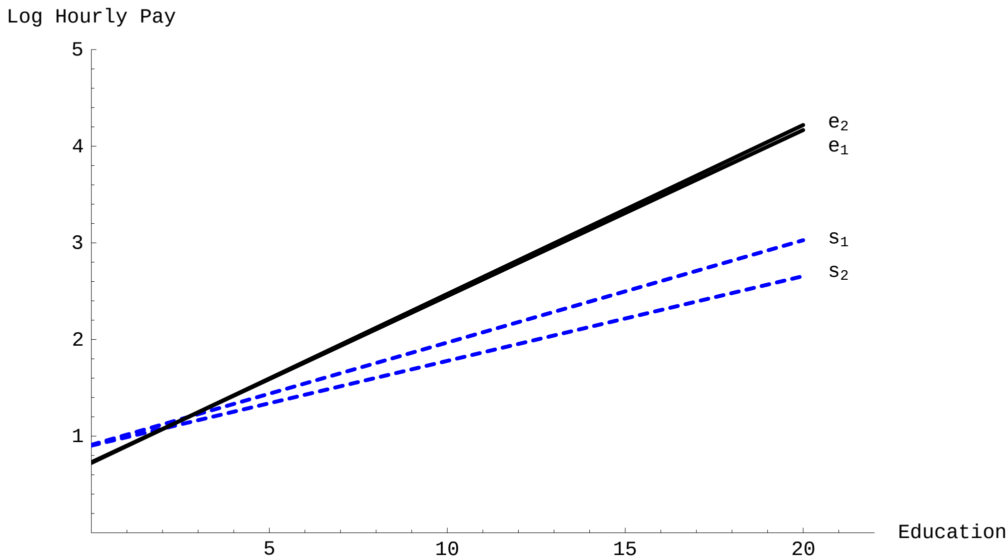
3.5.2 Underreporting of earnings

The second check pertains to recent evidence that entrepreneurs underreport their incomes more than employees do. In general, underreporting is not a problem for the estimation of the returns to education. As long as underreporting and education are unrelated, the estimated magnitude of the returns to education is unbiased. However, recent evidence by

[28] To be discussed in detail later on.

Lyssiotou et al. (2004) shows that this might not be the case. Blue-collar entrepreneurs underreport their incomes to a higher degree than white-collar entrepreneurs. Since blue-collar entrepreneurs have a lower average level of education than white-collar entrepreneurs the returns to education estimate for the total population of entrepreneurs could be upward biased. This in turn might explain the difference in returns to education between employees and entrepreneurs that we established.

If underreporting has an effect on our estimation results we would expect that the difference in returns to education between entrepreneurs and employees would be smaller if estimates are obtained for the group of blue-collar and white-collar workers separately. We estimate a three-way interaction to see if the differences in returns to education between entrepreneurs and employees diminishes when estimating the returns on the separate samples. Figure 3.1 shows that this is not the case.[29] Thus, we conclude that underreporting does not influence the returns to education estimates.



Note. s denotes salaried employees and e entrepreneurs. Subscript 1 and 2 are blue-collar workers and white-collar workers, respectively.

Figure 3.1: Returns to education: Blue- and white-collar employees and entrepreneurs

3.5.3 Returns to capital

As a third check, we address the issue raised by Fairlie (2005*b*) that some entrepreneurs in the NLSY might have erroneously included the returns to (business) capital in their reported

[29]The differences between entrepreneurs and employees are all significant at the one percent level and as large as before.

income. This could explain the result if more highly educated entrepreneurs have higher returns to capital than lower educated entrepreneurs. As proposed by Fairlie (2005*b*), income possibly needs to be adjusted for entrepreneurs who receive a business income from unincorporated businesses (others receive a 'wage' from their incorporated business that excludes returns to capital). Following Fairlie (2005*b*), the adjustment implies that five percent of their total business value is subtracted from their business incomes. As was indicated above, the variable 'total business value' has not been measured in every year, resulting in a smaller sample size. Therefore, Table 3.A-1 in Appendix 3.A shows the estimation results not only based on the adjusted entrepreneurial incomes, but also without this adjustment for the same sample, allowing a proper comparison. The comparison indicates that the adjustment for capital returns does not at all reduce the difference in returns between entrepreneurs and employees.

3.5.4 Hours worked

The fourth check concerns the number of hours worked by entrepreneurs and employees. Could the difference in returns to education (in terms of hourly earnings) between entrepreneurs and employees be explained by the inclusion of part-time entrepreneurs and employees in the sample? For instance, this could explain the result if working part-time is punished more heavily in terms of hourly earnings for entrepreneurs than for employees and if part-time workers have lower levels of education. Upon excluding all individual-year observations working less than 1000 hours per annum, the difference in returns to education between entrepreneurs and employees does not decrease at all, as is clear from the first two columns in Table 3.A-2 in Appendix 3.A.

3.5.5 Professional workers

The fifth check is based on the idea that professional workers such as lawyers and medical doctors have high earnings, are highly educated and are often self-employed. This might drive the result. However, as the right half of Table 3.A-2 in Appendix 3.A shows, excluding professional workers[30] from the sample does not decrease the estimated difference between entrepreneurs and employees.

[30]Including accountants, actuaries, pharmacists, health-diagnosing occupations, therapists, lawyers, dieticians and architects.

3.5.6 Locus of control

So why is education more valuable for entrepreneurs? A straightforward organization-oriented explanation could be that entrepreneurs have more freedom than employees to optimize their employment of education. Entrepreneurs are not constrained by rules from superiors and can decide on how to employ their education in such a way that its productive effect is the highest. In contrast to the entrepreneur, the organizational structure surrounding an employee makes it perhaps more difficult, or even impossible, to utilize education productively. Organizations cannot adapt their structure to every individual due to organizational inertia and individual incompatibilities. As a consequence, entrepreneurs are in a position to better control the profitable employment of their education. This might be an explanation for the higher returns to education for entrepreneurs vis-à-vis employees.

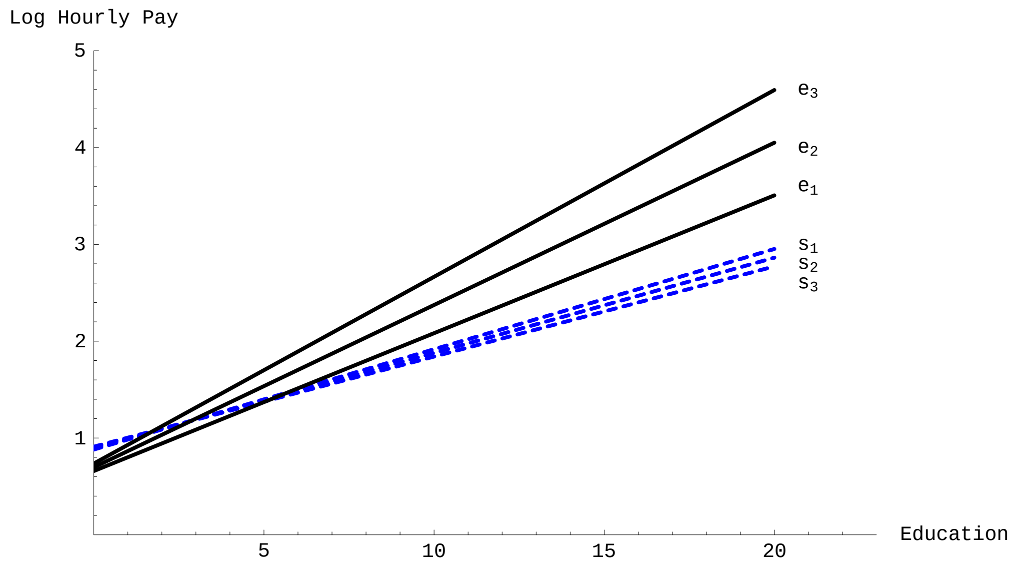
Ideally, we would like to test this explanation directly by randomly allocating entrepreneurs and employees to flexible and less flexible organizations (assuming that flexibility leads to more control) and observe the differences in returns to education between people working in the two groups of organizations. Unfortunately, such an experiment is very difficult to realize - if at all.

However, if it is true that a better control of the environment influences the possibility to optimize the returns to education, it might also be true that individuals' perceived control of the environment affects their returns to education. Those entrepreneurs and employees having the perception that they are in control of their environment should then experience higher average returns to education than others. This would support the control-related explanation indirectly. An individual's perceived control of the environment is measured by psychologists through the personality trait called 'locus-of-control'. This measure, introduced by Rotter in 1966 in the context of his social learning theory, is included in the NLSY.[31] Locus of control is defined as an individual's general expectancy of the outcome of an event as being either within or beyond her or his personal control and understanding (Rotter, 1966). Individuals with an external locus-of-control personality tend to perceive an event as beyond their control, and attribute the outcomes of the event to chance, luck, as under control of powerful others, or as unpredictable. Individuals with an internal locus of control tend to believe that events are contingent upon their own behavior or relatively permanent characteristics. In the psychological literature, there is ample evidence that locus of control is a fundamental and stable personality trait, with clear behavioral consequences (Boone

[31]The NLSY includes an abbreviated version of the original Rotter scale. The test was administrated in 1979, before the respondents had made any major decisions regarding their jobs or occupations. We re-scaled the Rotter score in such a way that 0 reflects high external locus of control and 10 indicates high internal locus of control.

and De Brabander, 1993; Boone, Van Olffen and Van Witteloostuijn, 2005).

Hence, we explore this control-related explanation by testing whether entrepreneurs and employees with an internal locus of control generate higher returns from their education than persons with an external locus of control. However, given that the control hypothesis would be supported, we might find that only entrepreneurs benefit from a higher return to education if they have more internally oriented 'locus-of-control' beliefs: The organizational structure in which employees operate might, on average, even turn it impossible to benefit from their internal 'locus-of-control' beliefs in terms of a higher return to their education. Thus, although we use one's 'locus-of-control' beliefs as a proxy for the extent to which one has an entrepreneurial position, i.e. one's control over the environment, the proxy might be ineffective for employees. Table 3.5 and Figure 3.2 show indeed that the returns to education are higher for individuals with a more internally oriented locus of control than for individuals with an external locus of control. However, this holds true for entrepreneurs only. We conclude that control matters and that it is likely to be an explanation for the higher returns to education obtained by entrepreneurs. Entrepreneurs who feel more in control of their environment extract higher returns from their investment in education. For employees, we do not find the same result, because the organizational constraints they experience possibly prevent the 'in control types' from 'controlling' the profitable employment of their human capital, i.e. education, in such a way that they cannot benefit from it.



Note. s denotes salaried employees and e entrepreneurs. Subscript 1, 2 and 3 are respectively, external locus, intermediate locus and internal locus (1 sd difference between 1, 2 and 3).

Figure 3.2: Returns to education evaluated at three levels of locus of control

Table 3.5: Second stage earnings equations including three-way interaction between education, entrepreneurship status and locus-of-control

Variable	**Coefficient**	**(Std. Err.)**
Education	0.098**	(0.010)
E	-0.198	(0.501)
Locus of control	0.007**	(0.003)
E*Education	0.069**	(0.023)
E*Locus of control	0.012†	(0.006)
Education*Locus of control	-0.003	(0.002)
E*Education*Locus of control	0.015**	(0.006)
General ability	0.043**	(0.007)
E*General ability	-0.015	(0.013)
Male	0.224**	(0.010)
E*Male	0.438**	(0.023)
Married	0.065**	(0.005)
E*Married	-0.083**	(0.020)
Not Healthy	-0.058**	(0.014)
E*Not Healthy	-0.002	(0.049)
Live outside city	-0.096**	(0.007)
E*Live outside city	-0.007	(0.011)
Live in South	-0.054**	(0.009)
E*Live in South	0.092**	(0.022)
Hispanic	0.038	(0.023)
E*Hispanic	-0.162*	(0.068)
Black	-0.060**	(0.016)
E*Black	-0.035	(0.049)
Education mother	-0.004	(0.003)
E*Education mother	-0.015†	(0.008)
Education father	0.002	(0.002)
E*Education father	0.003	(0.005)
Intercept	0.896**	(0.135)
N	50268	
R^2 Within	0.44	
R^2 Between	0.46	
R^2 Overall	0.44	

Significance levels : † : 10% $*$: 5% $**$: 1%

E denotes Entrepreneur.

3.6 Discussion and conclusion

We have estimated the effect of education on the performance of entrepreneurs. The performance measure used was earnings per hour (averaged over a year) such that the entrepreneurial returns to education can be estimated and compared to those of employees. The methodological rigor applied in studies of the returns to education for employees has been our benchmark, since this rigor has been lacking until recently in the comparable entrepreneurship literature.

Consistent with previous studies pertaining to the USA, our RE-estimates indicate that the return to education is slightly higher for entrepreneurs (6.9 percent) than for employees (6.0 percent). However, when we apply IV to account for the endogenous character of schooling in an income equation, the returns to education jump to 9.9 percent for employees and 18.3 percent for entrepreneurs. The first jump is comparable to previous findings using various identification strategies in labor economics. The second jump, which is larger, leads to the remarkable result that entrepreneurs' returns to education are not slightly higher, but are an impressive 85 percent higher than the returns to education for employees. The absence of any influence from selection bias and the further robustness of this result against various alternative explanations add to the credibility of our finding.

The explanation supported by the test outcomes that utilize the locus-of-control concept is that entrepreneurship gives better opportunities to optimize one's education and subsequent returns. All together, we believe that our findings bear implications for researchers and policy makers alike.

The observation that OLS estimates are biased and that the extent of this bias differs per labor market group is an interesting starting point for researchers investigating returns to education for entrepreneurs. We suggest that further research should first produce more evidence of the relative returns to education for entrepreneurs by using modern estimation strategies and clever instruments, and then aim at understanding the differences in terms of returns to education between entrepreneurs and employees.[32] As we shall see below, such research forms the basis of several policy implications.

Before discussing policy implications, we elaborate on the remaining untested assumptions that are required to translate the estimation results into policy implications. First, we assume that the development of more entrepreneurship is economically valuable. Second, we assume that the difference between the social and private returns to the education of entrepreneurial activity is at least as large as this difference is for employees. A successful

[32]Moreover, research that differentiates educational types would be insightful, such that we can compare the returns to, for instance, vocational education for entrepreneurs and employees.

entrepreneur is, for example, more likely to influence competition in a market positively than is an employee. Moreover, entrepreneurs can bring new and innovative ideas into the market more easily than employees. Third, we assume that individuals invest in schooling at a stage in their lives at which they do not know yet, in general, whether they will become entrepreneurs or employees, or a (sequential) combination of both. As a consequence, investment in schooling is not motivated by the specific expected returns when belonging to the group of entrepreneurs, but by some (weighted) average expected returns of both employment modes. Our fourth assumption is that individuals, policy makers, bankers and other parties involved, do not have more insight in the returns to education than we as researchers have. This implies that individuals and policy makers share the common opinion that the returns to education are similar or slightly different, at most, for entrepreneurs and for employees.

Clearly, our finding that the entrepreneurial returns to education are high, and that education is therefore a key success factor for a starting enterprise, is informative for individual labor market decisions, the development of educational policies, as well as for bankers' and capital suppliers' strategies with respect to (selecting) starters. From a managerial perspective, the explanation of our result indicates that the education of employees can become more profitable in organizations that allow for more decision authority to individual employees as to how they employ their human capital.

Our finding could motivate governments to stimulate higher education for (prospective) entrepreneurs. Alternatively, policy makers could stimulate higher educated individuals to opt for an entrepreneurial career. The first route would increase the likelihood that entrepreneurs will perform better and that they will generate more benefits that will not only accrue to the entrepreneurs themselves, but, under our assumptions, also to society as a whole. The second route appeals to the fact that entrepreneurship seems not to be the favored option among highly educated individuals. Both the meta-analysis as well as the results from this chapter indicate an insignificant relation between the choice for entrepreneurship and education level. We strongly believe in the benefits of governmental programs to stimulate the awareness among college and university students of the attractiveness of entrepreneurship. Future research into the entrepreneurial returns to education in general and of specific types of education may further increase the effectiveness of such policies.

Appendix 3.A

Table 3.A-1: Returns to education while adjusting for returns to capital

	Adjusted		Unadjusted	
Variable	**Coefficient**	**(Std. Err.)**	**Coefficient**	**(Std. Err.)**
Education	0.108**	(0.014)	0.108**	(0.014)
E*Education	0.101**	(0.028)	0.094**	(0.028)
General ability	0.043**	(0.009)	0.043**	(0.010)
E*General ability	-0.038*	(0.017)	-0.033†	(0.017)
E	-0.583†	(0.341)	-0.696*	(0.338)
Male	0.239**	(0.013)	0.240**	(0.013)
E*Male	0.432**	(0.030)	0.430**	(0.030)
Married	0.053**	(0.006)	0.052**	(0.006)
E*Married	-0.041†	(0.025)	-0.046†	(0.024)
Not Healthy	-0.041*	(0.016)	-0.042**	(0.016)
E*Not Healthy	-0.013	(0.057)	-0.018	(0.057)
Live outside city	-0.098**	(0.009)	-0.096**	(0.009)
E*Live outside city	-0.002	(0.013)	-0.009	(0.013)
Live in South	-0.068**	(0.012)	-0.068**	(0.012)
E*Live in South	0.129**	(0.026)	0.125**	(0.026)
Hispanic	0.027	(0.031)	0.027	(0.032)
E*Hispanic	-0.125	(0.080)	-0.121	(0.079)
Black	-0.058**	(0.022)	-0.058*	(0.022)
E*Black	-0.002	(0.058)	-0.004	(0.058)
Locus of control	0.007*	(0.004)	0.007*	(0.004)
E*Locus of control	0.003	(0.007)	0.007	(0.007)
Mother education	-0.006	(0.004)	-0.006	(0.004)
E*Mother education	-0.022*	(0.009)	-0.018*	(0.009)
Father education	0.001	(0.003)	0.001	(0.003)
E*Father education	-0.004	(0.007)	-0.002	(0.007)
Intercept	1.550**	(0.121)	1.547**	(0.123)
N	36186		36186	
R^2 Within	0.34		0.35	
R^2 Between	0.37		0.38	
R^2 Overall	0.35		0.36	

Significance levels : † : 10% $*$: 5% $**$: 1%
E denotes Entrepreneur.

Table 3.A-2: Earnings equations excluding people working part time and as professionals respectively

	Fulltime Workers		**Without Professionals**	
Variable	**Coeff.**	**(SE.)**	**Coeff.**	**(Se.)**
Education	0.092**	(0.011)	0.098**	(0.012)
E*Education	0.107**	(0.026)	0.086**	(0.026)
General ability	0.044**	(0.007)	0.043**	(0.008)
E*General ability	-0.046**	(0.013)	-0.021	(0.013)
E	0.101	(0.505)	-0.043	(0.506)
Male	0.210**	(0.010)	0.226**	(0.011)
E*Male	0.448**	(0.025)	0.439**	(0.024)
Married	0.069**	(0.005)	0.061**	(0.005)
E*Married	-0.071**	(0.020)	-0.093**	(0.020)
Not healthy	-0.052**	(0.014)	-0.055**	(0.014)
E*Not healthy	-0.090†	(0.052)	0.024	(0.048)
Live outside city	-0.089**	(0.007)	-0.086**	(0.007)
E*Live outside city	-0.005	(0.011)	-0.014	(0.011)
Live in South	-0.062**	(0.010)	-0.056**	(0.010)
E*Live in South	0.105**	(0.023)	0.107**	(0.022)
Hispanic	0.037	(0.025)	0.041	(0.027)
E*Hispanic	-0.212**	(0.074)	-0.173*	(0.070)
Black	-0.054**	(0.018)	-0.056**	(0.019)
E*Black	-0.124*	(0.052)	-0.035	(0.049)
Locus of Control	0.008**	(0.003)	0.007*	(0.003)
E*Locus of Control	0.007	(0.006)	0.006	(0.006)
Mother education	-0.002	(0.003)	-0.004	(0.004)
E*Mother education	-0.031**	(0.009)	-0.021*	(0.009)
Father education	0.003	(0.002)	0.002	(0.003)
E*Father education	0.001	(0.006)	0.004	(0.005)
Intercept	0.785**	(0.184)	0.868**	(0.138)
N	44996		49393	
R^2 Within	0.45		0.45	
R^2 Between	0.45		0.47	
R^2 Overall	0.44		0.43	

Significance levels : † : 10% $*$: 5% $**$: 1%

E denotes Entrepreneur.

Chapter 4

Returns to education for entrepreneurs and employees: Identification by means of changes in compulsory schooling laws[1]

4.1 Introduction

During the past decades governments and researchers have become increasingly interested in the benefits of education for individual performance. While governments have focused on increasing the average education level of the population by introducing several schooling laws and training programs, labor economists have developed strategies to estimate the private rate of returns to education consistently. This task has shown to be rather complicated since individuals choose schooling levels non-randomly, unobserved characteristics may affect both schooling choices and labor market outcomes, and schooling levels are measured with measurement error. Therefore, conditional correlations as measured by OLS are potentially biased. To get an unbiased returns to education estimate, one has to use an estimation strategy that deals with the endogenous nature of the education investment decision in an earnings equation when measuring the returns to a marginal year of formal education. Examples are approaches using instrumental variables (IV), twin studies, and natural experiments. Applying these strategies usually leads to higher estimates of the returns to education for *employees* than the conditional correlations between education and employee income that are estimated by means of OLS (Ashenfelter et al., 1999).

Until recently, returns to education for *entrepreneurs* had not yet been measured by

[1]This chapter is based on Van der Sluis and Van Praag (2006).

means of these estimation strategies. The lack of attention paid to entrepreneurs in this area of research is notable since entrepreneurs contribute strongly to our economy in the form of employment and innovation (Audretsch and Keilbach, 2004; Van Stel and Storey, 2004).[2] Moreover entrepreneurs make up, on average, ten percent of the labor force in most developed countries. This is a non-negligible number.

However, one of the identification strategies used in the returns to education literature for employees, i.e. the IV-approach, has recently been applied to entrepreneurs along with employees (cf. Chapter 3, Chapter 5, Van der Sluis and Van Praag (2004), Iversen (2004) and Garcia Mainar and Montuenga Gómez (2005)). The main result from these studies is that the returns to education for entrepreneurs are significantly higher than for employees, when measured in exactly the same fashion. This result turned out not to be driven by selectivity (into entrepreneurship) or any other unobserved differences between entrepreneurs and employees (see Sections 3.4 and 5.3). However, all these studies have their limitations (discussed in Section 4.2). The quality of these studies hinges on the quality and validity of the identifying instruments used.

The aim of this chapter is to apply the instrumental variables approach using a more valid and accepted identifying instrument than the ones used before to measure the returns to education for entrepreneurs. More in particular, we are the first to identify the returns to education for entrepreneurs - along with those for entrepreneurs - by means of estimating earnings equations in which education is instrumented based on reforms in laws that set a minimum compulsory school leaving age. Hence, the variation over time and geographical regions (states) in compulsory schooling laws in the USA is used as the main instrument for education.[3] The changes in compulsory schooling laws at different points in time for different regions thus generate a natural experiment for the estimation of the returns to education (Acemoglu and Angrist, 2001; Angrist and Krueger, 1991; Oreopoulos, 2003). Exploiting differences between regions and over time in exogenous changes of compulsory school leaving ages to identify the returns to education has recently become a quite common method for the estimation of the returns to education for employees (See Acemoglu and Angrist, 2001; Angrist and Krueger, 1991; Oreopoulos, 2003, 2006). It improves on the estimation strategies that have been used so far to estimate the returns to education for entrepreneurs in comparison to employees.

First, since changes in compulsory schooling laws affect the education level of an indi-

[2]For a more detailed discussion of the importance of entrepreneurship to the economy see Chapter 1.

[3]This also enables us to assess the effects of changes in the education system on the neglected segment of the labor market, i.e. entrepreneurs. Though the measurement of such effects has been restricted to one specific labor market group, i.e. employees, the reforms might also affect, more or less strongly, the labor market performance of entrepreneurs.

vidual during childhood, the variation in the number of years of schooling caused by these changes is unaffected by future wages and hence truly exogenous. Second, not all regions changed their compulsory schooling law at the same time. The effect of a change in a compulsory schooling law can therefore be compared with a group/region where no change has taken place, i.e. differences-in-differences can be estimated. In this way unobserved factors that change over time (such as trends) and possibly influence an individual's level of schooling can be separated from an education effect that is due to a change in compulsory schooling laws.

A possible drawback of using changes in compulsory schooling laws as an instrument is that the changes themselves might be related to omitted variables that also influence incomes and schooling attainment in a certain region and time period. If so, the changes in compulsory schooling laws are not exogenous. However, there is ample empirical evidence that this is not the case and that changes in compulsory schooling laws are valid instruments (Acemoglu and Angrist, 2001; Angrist and Krueger, 1991; Oreopoulos, 2003). Our own checks of the validity of the instruments indeed support these findings. Another possible drawback of using compulsory schooling laws as instruments is that the resulting returns to education estimates are specific to those individuals who react to the treatment (instrument) and not to the population as a whole. The estimate for the treatment group might therefore be different from the effect measured when the total population would be affected. Since we would like to measure the returns to education for the total population this might be a problem. However, recent work by Oreopoulos (2006) shows that returns to education estimated with IV and compulsory schooling laws as instrument are not specific to the treated population only but representative for the total population (See Section 4.2).

The contribution of this study is twofold. First and most importantly, using widely accepted instruments we confirm the findings of Chapter 3, i.e. the returns to education are substantially higher for entrepreneurs than for employees. Secondly, we show that the returns to education for entrepreneurs are sensitive to the way entrepreneurs' earnings are measured. Including negative incomes increases the returns to education for entrepreneurs.

A limitation of the current study is that we must take for granted that selectivity is of minor importance for the explanation of the differences between entrepreneurs and employees in terms of their returns to education: The dataset does not enable testing for the presence of selectivity. However, based on the results found in the previous and the next chapter, where we find no selection effects, we assume this limitation to be not too serious.

The remainder of the chapter is organized as follows. Section 4.2 gives an overview of the empirical literature on the returns to education for entrepreneurs. Section 4.3 describes the methodology and identification strategy using compulsory schooling laws. Section 4.4 describes the data and main variables. The dataset is taken from the US census and covers each decade from 1950 till 2000, hence six cohorts. Both the estimation strategy and the compulsory schooling law data that we use are largely based on work by Oreopoulos (2006, 2003). Section 4.5 presents the estimation results. The IV-estimates indicate that the returns to education in the USA are substantially higher for entrepreneurs than for employees, respectively 26.4 and 13.2 percent. Moreover, Section 4.5 shows that the returns to education estimate for entrepreneurs is sensitive to including or excluding negative incomes. Excluding negative incomes for entrepreneurs with an unincorporated business gives an returns to education estimate for the total group of entrepreneurs that is too low. Section 4.6 summarizes the findings and concludes.

4.2 Empirical literature

4.2.1 Returns to education for entrepreneurs

In Chapter 2 we provided a detailed overview of the current state of returns to education research for entrepreneurs in the form of a meta-analysis. Its most striking conclusion is that most studies thus far had measured the returns to education for entrepreneurs while treating education as exogenous. Hence, these studies risk a biased estimate of the returns to education (Ashenfelter et al., 1999; Card, 1999).

As we briefly and incompletely discussed in Chapter 2, a few studies – performed subsequent to the execution of the meta-analysis – accounted for the endogenous nature of education when estimating the returns to education for entrepreneurs. We have encountered five of such studies:

Van der Sluis and Van Praag (2004) have been the first to use IV to estimate the returns to education for entrepreneurs. Based on USA panel data and with several family background variables as instruments for education (including parental education) they measure returns to education for entrepreneurs that are significantly higher than for employees. Moreover, the difference between the returns to education for entrepreneurs and employees are much larger upon using IV than when using the benchmark method, OLS.[4] However, using parental education as an instrument for education is not without critique. For instance, Card (1999)

[4]The IV-estimate of the returns to education for entrepreneurs is 14 percent whereas the comparable estimate for employees is 10 percent. When using OLS, these numbers reduce to seven and six percent. Hence, for entrepreneurs, the difference between the two estimates is seven percent, whereas for employees the difference is only four percent.

criticizes parental education because of its direct effect on the child's income level. Therefore, the instrument would not be valid.

In their 2005 study that was discussed in Chapter 3, Van der Sluis, Van Praag and Van Witteloostuijn extend and slightly improve on Van der Sluis and Van Praag (2004), using the same data. They omit parental education from their set of instruments but maintain it as a control variable. Thus, parental education captures otherwise unobserved transfers from parent to child that affect both education level and income. This would render the remaining family background variables that are used as identifying instruments more valid.[5] Moreover, they estimate the returns to education for entrepreneurs and employees within one equation while accounting for selectivity. Again, the results reveal a significant difference in the returns to education between employees and entrepreneurs, with IV-estimates of respectively 10 and 18 percent. In Chapter 3 we argued that this difference is due to fewer (organizational) constraints faced by entrepreneurs than employees when optimizing the profitable employment of their most important asset, i.e. human capital, obtained by education.

Hence, both studies show significantly and substantially higher returns to education for entrepreneurs than for employees. This stands in contrast to the estimates reviewed in the meta-analysis, which did not control for the endogenous nature of education, (see Chapter 2), that showed only slightly higher returns to education for entrepreneurs than for employees (in the USA). The studies in the meta-analysis however did not correct for the endogeneity of education, a comparison with the result in this study is therefore not straightforward.

Two other relevant studies have also used family background variables as instruments for education, including parental education: Iversen (2004) and Parker and Van Praag (2006). Both studies use European data and their results raise additional questions. Iversen uses Danish data and finds lower returns to education for entrepreneurs than for employees given a sample of people with low levels of education. However the opposite is found in a sample of people with higher levels of education.[6] Parker and Van Praag find an returns to education estimate of 14 percent for entrepreneurs in the Netherlands. This is higher than comparable

[5] Family background variables used in Van der Sluis, Van Praag and Van Witteloostuijn (2005) are: Magazines in the household at age fourteen, newspapers in the household at age fourteen, number of children in the household and whether or not a stepparent was present.

[6] For active members of the labor force with a lower level of education the returns to education are 6 percent for employees and 2 percent for entrepreneurs. In contrast, based on a sample of entrepreneurs and employees with a higher level of education, the returns to education are 7 percent for employees and 14 percent for entrepreneurs.

returns to education estimates for employees in the Netherlands (Plug and Levin, 1999).[7]

The fifth study, Garcia Mainar and Montuenga Gómez (2005), uses a methodology by Hausman and Taylor (1981) to estimate the returns to education for entrepreneurs and employees in Spain and Portugal. This methodology has hardly been applied and implies using the within and between variation of the exogenous variables (their means) as instruments for education.[8] A disadvantage of this scarcely applied method is that the quality and validity of the instruments are difficult to assess. Applying this method results in a higher returns to education for employees than for entrepreneurs in Spain (respectively nine and five percent). For Portugal, the returns to education are ten percent for both segments of the labor market.

Thus, only a few studies have attempted to estimate the *causal effect* of education on the earnings of entrepreneurs (and compare these to employees). Moreover, there is room for improvement in terms of the identification strategies used in these studies. Research measuring the returns to education for entrepreneurs is still in its infancy. Many more studies are required, using various identification strategies - and for IV various identifying instruments - for various countries and periods. One of the least disputed instruments for education, i.e., variation in changes in compulsory schooling laws, has not yet been applied to estimate the returns to education of entrepreneurs. We will take this approach.

4.2.2 Applications of compulsory schooling laws as instrument

Labor economists have used changes in compulsory schooling laws to estimate the returns to education for *employees* by means of IV. However, there are multiple ways in which (changes in) compulsory schooling laws have been used to estimate the returns to education.[9]

Angrist and Krueger (1991) use quarter of birth as an instrument for schooling in the USA. Exogenous variation in the length of schooling is generated by the fact that students whose birthday is just after the school enrollment date for primary school have to wait up to a year before they can start their education. However, the minimum compulsory schooling age is the same for students born before or after the enrollment date. Students born in a quarter just after the enrollment date thus reach the compulsory schooling age with fewer years of education than individuals born in the quarter of the year just before the school enrollment date. Hence, birth quarters generate exogenous variation in schooling attainment. However,

[7] Parker and Van Praag account for the effect of education on capital constraints and subsequently on the entrepreneur's performance. They estimate this indirect effect of education on performance via capital constraints to be an additional 3 to 4.6 percent.

[8] The means of time-varying variables such as age, experience and marital status are thus used as instrument.

[9] Below we will describe only a limited number of studies that have used changes in compulsory schooling laws. For more studies on this topic see Card (2001) and Pischke and Von Wachter (2005).

the low correlation between quarter of birth and years of education causes the instrument to be of low quality, possibly leading to the same sort of biases as the biases introduced by not at all acknowledging the endogenous nature of the education variable and using OLS (Bound, Jaeger and Baker, 1995; Cruz and Moreira, 2005). Using this variation Angrist and Krueger (1991) estimate the returns to education for USA employees to be between 6 and 10 percent.

A second way in which changes in compulsory schooling laws are used as instruments is by considering the effect of discontinuities over time in average individual schooling levels due to sudden changes in compulsory schooling laws on individual earnings (Harmon and Walker, 1995; Oreopoulos, 2006). The main advantage of the discontinuity approach is that no extra assumptions are required concerning the comparability of the 'treatment' and the 'control' group. Individuals who are 'treated' by the compulsory schooling law and those who are not are virtually identical apart from the fact that one group obtained their education just before the change in laws and the others just after the change. Harmon and Walker (1995) use this approach and find an returns to education estimate of around 15 percent for UK employees. Oreopoulos (2005, 2006) finds a comparable result for the UK and Northern-Ireland.

Finally, the differences in the timing of changes in compulsory schooling laws between regions are exploited as a source of exogenous variation based on which the returns to education can be estimated (Acemoglu and Angrist, 2001; Lleras-Muney, 2002; Lochner and Moretti, 2004; Oreopoulos, 2003, 2005, 2006). The impact of compulsory schooling on income can be estimated using the changes in compulsory schooling laws in one region and comparing the results to a region where no change has taken place in the same time period. In this way unobserved factors that influence individuals' schooling choices over time can be separated from an education effect due to a change in compulsory schooling laws. If two or multiple regions implement a change in compulsory schooling laws and one region does not, this 'no change region' functions as a control group. This 'difference in differences' approach (dif-in-dif) ensures that the effects measured are only due to exogenous changes in education levels. As was discussed in the introduction of this chapter, a point of attention when using this approach is that it should be reasonable to assume that region-specific economic conditions do not influence both the change of the compulsory schooling law and the future income of the individual. Several studies have performed tests to validate this assumption. Based on their evidence we conclude that this assumption is valid (Acemoglu and Angrist, 2001; Oreopoulos, 2003, 2006). Returns to education estimates found using the dif-in-dif approach range from 10 to 14 percent for USA employees (Acemoglu and Angrist, 2001; Oreopoulos, 2003).

Another point of attention when using compulsory schooling laws as instruments in an Instrumental Variables regression, which concerns all above presented approaches, is that the resulting returns to education estimates are specific to those individuals who react to the treatment (instrument) and not to the population as a whole. The estimate for the treatment group, the Local Average Treatment Effect (LATE), might therefore be different from the Average Treatment Effect (ATE), i.e. the effect measured when the total population would be treated. Since changes in compulsory schooling laws often have an effect on only a small proportion of the population it is likely that the treatment and the non-treatment groups, and therefore the true returns to education for each, are different from each other. Changes in compulsory schooling laws most often influence those individuals that would otherwise have relatively low schooling. If these individuals stay out of school because they are more credit constrained these credit constraints will cause the LATE estimates of the returns to education to be higher than the returns to education estimate obtained with OLS (Card, 2001).

Oreopoulos (2006) sets out to investigate the potential differences between ATE and LATE measurements. Oreopoulos makes use of changes in compulsory schooling laws in the UK and in Northern-Ireland. For both countries dropout rates were extremely high. Therefore, the changes in compulsory schooling laws affected half of the population studied, in contrast to the changes in compulsory schooling laws in the USA we exploit, which have an effect on a small proportion of the population. The high proportion of would-be dropouts in the countries studied by Oreopoulos (2006) creates a situation where the LATE estimates are much closer to ATE estimates than usual when the changes in compulsory schooling laws are used as instruments. Oreopoulos does not only show that the ATE and LATE effects are close to each other when a high fraction of the population belongs to the treatment group, he also compares these results with USA and Canadian LATE estimates which turn out to be equally high as the UK and Northern-Ireland estimates (between 10 and 14 percent). Oreopoulos concludes that the higher returns to education found when using IV (instead of OLS) is not likely to be due to LATE effects being estimated instead of ATE effects. ATE and LATE are fairly close to each other, even if the proportion of the population belonging to the treatment group is small.

We will use the last method (the dif-in-dif approach) for several reasons. First, we discard the use of the discontinuity approach since the compulsory schooling changes in our country of interest, the USA, had an effect on less than 10 percent of the population. This makes the discontinuity generated by the schooling laws difficult to observe and the results imprecise (Oreopoulos, 2006). Second, based on own analyses, we agree with Bound et al. (1995) that quarter of birth is too weak as an instrument for education. The most suitable approach for our purpose is the dif-in-dif approach. This approach is possible since our data, the US census, contain enough information to identify which individuals were treated by which state-specific compulsory schooling law at what point in time.

In the next section we will explain in detail how the changes in compulsory schooling laws are turned into valid instruments. We will also address the effect of the instruments on education and the possible interference of economic conditions. We will complete the section by a discussion of the estimation model.

4.3 Methodology

4.3.1 Identification by means of changes in compulsory schooling laws

Starting in the beginning of the 20th century up till the mid seventies most of the states in the USA changed their compulsory schooling laws several times. Table 4.1 shows that both increases and decreases in minimum school leaving age were observed in this period. Minimum school leaving ages ranged from none to 16, with the bulk lying between 14 and 16. For a state whose minimum school leaving age was raised or lowered this meant an increase or a decrease in the average schooling level of the individuals within that state as compared to other states. Since changes in compulsory schooling laws work at the state-level and not at the individual level these law changes are likely to be exogenous.

Table 4.1: Timing of changes in compulsory school leaving ages in the USA by state from 1915-1975

	Minimum school leaving age				
State	≤12	13	14	15	16
Alabama	1915		1921		1950
Arizona			1915		
Arkansas	1915		1917		
California	1915		1917	1921	
Colorado	1915			1965	
Connecticut			1915		1959
Delaware	1915		1921		
District of Columbia			1915		
Florida	1915		1935		1959
Georgia	1915		1921		1946
Idaho			1915;1921	1917	
Illinois			1915		1950
Indiana			1915		
Iowa	1915		1917		
Kansas			1915		
Kentucky			1915		1950
Louisiana	1915		1921		1946
Maine			1915;1939	1921;1946	
Maryland	1915	1917	1921		1959
Massachusetts			1915;1935;1959	1929	1946
Michigan			1915;1950	1921	
Minnesota			1915		
Mississippi	1915		1924		
Missouri			1915		
Montana			1915;1935		1921;1959
Nebraska			1915		
Nevada			1915		
New Hampshire			1915		
New Jersey			1915		1946
New Mexico	1915		1924		
New York			1915		1959
North Carolina	1915		1921		1946
North Dakota			1915		
Ohio				1915	1924
Oklahoma			1915;1946		1935
Oregon			1915		
Pennsylvania			1915		1946
Rhode Island			1915		1946
South Carolina	1915		1917		1946
South Dakota	1915		1924	1917	
Tennessee			1915		1950
Texas	1915;1935			1921;1946	
Utah	1917		1915;1935;1959		1946
Vermont			1921	1915	
Virginia	1915		1917		1950
Washington	1917		1915;1921		
West Virginia			1915		1946
Wisconsin			1915		1946
Wyoming	1915;1946		1935		1965

However, compulsory schooling laws, i.e. minimum school leaving ages, were not the only laws that influenced the decision to stay in or leave school. The minimum age at which individuals are allowed work permits also influences this decision. According to Acemoglu and Angrist (2001) and Lleras-Muney (2002) these work permit restrictions for young individuals are sometimes more binding than compulsory schooling laws that are thus not as compulsory as is implied. The effect of a change in compulsory schooling laws could thus be obstructed by a strict law with respect to work permits. For example in Alabama in 1949, students could obtain a work permit at age 14 while the compulsory schooling law was set at age 16. To solve this problem we follow Oreopoulos (2006) by taking the minimum between the compulsory minimum school leaving age and the minimum age at which a work permit can be obtained within each state. In this way we use the most binding "age restriction" for each individual. For the year 1949, in the state Alabama we thus use age 14 (work permit age) as the year at which the student could leave school instead of age 16 (compulsory schooling age). For convenience we will use the name (compulsory) schooling age in the rest of this chapter when we actually mean the minimum of the compulsory minimum school leaving age and the minimum age at which an individual can get a work permit.

Since we know the compulsory schooling laws and work permit laws in place in each state and each point in time we can identify which individuals were influenced by which state-specific compulsory schooling law. However, since we do not know an individual's place of residence at the time each schooling law was in place we have to use a proxy. Following Acemoglu and Angrist (2001), Lleras-Muney (2002) and Oreopoulos (2003, 2006) we use the combination of state of birth and the year an individual was aged fourteen to determine the laws in place for each individual. We thus assume that individuals go to school in the state in which they were born. Previous research indicates that this assumption has no impact on the results (Lleras-Muney, 2002). The age of fourteen was chosen because most of the changes in compulsory schooling laws were changes from a compulsory schooling age of 14 to a compulsory schooling age of 15 or 16 (See Table 4.1). Moreover earlier research indicates that the influence of the schooling laws was largest at this age (Schmidt, 1996).

In Section 4.5 we will check the validity and quality of the compulsory schooling laws as instrument. Both our own tests and previous research indicate that compulsory schooling laws are indeed valid instruments of good quality.

4.3.2 Estimation strategy

Our estimation strategy is based on the dif-in-dif approach used for the estimations on USA data in Oreopoulos (2006). This dif-in-dif estimation comes down to IV-estimation with the changes in compulsory schooling laws used as instruments.

Just as Oreopoulos (2006) we collapse the data into cell-means by survey year, state, birth year, years of education, gender and race. Since we want to differentiate entrepreneurs from employees we also collapse by employment status, as opposed to Oreopoulos (2006). Collapsing the data will reduce within state heterogeneity and computation speed substantially. Since compulsory schooling laws vary by state and not by individual this transformation does not affect the estimates.

To separate the returns to education for entrepreneurs from the returns to education for employees we further adjust the estimation strategy of Oreopoulos (2006). We interact all explanatory variables with a dummy for entrepreneurship. As we will show, this comes down to estimating two separate income equations for entrepreneurs and employees, except for differences in the error structures of the two models. The advantage of estimating one income equation for both labor market segments is that the difference in returns to education between the two labor-market groups is directly observable in the form of the interaction effect between the entrepreneurship dummy and the instrumented schooling variable. Not only do we proceed in this manner for the second stage income equation, but also for the first-stage education equation 4.1:

$$\overline{S}_{ijklm} = \delta L_{ij} + \theta E_{ijklm} + \lambda EL_{ijklm} + \alpha X_{ijklm} + \kappa EX_{ijklm} + v_i + v_j + v_k + v_l + v_m + v_{ijklm} \quad (4.1)$$

$\overline{S}_{ijklm}$ is the average level of schooling for individuals in birth cohort i, in state j in survey year k with race l and employment status m. L indicates three school-leaving-law-dummies for minimum school leaving ages of fourteen, fifteen and sixteen for birth cohort i in state j. The omitted category covers the minimum school leaving ages lower than fourteen. E_{ijklm} is a dummy variable indicating if the cluster of individuals is entrepreneur or employee. EL_{ijklm} is the interaction term of the schooling laws and the entrepreneurship dummy. To reflect economic conditions when the compulsory schooling laws were in place X_{ijklm} includes, per state, the fraction of individuals living in urban areas, the fraction of females, the fraction of black individuals, the fraction of individuals that are part of the labor force, the average age of the population, and the fraction of individuals working in manufacturing (see Oreopoulos (2006) for a motivation of this procedure). EX_{ijklm} is the interaction term of the control variables with the entrepreneurship dummy. v_i till v_m are cluster specific effects and v_{ijklm} is the error term. The predicted schooling estimates $\widehat{\overline{S}}_{ijklm}$ obtained from estimating equation (4.1) will be used to estimate the final wage equation (4.2):

$$\overline{lnY}_{ijklm} = \beta\widehat{\widehat{S}}_{ijklm} + \gamma E_{ijklm} + \eta E\widehat{\widehat{S}}_{ijklm} + \kappa X_{ijklm} + \zeta EX_{ijklm} + e_i + e_j + e_k + e_l \\ + e_m + e_{ijklm} \tag{4.2}$$

In equation (4.2) $\overline{LnY}_{ijklm}$ is the average log weekly income for all subgroups defined before. $\widehat{\widehat{S}}_{ijklm}$ is the predicted education variable obtained from estimating equation (4.1). Just as before E_{ijklm} is a dummy indicating one if the cluster of individuals is entrepreneur and zero if the cluster of individuals is employee. $E\widehat{\widehat{S}}_{ijklm}$ is the interaction term between schooling and the entrepreneurship dummy.

A potential problem when estimating equation 4.2 is that E_{ijklm}, i.e. the choice between being an entrepreneur or employee, might be endogenous. As we already discussed in the introduction of this chapter we have no means to address this issue empirically and therefore have to rely on results from the previous and next chapter that indicate that selectivity plays no role.

4.4 Data

We estimate our model on US census data for the period 1950–2000.[10] For each of the decades from 1950 till 2000 we use the publicly available 1 percent sample of the USA population.[11] Our initial sample, after the exclusion of those not in the labor-force, consists of approximately 3 million individuals between the age of 25 and 65. As we already discussed briefly in the section on the estimation strategy, we collapsed the data into cell means. Cell-mean groups are made by survey year, state[12], birth year, years of education, gender, race and employment status. This creates a comprehensive database of 350385 cells. Each cell is weighted by their population size to make sure that the final sample is representative of the USA population.[13] Before showing the descriptive statistics of this sample, we shall first discuss the most important variables in this dataset, i.e. education, employment status and income.

[10]Data were obtained from www.ipums.org, (Ruggles et al., 2004).

[11]As in Oreopoulos (2006) we extract all labor force participants from the 2000 1% sample, the 1990 1% Unweighted sample, the 1980 1% Metro (B Sample) sample, the 1970 Form 2 State sample, the 1960 General sample and the 1950 General sample, available at the US-census bureau.

[12]Individuals born in Alaska or Hawaii were excluded since these states were not part of the USA in 1950.

[13]Weights are obtained from the US Census Bureau (Ruggles et al., 2004).

4.4.1 Main variables

Education

We measure the education level of each individual in years. From 1950 till 1980 the US census directly reports years of schooling. However, for the years 1990 and 2000 education levels are administered in categories. Therefore, we follow Acemoglu and Angrist (2001) and Oreopoulos (2006) in replacing the categorical schooling variables for those years by average years of schooling. This imputation is done separately for white women, white men, black women and black men according to the imputation method by Park (1994).[14] To promote comparability across years Oreopoulos (2006) used a uniform top code of 17 across all census years. He did so because in 1950 the education level was top coded at 17 years of education. We do not follow Oreopoulos (2006) on this point because those individuals who have more than 17 years of education probably also have higher income levels. This group would then raise the average income level of the group of individuals with 17 years of education and thereby bias the returns to education estimates upwards. To circumvent this problem, we omitted all observations with more than 16 years of education from the initial sample.[15]

Employment status: Entrepreneurs and employees

We define two labor market groups – i.e., entrepreneurs and employees. We define an entrepreneur as a person who is mainly active in the labor market on a self-employed basis ("unincorporated") or who is the director-owner of an incorporated business. We exclude "hobby" entrepreneurs from the sample by using a lower boundary of 15 weeks per year worked as entrepreneur.[16] An employee is defined as a person whose main occupation is a salaried job.

[14] In his study Park (1994) uses the USA February 1990 CPS data where both the old (1950-1980) and the new question (1990-2000) on schooling attainment were included. In this way, Park was able to get at the actual years of education attained when an individual answered, for example, "some college". As it turned out there were small differences in the average years of education associated with the reported categorical levels of education for the separate groups of white women, white men, black women and black men (see Table 5 in Park (1994)). This is the reason why we imputated the average years of education separately for these four groups.

[15] This reduces our sample from 3 to 2.7 million observations. As it turns out there is almost no difference in results when using the approach by Oreopoulos (2006) or the method used here.

[16] Results do not change substantially when we remove or change this lower boundary. Farmers are also excluded from the sample because their economics are very different from other occupations.

Income

The income measure we use is log weekly income. For employees log weekly income is created by dividing before tax yearly income by weeks worked and than take logs. Unfortunately, creating an income measure for entrepreneurs is not straightforward. Four potential problems arise that might have an effect on the correct estimation of the returns to education – i.e., underreporting of income, non-reporting of income, erroneous reporting of income and negative incomes (Parker, 2004).[17] Each of these factors can possibly influence the correct measurement of income and subsequently bias the returns to education. Especially for entrepreneurs, researchers have not paid much attention to the effects the above presented problems can have on estimates of the returns to education. Exceptions are Devine (1995), Hamilton (2000), Lyssiotou et al. (2004), Fairlie (2005*b*) and the previous chapter. We will address each of the above listed problems shortly.

Recent evidence suggests that entrepreneurs underreport their income more than employees and that blue collar entrepreneurs underreport more than white collar entrepreneurs (Carson, 1984; Lyssiotou et al., 2004). Under certain circumstances underreporting could lead to biased estimates of the returns to education for employees. For instance, the returns to education will be upward biased if blue collar entrepreneurs have lower average education levels than white collar entrepreneurs. If underreporting has an effect on the estimation results we would expect that the difference in returns to education between entrepreneurs and employees would be smaller if the estimations are performed separately for the samples of blue collar entrepreneurs and employees and white collar entrepreneurs and employees. The previous chapter showed, by performing the estimations for these distinct groups separately, that underreporting does not bias the returns to education estimates. Hence we conclude that underreporting does probably not influence our results.

The second potential problem for the estimation of the returns to education is non-reporting of income. If non-reporting happens in a selective manner and, more importantly, if this selection is closely related to an individual's education level the returns to education could be biased. However , Devine (1995, Table A.1, p.249) shows that non-reporting does not vary in a structural manner with any personal characteristics. This makes a bias in the returns to education due to non-reporting unlikely.

The third potential bias for the returns to education estimate is erroneous reporting of in-

[17]For a complete description of problems with regard to measuring entrepreneurial income see Parker (2004, p.14). Erroneous reporting of income (including returns to capital) and the occurrence of negative incomes is only possible for entrepreneurs with an unincorporated business.

come (not on purpose).[18] Erroneous reporting happens, for instance, when the entrepreneur with an unincorporated business includes the returns to capital in his or her reported income (Devine, 1995; Fairlie, 2005*b*). This influences the absolute returns of entrepreneurs, but the returns to education is only influenced when erroneous reporting is non-random. For instance, if higher educated entrepreneurs with an unincorporated business earn a higher return to capital than lower educated entrepreneurs with an unincorporated business then the returns to education for the total group of entrepreneurs would be *upward biased* if the higher educated include the return to capital in their income report at least as often as the lower educated. Based on the findings in Chapter 3, which show that erroneously including the returns to capital results, if anything, in a very small *downward bias*, we conclude that the credibility of our returns to education estimates is not influenced by this problem.

A final way in which the returns to education can be biased is when negative incomes are omitted for those entrepreneurs who have an unincorporated business. Omitting negative incomes causes returns to education estimates to be *upward biased* when negative incomes are distributed evenly over all levels of education or if negative incomes are mostly observed for the higher educated. However if negative incomes are omitted and are mostly observed for lower educated entrepreneurs, returns to education estimates are *downward biased*. The main reason why negative incomes are often omitted by researchers is because the returns to education is best measured with income in logs.[19] When taking logs of income, observations with zero and negative income are automatically deleted from the sample. Another reason is that questionnaires often fail to ask if income is negative or simply report zero earnings instead of the amount of negative income. In our data, the US census, negative incomes are included. Most often, one of four approaches is taken when researchers deal with negative incomes (Devine, 1995; Parker, 2004).

The first approach simply removes observations with negative or zero income from the data. This can be done in two ways: (1) By simply deleting all observations with a negative income, (2) By taking logs of the income measure. Observations with negative or zero income then automatically drop out of the sample. This approach might lead to biased estimates of the returns to education as described above.

The second approach aims at including the observations with negative incomes in the analysis. As we explained above taking logs of the income variable automatically removes those observations with an income of zero or lower. A solution to keep those observations

[18] Making mistakes in reporting income is not a specific characteristic of entrepreneurs but of individuals in general. A discussion of this issue is beyond the scope of this chapter.

[19] Taking logs ensures that the returns to education is measured in percentage gains instead of absolute gains. When researchers take log income they make the plausible assumption that the percentage returns to education is constant across education levels.

from dropping out is to recode all negative incomes to zero and add a value of 1 to all observations. Taking logs of this new income measure now returns a value of 0 for those with a negative or zero income. Hence, these observations are kept in the sample and included in the analysis. The disadvantage of this approach is that the absolute value of income is changed for those individuals with incomes lower than zero. This is not ideal, but might be preferred above excluding those observations with negative incomes.

The third approach uses income-categories instead of absolute income. However, this implies a huge loss of information. Moreover, it makes the measurement of returns to education in percentages impossible.

The fourth and final approach is to omit taking logs and use negative incomes as they occur, i.e. include them as observations of the dependent variable together with all the positive income observations. Also for this approach measuring the percentage gain from an increase in years of education is difficult. Moreover it is not so plausible to assume that absolute income gains are constant across education levels.

None of these approaches is completely satisfactory for solving the problems described before. Given our expectations about the possible effects of the described problems on the correct measurement of the returns to education, we are particularly interested in assessing the potential effect on the returns to education of omitting negative incomes. Comparing the results obtained by following the first and the second approach renders an indication of the bias in returns to education associated with omitting negative incomes. The same is possible for the fourth approach where we can create a comparison group where negative incomes are omitted. In this way we can compare the estimation results that follow from the fourth approach where negative incomes are included and the newly created comparison group where negative incomes are excluded. We will proceed as follows.

To this end, we create four different income measures for entrepreneurs, i.e. four dependent variables. The first serves as a benchmark and is associated with approach one. For entrepreneurs with an incorporated business each respondent's total annual pre-tax wage and salary income is divided by the number of weeks worked, and then logs are taken. For unincorporated entrepreneurs we take the respondent's annual pre-tax income and divide this by the total number of weeks worked and take logs.[20] We compare this measure to our second "log"-measure, corresponding to approach two. Instead of dropping zero and negative income we recode all negative incomes for entrepreneurs with an unincorporated business to zero and add a value of 1 to all observations and then take logs. In this way negative incomes can be included somehow while using a log specification. Our third income

[20]The pre-tax income corresponds to the amount earned after subtracting expenses from gross receipts. If expenses are bigger than total receipts, the income for the unincorporated entrepreneur can be negative.

measure, which corresponds to approach four automatically includes negative incomes. The third income measure is compared with the last and fourth measure of entrepreneur income where we simply drop all negative and zero incomes.

Both the comparison of income measure one and two and income measure three and four give us an indication of the bias in the returns to education associated with omitting negative income.

4.4.2 Descriptive statistics

Table 4.2 shows the weighted means and standard deviations for both employees and entrepreneurs over the period 1950-2000. The first four rows show the different income measures we discussed above. Row three shows that earnings are slightly higher for entrepreneurs than for employees. As expected, including negative incomes in the income measure causes the average income of entrepreneurs to be slightly lower than for employees (see row 4).[21]

We furthermore observe that the average years individuals spend in school lies around 12 years. This comes down to completing high school. Except for differences similar to the differences observed in Chapter 3 the summary statistics for entrepreneurs and employees look quite alike.[22] On average 11% of the active working population was entrepreneur in the period 1950-2000.

[21] In log-terms, i.e. rows 1 and 2 in the table, the average earnings of entrepreneurs are lower relative to employees. The difference in ratios of entrepreneurs' and employees' incomes between the log and non-log specifications is due to the larger spread in the income distribution of entrepreneurs vis-à-vis employees.

[22] More males are entrepreneur than females, more blacks are employees than whites, and entrepreneurs are slightly older than employees.

Table 4.2: Descriptive statistics

Variable	Employees Mean	Employees SD	Entrepreneurs Mean	Entrepreneurs SD
Individual level				
I Log Weekly income no negatives	5.31	0.93	5.10	1.16
II Log Weekly income with negatives	5.31	0.93	4.42	1.61
III Weekly income no negatives	352.36	341.52	382.62	427.24
IV Weekly income with negatives	352.36	341.52	338.57	411.75
Education level (years)	11.78	2.82	11.69	3.02
Compulsory schooling age (years)	14.76	0.95	14.59	0.91
Female (dummy)	0.42	0.49	0.21	0.41
Black (dummy)	0.11	0.32	0.04	0.20
Survey year	1978.70	15.28	1976.75	16.72
Year when respondent was 14	1950.29	15.98	1946.16	16.97
State individual was born	30.36	14.43	30.03	14.39
Age	42.41	10.13	44.59	9.75
State level				
% urban in state	0.52	0.29	0.48	0.30
% farms in state	0.17	0.15	0.20	0.17
% black in state	0.11	0.11	0.10	0.11
% in labor-force in state	0.55	0.03	0.55	0.03
% in manufacturing in state	0.11	0.05	0.10	0.05
Age in state (years)	30.67	2.25	30.32	2.46
% entrepreneur (dummy)		11.2%		
N		346911		

Note. This table reports population-weighted statistics

4.5 Results

In this section we discuss the main estimation results. As a benchmark, we first discuss results from applying OLS. After presenting the estimates, we discuss the outcomes of testing the quality and validity of the instrument based on compulsory schooling. Finally, we will check the robustness of our results when we use different income measures, i.e. when we include or exclude negative income.

4.5.1 OLS benchmark

As a benchmark we estimate the difference in returns to education between employees and entrepreneurs by OLS. All regressions include fixed effects for birth cohort, region survey year, gender, a quartic in age, and employment status, explaining the very high R^2 statistic of 88%. As our benchmark dependent variable we use log weekly income (see Table 4.2 row 1) where negative and zero earnings are omitted. This one is preferred as a benchmark to the three other income measures for reasons of comparability with previous studies. The first column of Table 4.3 shows the OLS estimates for the combined sample of employees and entrepreneurs. Entrepreneurs seem to have slightly higher returns to education than employees, respectively 8.3 and 7.3 percent. This difference is slightly larger than usually found in comparable studies employing OLS. In absolute terms, these estimates are also slightly higher than in previous studies (see Chapter 2 and Table 3.2 in the previous chapter).

Table 4.3: Income equations; OLS and IV

	OLS		**IV**	
Variable	**Coeff.**	**(SE)**	**Coeff.**	**(SE)**
Individual level				
Education	0.073**	(0.000)	0.132**	(0.009)
E*Education	0.010**	(0.001)	0.132**	(0.036)
E	0.520	(2.335)	0.631	(2.824)
Female	-0.635**	(0.003)	-0.646**	(0.004)
E*Female	-0.187**	(0.007)	-0.228**	(0.012)
Black	-0.173**	(0.004)	-0.118**	(0.009)
E*Black	-0.168**	(0.013)	0.055	(0.057)
State level				
% urban in state	-0.041**	(0.008)	-0.047**	(0.008)
E*% urban in state	-0.045	(0.030)	-0.047	(0.032)
% farms in state	-0.291**	(0.024)	-0.144**	(0.034)
E*% farms in state	-0.546**	(0.081)	-0.174	(0.128)
% blacks in state	-0.192**	(0.048)	-0.115*	(0.049)
E*% blacks in state	0.142	(0.167)	0.111	(0.180)
% in labor-force in state	-0.403**	(0.070)	-0.035	(0.089)
E*% in labor-force in state	-1.650**	(0.218)	-1.612**	(0.254)
% in manufacturing in state	-0.441**	(0.082)	-0.293**	(0.087)
E*% in manufacturing in state	1.962**	(0.284)	1.623**	(0.324)
Age in state	-0.007**	(0.001)	-0.011**	(0.001)
E*Age	-0.014**	(0.004)	-0.027**	(0.006)
Intercept	1.263	(0.814)	1.687†	(0.896)
N	339087		339087	
R^2	0.884		0.839	

Note. E denotes entrepreneur. Regressions include dummies for year, dummies for the year the persons in the cell were 14 and dummies for the state they were born in. Control variables as described in Table 4.2 are also included. Robust standard errors have been calculated taking account of the fact that the observations are clustered into cell means. Estimations are calculated using population weights. Significance levels: † : 10% $*$: 5% $**$: 1%

4.5.2 Returns to education estimated by IV

The more interesting result is the IV-estimate shown in column two of Table 4.3. Again, all regressions include fixed effects for birth cohort, region, survey year, gender, a quartic in age, and employment status, explaining the very high R^2 statistic of 84%. A big jump in returns to education from OLS to IV-estimates is observed for both employees and entrepreneurs. Moreover the IV-estimates indicate much higher returns to education for entrepreneurs than for employees, of 26.4 and 13.2 percent respectively. Using widely accepted instruments we thus confirm the findings presented in the previous chapter where we also found a large and significant difference between both labor market groups. The difference found here is however substantially larger, the returns to education for entrepreneurs are twice as large as the returns to education for employees. In the previous chapter we found that the returns to education for entrepreneurs are approximately 1.8 times as large as for employees. Table 4.A-1 in Appendix 4.A shows that exactly the same results are obtained when estimating separate income equations for entrepreneurs and employees. In the next section the results with respect to education are discussed in more detail.

In the remainder of this section, we shall first discuss the estimated effects of the control variables included into the income equation. We shall then provide a comparison of our results with the most similar of previous studies, i.e. Oreopoulos (2003, 2006). We will then assess the quality and validity of the identifying instruments that are used to obtain these results. Finally, we will address the effects of using various measures of entrepreneurial incomes.

4.5.3 Control variables

Besides the returns to education for employees and entrepreneurs, Table 4.3 gives interesting insights in some other determinants of the incomes of entrepreneurs and employees and the differences between those groups. As was found in previous studies (see Chapter 2, Table 2.B-1) and consistent with the result of the previous chapter, females earn significantly less than males, in particular when they hold entrepreneurial positions. The gender gap found in this study is larger than the one found in the previous chapter. However the current study pertains to a much longer time period and is based on weekly instead of hourly incomes. If the gender wage gap has decreased over time and if women work fewer hours per week than men, the sample differences could very well contribute to the explanation of the larger gender wage gap found in this study, i.e. 65% for wage workers and 87% for entrepreneurs. Notably, in Chapter 3 we found that the gender wage gap for entrepreneurs was 68% and for employees 23%. Hence the relative disadvantage of females in entrepreneurial positions vis-à-vis wage employment is much smaller here. As was discussed in Chapter 3, previous studies have also found that the disadvantage of being female in terms of income is larger for entrepreneurs than for employees (See Chapter 2).

People who belong to the ethnic minority of blacks earn almost 11% lower incomes than the reference group, i.e. the white majority. This earnings differential is not significantly different for entrepreneurs than for employees. It is consistent with the result shown in Table 3.2 of the previous chapter, although the disadvantage of blacks in this study is somewhat larger. However, as we noted already, this might be due to the fact that we had a more recent sample consisting of younger people as well as an hourly income measure in the previous study. The remainder of the control variables is at the state level. They serve as controls for economic and demographic conditions.

4.5.4 Comparison with Oreopoulos[23]

Since we use the same data and estimation method as in Oreopoulos (2006) it seems logical to compare the findings from both studies, i.e. when it comes to the returns to education estimates for employees. The results of both studies are however not directly comparable. Oreopoulos's sample of employees actually includes part of our entrepreneurs' sample: It includes individuals with an incorporated business but excludes entrepreneurs who have an unincorporated business, i.e. who are self-employed. Since we find higher returns to education for entrepreneurs than for employees, we would expect the returns to education Oreopoulos finds to be higher than our returns to education estimate for employees, but lower than the returns estimated for the sample of entrepreneurs. Indeed, the returns to education of 14.2 percent that Oreopoulos (2006) finds are 1 percent higher than our returns to education estimate for employees. We obtain the same results when including individuals being employed in their own incorporated business into a sample of employees.

4.5.5 Quality and validity of the instrument

Before we continue with several robustness checks related to the measurement of income for entrepreneurs we will first discuss the quality and validity of the instruments used in the IV-estimates above. Table 4.4 gives an indication of the quality of minimum school leaving age as instruments. It presents the first stage effects of changes in compulsory minimum

[23]Oreopoulos (2003, 2006) only shows the results pertaining to the main effects, schooling in this case. Hence the effects of control variables and the statistics pertaining to the fit of the regressions cannot be compared.

school leaving age on the number of years of education people attain (see equation 4.1). The effects of raising the minimum school leaving age on educational attainment is supposed to be linear.

Panel I in Table 4.4 shows that individuals who face a minimum school leaving age of fourteen or higher have substantially more years of education than individuals who face compulsory schooling laws lower or equal to thirteen. In particular, the average number of years of schooling for the full sample is 0.245 years higher for those who faced a minimum school leaving age of 14 compared to those who experienced a lower or no compulsory school leaving age. The average school leaving age for those individuals who couldn't drop out of school before they were 15 years old, is on average, 0.273 years higher than for those facing a minimum school leaving age of 13, or younger. In the same fashion, a minimum school leaving age of 16 even increases the average number of years of schooling by almost half a year as compared to the reference group. All these effects are statistically significant and very similar to those in Oreopoulos (2003, 2006).[24] Taken together, all three schooling age dummies are significantly different from zero ($F_{(1,2988)} = 107.9, p = 0.000$) and hence of good predictive quality. We conclude that increasing the minimum school leaving age for students has a strong positive relationship with the total years of education obtained by an individual.

Panel III shows similar estimated effects if the sample is restricted to individuals with 12 years of completed schooling at maximum, i.e. those with no more than high school.[25] Hence this particular sample contains individuals who are most likely to be affected by the changes in minimum school leaving laws. The estimated effects from changes in minimum school leaving ages are similar to the effects shown in the first panel of Table 4.4. Oreopoulos (2003, 2006) also shows very similar results for this subsample as for the full sample.

Since compulsory schooling laws only affect pupils who intended to leave school at a young age, changes in the minimum school leaving age should not affect the schooling choices of individuals who opted for post-secondary education anyway. Hence, this intuition should be corroborated by insignificant effects of changes in compulsory school leaving ages on the educational level of pupils with more than 12 years of schooling. Panel II shows the evidence.[26]

[24]No state in the period included in our sample enforced a higher minimum school leaving age than 16.

[25]Again, the individual data have been compressed into cell means.

[26]Oreopoulos (2003) notes that the coefficients in Panel II could be positive as a consequence of some kind of spillover effect of the law changes: Pupils who intended to complete one or two years of schooling in addition to the minimal requirement, to avoid signalling to potential employers they are in the minimum category, might also increase their schooling attainment.

Table 4.4: First stage equations: The effect of compulsory schooling laws on education

	Variable	Coeff.	(SE)
	Estimates for the full sample (all education levels)		
	Compulsory schooling age = 14	0.245**	(0.033)
Panel I	Compulsory schooling age = 15	0.273**	(0.034)
	Compulsory schooling age = 16	0.495**	(0.038)
	N = 339087		
	R^2 = 0.266		
	Estimates for individuals with education> High School		
	Compulsory schooling age = 14	0.021	(0.019)
Panel II	Compulsory schooling age = 15	0.025	(0.017)
	Compulsory schooling age = 16	0.018	(0.020)
	N = 107596		
	R^2 = 0.019		
	Estimates for individuals with education≤ High School		
	Compulsory schooling age = 14	0.180**	(0.033)
Panel III	Compulsory schooling age = 15	0.179**	(0.037)
	Compulsory schooling age = 16	0.441**	(0.038)
	N = 231491		
	R^2 = 0.256		

Note. Regressions include dummies for year, dummies for year person was 14 and dummies for the state an individual was born in. Control variables as described in Table 4.2 are also included. Robust standard errors have been calculated taking account of the fact that the observations are clustered into cell means. Estimations are calculated using population weights. Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Next to the quality of the instrument, we check its validity. First, we have to be sure that it is the change in school leaving laws that affects the schooling attainment and that no other changes that happened simultaneously interfere. This happens for example when states that faced a change in schooling law also changed their curriculum. It might then be the effect of the curriculum on schooling attainment we are measuring and not the effect of the schooling laws. Fortunately there is ample recent evidence, pertaining to the USA, i.e. Lleras-Muney (2002), Lochner and Moretti (2004) and Oreopoulos (2003), that the school leaving laws caused the increase in schooling attainment and not something else. In addition, we perform our own checks. If other state specific policies or economic conditions changed simultaneously with the change in compulsory schooling laws, the effect of changes in school leaving ages on educational attainment or earnings might pick up the effect of this unobserved source of heterogeneity. In that case, the measured effect of changes in compulsory schooling laws on the educational attainment and earnings of the unaffected subsample of pupils with more than 12 years of schooling might be non-zero.

First, Panel II in Table 4.4 showed that the number of years of schooling is not affected by any of the law changes given the subsample of pupils with post secondary schooling. Hence this supports the absence of state specific changes that are correlated with changes in compulsory school leaving ages.[27]

Table 4.5: Reduced form equations; The effect of compulsory schooling laws on income

	Variable	**Coeff.**	**(SE)**
	Estimates for the full sample (all education levels)		
	Compulsory schooling age = 14	0.028**	(0.007)
Panel I	Compulsory schooling age = 15	0.048**	(0.008)
	Compulsory schooling age = 16	0.064**	(0.008)
	N = 339087		
	R^2 = 0.849		
	Estimates for individuals with education> 12		
	Compulsory schooling age = 14	0.001	(0.011)
Panel II	Compulsory schooling age = 15	0.022*	(0.010)
	Compulsory schooling age = 16	0.014	(0.011)
	N = 107596		
	R^2 = 0.819		
	Estimates for individuals with education≤ 12		
	Compulsory schooling age = 14	0.021**	(0.007)
Panel III	Compulsory schooling age = 15	0.032**	(0.008)
	Compulsory schooling age = 16	0.057**	(0.008)
	N = 231491		
	R^2 = 0.853		

Note. Regressions include dummies for year, dummies for year person was 14 and dummies for the state an individual was born in. Control variables as described in Table 4.2 are also included. Robust standard errors have been calculated taking account of the fact that the observations are clustered into cell means. Estimations are calculated using population weights. Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Second, we have to be sure that there are no unobserved state-specific economic conditions that influence both the change of the compulsory schooling law and the future income of the individuals affected by the change in compulsory schooling law. Evidence from the empirical literature suggests that this is usually not the case (Acemoglu and Angrist, 2001;

[27] Oreopoulos (2003, 2006) provides comparable results.

Oreopoulos, 2003). Our own check provides additional evidence of the absence of such a relationship. We regress log weekly income on the three compulsory schooling age dummies to estimate reduced form effects. For those individuals influenced by the change in schooling laws we expect to see a significant positive effect of the compulsory schooling age dummies on income. This simply shows the indirect effect of the change in schooling laws on income via education. For those individuals less likely to be influenced by the school leaving laws we expect no result. If we established a significant effect for this group, this could indicate that the law change was implemented due to, in anticipation of, or simultaneously with a change in economic conditions or policies. Table 4.5 shows that our instruments have a significant (indirect) effect on the incomes of the individuals who should be influenced by the change in school leaving laws (Panel III) but not on the incomes of those individuals that should not be influenced by the change in schooling laws (Panel II). Hence we conclude that economic conditions do not confound our results and that our instruments are valid.

4.5.6 Returns to education based on various measures of entrepreneurial incomes

In the data section we elaborated extensively on the measurement of income for entrepreneurs with an unincorporated business. We argued that omission of negative income could bias the returns to education estimate for entrepreneurs. We constructed four income measures which will allow obtaining an indication of the direction and magnitude of the bias that arises when negative incomes for entrepreneurs with an unincorporated business are omitted. Table 4.6 shows the results. The first row shows the IV-results for entrepreneurs as in Table 4.3, which serves as a benchmark. Row two is based on estimates obtained with the second definition of entrepreneurial incomes as the dependent variable, i.e. where estimates obtained with all negative incomes are recoded to zero and one dollar per week is added to the incomes of each individual and logs are taken. To deal appropriately with all the zero values that are created with this approach we use a Tobit model in the second stage. Comparing the results in row two of Table 4.6 with the results in row one shows that our original returns to education estimate was biased downward by approximately 5 percentage points.

Table 4.6: Entrepreneur returns to education estimates with various measures of the entrepreneurs' income

Variable	**Coeff.**	**(SE)**	N	R^2
Log Weekly income no negative income				
Education	0.264^{**}	(0.037)	339087	0.839
Log Weekly income negative income Tobit regression *				
Education	0.313^{**}	(0.064)	346911	n.a.
Weekly income no negative income				
Education	41.462	(16.466)	339087	0.646
Weekly income negative income				
Education	$51.175^{\dagger}$	(16.086)	346911	0.643

Note. Regressions include dummies for year, dummies for year person was 14 and dummies for the state an individual was born in. Control variables as described in Table 4.2 are also included. All variables included in Table 4.3 are also included. Robust standard errors have been calculated taking account of the fact that the observations are clustered into cell means. Estimations are calculated using population weights. Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

The two income measures presented above do not fully use the information available in the data since they either omit negative incomes or recode them to one common positive value. We therefore use the two non-log income measures, one including negative incomes and one excluding negative incomes as we defined previously. Thus we compare returns to education estimates with and without negative income observations. Row 3 and 4 of Table 4.6 show the results of using these two income measures, respectively. The coefficients reported in rows 3 and 4 represent the absolute return to one extra year of education. However before discussing these results, we first discuss two related disadvantages of the non-log estimation procedure.[28]

First, as was mentioned in Section 4.4.1, the no-log transformation implausibly assumes that the relationship between absolute earnings gains and education is linear. Second, the estimates of the effect of education on income are very imprecise, as is shown by the large standard errors in Table 4.6.

Keeping these disadvantages in mind we compare the absolute returns to education shown in row 3 with the absolute returns to education in row 4, where we use the full information on incomes. Just as in rows 1 and 2 the returns to education increase when negative incomes

[28]It is no wonder that the estimation of the determinants of absolute income levels is hindered by some serious drawbacks: One almost never observes this practice.

are included. We conclude on the basis of the results in Table 4.6 that excluding negative income biases the returns to education downwards.

4.6 Discussion and conclusion

In this chapter we estimated the returns to education for entrepreneurs in the USA and compared them to the returns to education for employees in the same country. We used changes in compulsory schooling laws as instruments for education. These have turned out to be qualified and valid instruments of an individual's number of years of education, both in various previous studies as well as in our own application. By using these instruments, we seek to confirm and improve on earlier studies, which used less accepted instruments to estimate the returns to education for entrepreneurs. In particular, we are aiming at validating the results found in the previous chapter.

Estimating the effects of changes in compulsory schooling laws actually leaves room for two interpretations that are both policy relevant. The first interpretation is the most direct one, increasing the minimum age at which pupils are permitted to leave school has a positive effect on individuals' earnings. More in particular, implementing minimum school leaving ages of 14, 15 and 16 years respectively (where a lower or no minimum school leaving age is the benchmark), increases the average number of years of schooling in a state by 0.25, 0.27 and 0.50 years respectively. These increases are significant. Moreover, the application and enforcement of these compulsory school leaving ages have, on average, a significantly positive direct effect on incomes of three, five and six percent, respectively.

Second, by using compulsory minimum school leaving ages as instruments for individuals' education levels in the estimation of incomes, one can retrieve estimates of the returns to education. This interpretation is implied by the main research question of the current study. We find evidence of substantially higher returns to education for entrepreneurs than for employees, i.e. 26.4 percent and 13.2 percent respectively. These estimates are much higher than the OLS benchmark estimates of 8.3 percent for entrepreneurs and 7.3 percent for employees. The returns to education for entrepreneurs -and thereby the difference with the returns to education for employees- become even larger if we allow the incomes of entrepreneurs with unincorporated businesses to be negative. Including negative incomes conforms to reality, though not to common practice in previous studies. Albeit imperfectly measured, upon including negative incomes, the returns to education for entrepreneurs become even five percentage points larger, i.e. they amount to 31%. This finding is corroborated by the results from using another imperfect income measure that includes negative incomes.

4.6.1 Comparisons to previous studies

It is useful to compare our results to Chapter 4 and Oreopoulos (2006). The first comparison adds to the credibility of the results presented in Chapter 4, whereas the second contributes to the credibility of this study. Using widely accepted instruments for education we confirm the results from the previous chapter that the returns to education are higher for entrepreneurs than for employees. However, in this study we find substantially higher estimates for both segments of the labor market than in Chapter 3 where we measured the returns to education in the USA for both labor market groups by using family background variables as instruments for education.

Unfortunately it is not possible to completely attribute the different findings of the two chapters to the different instruments used. Various differences between the samples used in the two studies, both based on USA data, might also explain the differences in the estimates. The average age of the individuals in the NLSY sample that was used in the previous chapter is lower than the average age of the individuals in the US Census, the current study's sample. Moreover, the time period the NLSY covers is more recent and shorter than the one studied in this sample. Finally, the definitions of incomes differ across the two datasets. In the previous chapter, the dependent variable in all income equations was hourly earnings averaged over a year, whereas in the current sample, the main indicator of income is earnings per week, averaged over a year.

As a matter of fact, the US census is not rich enough to allow a within sample comparison of both instruments. Neither can we define hourly incomes (hours are not available) or use only the most recent decennial census (too few changes in compulsory school leaving ages would remain). The only adjustment of the sample that would make the two studies more similar would be to limit the sample of the current study to younger people only. In general, we cannot address the source of the differences of the findings of this and the previous chapter.

Comparing the results of the (employee part) of the current study to the study that is most similar in design and sample used, i.e. the USA results contained in Oreopoulos (2006), increases the confidence we have in the results found in this study. We find the almost exact same effects of compulsory schooling laws on both educational attainment and earnings. Moreover, the estimated returns to schooling are similar to the estimate provided by Oreopoulos. In addition, Oreopoulos' results from testing the usefulness of the instruments are consistent with ours and support the quality and validity of the instruments we use. Finally, the most novel aspect of Oreopoulos (2006) is that he creates an opportunity to assess the differences between Local Average Treatment Effects (LATE) and general Average Treatment Effects (ATE). He concludes that the usual high estimates of the returns to education based on IV-approaches are not due to peculiarities of the -often small- sub-sample that is treated by the interventions used as instruments. Hence, LATE can be treated as ATE and therefore be generalized.

4.6.2 Limitations

Before turning to possible policy implications of the current study, we wish to highlight some of its limitations (and thereby possibly provide suggestions for future research). First, we must perforce assume that selectivity (into entrepreneurial positions) does not confound our results. Although the evidence suggests that the probability of biased results due to selectivity is limited, see both the previous and the next chapter, a test within the current framework would be useful but impossible to design, due to a lack of possible instruments of employment status. A second concern is the relatively large standard errors of the coefficient pertaining to the estimated bonus effect of education on earnings for entrepreneurs. Hence, the 95% confidence interval of this bonus, of which the point estimate is 13.2% lies in the range of 6.6 and 19.9 percent. Our third concern is related to the sensitivity of the results with respect to alternative definitions of entrepreneurs' incomes. We deduced from the literature that the occurrence of negative incomes for entrepreneurs, as opposed to misreported or underreported incomes, has the highest probability of affecting the estimates of the returns to education. This was indeed shown to be the case: The returns to education are underestimated when negative incomes for entrepreneurs are omitted. This is an interesting finding for researchers who want to use entrepreneurial incomes as a dependent or independent variable in their analysis. We have not yet provided a definite solution to deal with negative incomes in future research. None of the income measures we used is optimal. Temporarily, using the log specification where negative income is recoded to a small positive value seems to be the best alternative. It is however quite likely that a better alternative exists that has the properties of the log-transformation but keeps the information contained in the negative income observations.

4.6.3 Policy implications

The positive effects of changing compulsory schooling laws on school attainment and incomes suggest that (at least a proportion of the) USA pupils choose below optimal schooling levels. Viewing education as an investment decision would plea for the absence of any compulsory schooling laws, since such laws restrict the choices available to optimizing individuals.

Oreopoulos (2003) digs further into the policy relevant consequences of the direct positive effect of increasing minimum school leaving ages. The policy measures, in terms of changing compulsory school leaving ages, obviously depend on the underlying causes of dropping out, which are in contrast to the results that suggest that staying in school would lead to substantial income gains, on average.

From the viewpoint of schooling decisions as part of an investment plan, the results imply school attendance costs would have to be extremely high to explain dropout decisions. If true, addressing these costs could mitigate dropout behavior. However, this is unlikely to be the true cause of dropping out because only 'sky high' attendance costs could offset the predicted gains from staying in school. However, as Oreopoulos notes:

"There is no a priori reason to prefer an investment model of school attainment over the many other possibilities - cultural or peer pressures may predominate adolescent decision-making; youths may underestimate the rewards from staying in school; they may temporarily ignore longer-term consequences of their decisions; or parents may not be able to recover costs off of their children's gains. Each explanation carries vastly different policy implications." Oreopoulos, 2003, p.25.

Hence, more research is required to distil the appropriate policy measures from the finding that changes in compulsory school leaving ages matter for outcomes in terms of educational attainment and incomes.

The study leaves room for a second strand of policy implications based on the high measured returns to education that are even much higher for entrepreneurs. These implications were already presented at the end of the previous chapter, based on qualitatively the same findings: Since entrepreneurs have high private benefits from education, and seem not be aware of that -based on the insignificant effects we found in Chapter 2 of education on entrepreneurship selection-, an information campaign could stimulate more highly educated people to prefer entrepreneurship over wage employment. For instance, entrepreneurship programs available to or even obligatory for broad groups of students could provide such a stimulus. Moreover, a different campaign targeted at prospective entrepreneurs could stimulate them to opt for higher levels of education in their own interest.

However before we know that (i) Stimulating individuals with higher levels of education to become entrepreneurs and (ii) Stimulating entrepreneurs to attain higher levels of education, is beneficial for society we must assume that the difference between the social and private benefits of entrepreneurial activity is at least as large as this difference is for employees. We think this is a very plausible assumption. A successful entrepreneur is, for example, more likely to influence competition in a market positively than is an employee. Moreover the evidence, so far, pertains to the USA only. Hence we cannot conclude yet that the same policy measures would be effective in other countries, i.e., Europe. It is up to future research to investigate these questions.

Appendix 4.A

Table 4.A-1: IV income equations: Separate equations for employees and entrepreneurs

Variable	Employees Coeff.	(SE)	Entrepreneurs Coeff.	(SE)
Education	0.132**	(0.009)	0.264**	(0.037)
Female	-0.646**	(0.004)	-0.874**	(0.012)
Black	-0.118**	(0.009)	-0.063	(0.058)
% urban in state	-0.047**	(0.008)	-0.094**	(0.032)
% farms in state	-0.144**	(0.034)	-0.318*	(0.130)
% blacks in state	-0.115*	(0.049)	-0.004	(0.182)
% in labor-force in state	-0.035	(0.089)	-1.647**	(0.254)
% in manufacturing in state	-0.293**	(0.087)	1.331**	(0.328)
Age in state	-0.011**	(0.001)	-0.038**	(0.006)
Intercept	1.652†	(0.896)	2.522	(2.900)
N	251467		87620	
R^2	0.894		0.542	

Note. Regressions include dummies for year, dummies for year person was 14 and dummies for the state an individual was born in. Control variables as described in Table 4.2 are also included. Robust standard errors have been calculated taking account of the fact that the observations are clustered into cell means. Estimations are calculated using population weights. Significance levels: † : 10% $*$: 5% $**$: 1%

Chapter 5

Returns to intelligence: Entrepreneurs versus employees[1]

5.1 Introduction

Previous empirical analyses and the previous chapters have demonstrated that human capital is an important determinant of entrepreneur performance, cf Wagner (2003), Silva (2007), Lazear (2005) and Van Praag (2006). Both education and specific sorts of labor market experience have significant effects. For instance, in Chapters 3 and 4 we demonstrated that the returns to education are higher for entrepreneurs than for employees.

In addition to education and experience, intelligence might also affect the performance of entrepreneurs. Little is known yet about its importance for the performance of entrepreneurs. More in particular, no empirical research has yet taken place to measure the labor market value of (various areas of) intelligence for entrepreneurs and compared them to the value for employees (Schmidt and Hunter, 1992). Such a comparison is the aim of this study.

Comparing the returns to human capital for entrepreneurs and employees requires a single framework for measuring returns, as well as a measure of labor market performance that is identical for entrepreneurs and employees. In microeconomics, labor market performance has mostly been measured in terms of income levels individuals generate from employment. Within the Mincer framework (Mincer, 1970) education and experience are the main determinants of income. Intelligence has mostly been considered a source of unobserved heterogeneity. We shall extend the application of the Mincer framework to a comparison of the returns to measured intelligence between employees and entrepreneurs.

Using measures of various types of intelligence in the NLSY data we seek an answer to

[1]This chapter is based on Hartog, Van der Sluis and Van Praag (2006).

three questions for both entrepreneurs and employees: (1) To what extent does an individual's general intelligence level affect income? (2) Do different areas of intelligence (such as math, clerical, language, technical and social intelligence) affect incomes differently?, and (3) To what extent does the balance in an individual's scores in these areas of intelligence affect an individual's income?

The third analysis will further our understanding of Lazear's Jack-of-all-Trades (JAT) theory, which claims that entrepreneurs -in contrast to employees- need a wide variety of skills. We extend Lazear's theory and its empirical testing in two ways. First, by using the balance in an individual's scores on various areas of intelligence, instead of the breadth in educational or labor market activities as a measure of being a JAT, we measure a more exogenous effect. Second, empirical testing of Lazear's JAT-theory so far, has been limited to the effect of being a JAT on the probability of becoming an entrepreneur. We extend the analysis by measuring the effect of JAT on the *performance* of entrepreneurs. Moreover, we account for the possible endogeneity of the decision to become an entrepreneur as well as the endogenous nature of education in income equations.

We find that an individual's level of *general* intelligence increases both entrepreneurs' and employees' incomes to the same extent. Entrepreneurs benefit differently from specific *areas* of intelligence than employees. The earning power of both groups is affected by social and technical intelligence. Both types of intelligence render stronger returns for entrepreneurs than for employees. Mathematical intelligence is also beneficial for both groups, though only for the upper end of the intelligence distribution. Only employees benefit from clerical intelligence. Verbal intelligence does not have a strongly significant effect on earnings for either group. The balance in the various areas of intelligence is indeed a strong determinant of earning power for entrepreneurs only, and in the expected direction. This finding supports and extends Lazear's JAT theory: Entrepreneurs benefit from a balanced intelligence set.

The chapter proceeds as follows. In Section 5.2 we describe the literature on intelligence and earnings. From this literature we derive the main hypotheses about the effects of various types of intelligence on entrepreneur and employee earnings. Section 5.3 presents a description of the data and methodology we use. In Section 5.4 the results are presented. Finally, conclusions and implications are provided in Section 5.5.

5.2 Literature and main hypotheses

5.2.1 Intelligence

In 1869 the British psychologist Francis Galton (1822-1911) published his work 'Hereditary Genius' in which he concluded that intelligence is indeed hereditary.[2] In 1890 Alfred Marshall (1890;1930, p.207) was the first to link performance of entrepreneurs to intelligence. Inspired by Galton, he postulates that successful entrepreneurship requires a high level of general ability as well as various specialized forms of ability.[3]

The measurement of intelligence goes back to the beginning of the 20th century when Alfred Binet was commissioned by the French government to develop an instrument to measure differences in child intelligence. His intelligence test, which was the first paper and pencil test to measure intelligence, is still a central element of many of the current tests, e.g., SAT, Stanford-Binet Test, Wechsler adult intelligence scale and ASVAB (Binet and Simon, 1911). Binet assumed that intelligence was essentially unitary, i.e. intelligence is one overriding quality that helps an individual deal with the environment. This assumption has been challenged in a still ongoing discussion.

On the basis of correlation studies Spearman (1904) concluded that intelligence can be divided into a general factor (g) which is associated with all activities, and specific factors (s_n) of which one or more additionally influence an activity. Thorndike (1904) denied the fact that there is such a thing as "g" and acknowledged the specific factors only. Using factor analysis Thurstone (1941) positiones himself between Spearman and Thorndike. According to Thurstone, intelligence is not just "g" plus "s" but a primary group of factors (6 or 7) and one general factor "g". Recently Carrol (1993) came up with the idea of different hierarchies, with one general factor on top, several major group factors in the middle and some 50 specific factors at the bottom. A year later Hernstein and Murray (1994) promoted again "g" as the most important factor. The discussion on the nature of intelligence has not yet been closed. We take an eclectic position and consider three measurable aspects; general ability, specific abilities and balance in abilities. Attention for the balance in abilities is based on Lazear (2005), the so called Jack-of-all-trades theory (JAT).

[2]He was inspired by his cousin Charles Darwin's publication, 'Origin of Species' (Darwin, 1859). We use the terms intelligence and general ability as synonyms (English and English, 1958).

[3]For an overview of several important classical thoughts on entrepreneurship and the effect of ability on performance, see Van Praag (1999). For a more general discussion and a survey on the relationship between human capital and individual abilities see Hartog (2001).

5.2.2 The returns to general ability

Since the introduction of the Mincer earnings equation, general ability has been recognized as an important form of unobserved heterogeneity that influences the correct estimation of the returns to education and experience (Mincer, 1970; Card, 1999). Many researchers have included measures of intelligence into Mincerian earnings equations. Results from these studies show a significantly positive effect of general ability on employee incomes (for example, Hartog, 1980, 1992; Blackburn and Neumark, 1993; Neal and Johnson, 1996; Goldsmith, Veum and Darity Jr., 1997; Plug and Levin, 1999; Ferris, Hochwachter and Witt, 2001; Zax and Rees, 2002; Zetterberg, 2005). Less evidence has been obtained about the effect of general ability on entrepreneurial incomes, though the scarce results indicate a positive effect (De Wit and Van Winden, 1989; Van Praag and Cramer, 2001). Even less is known about the relative benefits of intelligence for entrepreneurs and employees.

According to psychologists the effect of intelligence on job performance is mainly caused by the impact of intelligence on the acquisition of job knowledge. That is, more intelligent individuals acquire relevant job knowledge more rapidly and in larger quantities and it is this knowledge that causes better job performance (Schmidt, Hunter and Outerbridge, 1986; Ree and Earles, 1992; Schmidt and Hunter, 1992; Guion, 1998; Schmidt and Hunter, 1998, 2004). Moreover, general ability is believed to enhance the ability to deal with complex situations (Schmidt and Hunter, 2004). Since both employees' and entrepreneurs' tasks obviously require job knowledge and dealing with complex situations, a positive return to general ability is expected for both labor market groups. But it is hard to speculate on the relative benefits for both groups. Perhaps one may conjecture that an entrepreneur inevitably faces complex situations, and an employee will face them only if assigned to that particular sort of job. Then, the returns to general ability might be more certain for entrepreneurs than for employees.

5.2.3 The returns to specific abilities

As was pointed out, Marshall was the first economist who stressed the importance of various specialized forms of ability for achieving successful entrepreneurship. Especially the management literature has pursued the idea that certain specific abilities[4] (e.g. social ability) are relatively important for entrepreneurs as compared to employees (Baron, 2000; Baron and Markman, 2003; Shane, 2003; Hmieleski and Ensley, 2004). However, little empirical evidence has supported such ideas so far. One reason for this is probably the availability of

[4] Another commonly used term for specific ability is 'specific aptitudes' or just aptitudes.

data. Most representative datasets do not include any measures of ability or, if so, only one. Moreover, there are few datasets that include both entrepreneurs and employees as well as comparable measures of their labor market performance.

Our data, the NLSY, allows us to distinguish and analyze five types of specific ability. We distinguish: (i) Verbal ability, which consists of the knowledge to understand and process written material; (ii) Mathematical ability, which is the knowledge to perform mathematical calculations and logical thinking; (iii) Technical ability, which is the ability to understand physical and mechanical principles; (iv) Clerical ability, which is the ability to process information quickly; and, finally (v) Social ability, which is the ability to form social contacts. All of these measures (loosely) correspond to specific abilities used and analyzed in earlier studies. However, the joint analysis of these five specific abilities and the evaluation of their "relative" value has not been carried out before.

The relationship between labor market performance and verbal ability has received substantial research attention. However, the results are ambiguous. Verbal ability is reported to have no value (Paglin and Rufolo, 1990; Dougherty, 2000), a negative impact (Bishop, 1991) and sometimes a positive effect on labor market outcomes (Hause, 1972). An explanation for these mixed results might be the presence of non-linearities in the returns to verbal ability (Dougherty, 2000; McIntosh and Vignoles, 2001). The idea is that performance requires only a small/basic level of verbal ability such that returns are positive only at the lowest levels of the verbal ability distribution, whereas at higher levels there are no extra effects or even negative effects. No distinction has been made in this respect in the existing literature between entrepreneurs and employees. We expect to find a small positive non-linear effect of verbal ability on the performance of both employees and entrepreneurs and we do not have any expectations about the difference in returns to verbal ability across entrepreneurs and employees.

The relationship between labor market performance and mathematical ability has received most attention from researchers. The returns to math ability are mostly found to be significantly positive (Taubman and Wales, 1974; Willis and Rosen, 1979; Paglin and Rufolo, 1990; Murnane, Willett and Levy, 1995; McIntosh and Vignoles, 2001). A minority of (some older) studies reports an insignificant or even negative return to math ability (Hause, 1972; Bishop, 1991). Again, no distinction has been made in this respect between entrepreneurs and employees. Therefore, and based on the majority of the empirical findings, we expect math ability to have a positive effect on labor market outcomes for both employees and entrepreneurs. Again, we have no upfront expectation about the difference in returns across entrepreneurs and employees.

The third ability type that we study is technical ability, i.e. the ability to understand physical and mechanical principles. Evidence of the relation of technical ability to performance comes from a study by Blackburn and Neumark (1993).[5] They find that technical ability has a significant and positive impact on employee performance. For entrepreneurs there is no empirical evidence. We expect that technical ability is even more valuable for entrepreneurs than for employees, since successful entrepreneurship often depends on process or product innovation.

The fourth type of specific ability we explore is clerical ability, i.e. the ability to process information quickly. This type of ability has almost entirely been neglected in studies regarding the labor market effects of intelligence. The exceptions are Bishop (1991) and Murnane, Willett, Braatz and Duhaldeborde (2001) who used the NLSY (for the periods 1979 to 1986 and 1979 to 2000, respectively) to find that clerical ability enhances employees' performance. We include clerical ability in our analysis to explore whether entrepreneurs and employees have a different need for clerical ability. We expect however that clerical ability is less beneficial for entrepreneurs than for employees since clerical ability is often associated with administrative tasks that can easily be outsourced by entrepreneurs.

The last type of ability we consider is social intelligence, which is receiving increasing attention in research.[6] Various studies have shown that the ability to disentangle patterns of social relationships and being able to deal with social relationships accordingly have a positive influence on performance (Baron, 2000; Wong and Law, 2002; Baron and Markman, 2003). Baron and Markman (2003) suggest that social ability is more important for the performance of entrepreneurs than of employees.[7] They argue that social ability is relatively more valuable for entrepreneurs because entrepreneurs must interact with many different persons inside and outside the firm in environments that are often unstructured and uncertain. Baron and Markman indeed find that social perception, adaptability and expressiveness are important determinants of entrepreneurial performance.[8] However they cannot test whether social ability is indeed more important for entrepreneurs than for employees as their sample only

[5]Their measure of technical ability is not identical to the one used in the current study. Our measure of technical ability is included in their measure of 'non-academic ability'. The latter corresponds to our technical ability measure and also includes what we call clerical ability (discussed next), auto and shop knowledge and electronics knowledge.

[6]A concept related to social ability which is also receiving increasing attention is "self-esteem". Self-esteem probably influences an individual's social ability, which on its turn could influence performance. In this study we do not take any indirect effects of our ability measures into account. For a study on the relationship between self-esteem, general ability and employee performance see Zetterberg (2005).

[7]Our notion of social ability corresponds to the term 'social competence' in Baron and Markman (2003).

[8]Moreover, the combination of general and social ability has proven to be of importance for labor market performance in general. Ferris et al. (2001) assess the interactive effect of social ability and general ability and find a positive interaction between the two.

includes entrepreneurs. We expect social ability to have a positive and significant effect on labor market outcomes for both employees and entrepreneurs. Based on the suggestion by Baron and Markman, we expect the effect to be stronger for entrepreneurs than for employees.

5.2.4 The returns to a balanced set of abilities

A recent stream of papers, initiated by Lazear (2002, 2005) and further built on by Wagner (2003), Lee (2005) and Silva (2007), pays attention to the combination of different competencies instead of merely the level of these competencies. Lazear's theory (2005) poses that individuals with a broad set of balanced competencies across different fields, i.e. "Jacks-of-all-trades" (JAT), are more apt for entrepreneurship than those who have a very unbalanced set of competencies, i.e. specialists.[9]

According to Lazear, the JAT characteristic determines the *selection* into entrepreneurship since being a JAT is more valuable for entrepreneurs than for employees. An employee may work in a specific job needing a specific skill, but entrepreneurs are assumed to need all competencies themselves and they are as strong as the level of their weakest skill.[10] Employees are the pawns in the division of labor, employers organize this division.[11] Lazear indeed finds that JATs have a higher probability of becoming an entrepreneur. Lazear's definition of a JAT is based on an individual's schooling curriculum and the number of different job roles individuals have had. Hence, people can either have chosen to develop as a JAT, for instance because they wanted to become entrepreneurs, or they may have been born as such and have therefore pursued a general schooling curriculum and assumed many different job roles.[12]

We modify Lazear's way of testing his theory in three dimensions. First, we use an alternative measure of JAT. We use the variation in an individual's scores across the five

[9] Although the JAT theory has come up recently, the idea that a combination of various skills is important for entrepreneurs is not entirely new. As early as 1959, Edith Penrose mentioned that *" 'Enterprise' is by no means a homogeneous characteristic, and that the 'quality' of enterprise, that is to say, the particular types of entrepreneurial services available to a firm, is of strategic importance in determining its growth"* (Penrose, 1959, p.35 fifth impression 1972). Moreover Hartog (1980) tests whether different abilities are separable in an earnings equation. In a sense, Hartog's test comes down to (an inverse test) of the JAT theory for employees: He tests to what extent the value of specific abilities depends on the value of other specific abilities in particular professions. In case of a dependence, the function can be classified as a JAT function. It turns out that for employees being a specialist is worth more than being a generalist.

[10] Team entrepreneurship is an exception.

[11] For a vivid illustration of this division of labor in industrial processes, see the quotes and references in Hartog (1992, Section 2.3).

[12] See Silva (2007) for a discussion on the issue of the source of being a JAT, i.e. 'nature vs. nature'.

types of specific abilities we distinguish. Lazear used two measures of JAT: (i) The relative variety of fields from which MBA courses have been taken by Stanford graduates; and (ii) The number of job roles assumed by these graduates in their professional lives. Our intelligence measures have a clear advantage. We do not think they are influenced by the anticipated decision to become an entrepreneur or by the anticipated relative earnings as such, whereas an individual's curriculum or job role choice might very well be influenced by this anticipation, leading to endogeneity and unclear causality. We believe that our JAT measure -measured at a relatively young age- only measures the effect of being 'endowed'[13] as a JAT and does not mix this effect with possible effects of investments in schooling and labor market roles to become a JAT.

Second, we do not evaluate the effect of being a JAT on the *selection* into entrepreneurship, but rather on the *performance* of entrepreneurs. The JAT theory shows that JATs have a comparative advantage over others to become an entrepreneur. A relevant way to test the existence of a comparative advantage would be to measure whether being a JAT has a more positive effect on the performance of entrepreneurs than of employees.

Third, most of the studies that examine the effect of being a JAT on entrepreneurship focus on only a particular part of the schooling distribution such as university -or even MBA-graduates (Lazear, 2005; Silva, 2007). The NLSY data used in this study allow studying the relationship between being a JAT and performance across the entire schooling distribution.

5.3 Data

5.3.1 Data description

We test the hypotheses developed in the previous section using a sample drawn from the 1979 National Longitudinal Survey of Youth (NLSY) in the USA.[14] The nationally representative part of the NLSY consists of 6,111 individuals aged between 14 and 22 years in 1979.[15] They have been interviewed annually up to 1994, and since then on a bi-annual basis. Within each observed year, our sample selection includes all persons who are entrepreneurs or employees (defined below), while excluding students and people who are unemployed or otherwise not working. The resulting sample includes, on average and per annum, 2,490 entrepreneurs/employees. On average, each individual is included in the sample in 10.3

[13] In the next section we will explain in what way our ability measures can be seen as an endowment.

[14] This is the same dataset as was used in Chapter 3.

[15] The original NLSY sample consists of 12,686 individuals. From this sample we excluded the supplementary military sample and the supplementary minority sample.

waves. Before turning to the descriptive statistics, we first define the variables used in the empirical analyses.

Occupational status

An entrepreneur is defined as a person whose main occupation in the labor market is on a self-employed basis or who is the owner-director of an incorporated business. Farmers are excluded from the sample.[16] Furthermore, we exclude "hobby" entrepreneurs from the sample by using a lower boundary of 300 hours per year worked as an entrepreneur.[17] An employee is defined as a person whose main occupation is a salaried job.

An important feature of the sample is that it includes both entrepreneurs and employees, and it records individuals' switches between these states over time. All entrepreneurship spells, except the very short ones (less than half a year), are recorded. Therefore, the subsample of entrepreneurs does not suffer from survival bias – i.e., the returns to ability will not pertain to surviving entrepreneurs only.

Hourly income

We need a variable that measures labor market performance in a comparable manner for entrepreneurs and employees. Obviously, the only performance measure that is defined for both labor market segments is income. Specific variables such as supervisory ratings or firm size cannot be used.

Therefore, we use gross hourly income as a performance measure for both entrepreneurs and employees. Hourly income is constructed as the average earnings (for entrepreneurs, the average income withdrawn from their firm) over a year divided by the number of hours worked in that year.[18]

General ability, specific ability and balance in abilities

The intelligence test we use is the Armed Service Vocational Aptitude Battery (ASVAB). The ASVAB is a test developed by the USA Department of Defense in the 1960s for the

[16] Their economics are very different from other occupations. From 1979 to 2000, we left out 299 farmers. Most studies drop farmers or study them separately.

[17] This number of hours has been chosen arbitrarily. We tested whether the results we present in Section 5.4 are sensitive to increasing this number. We find that the results do not change substantially.

[18] See Fairlie (2005*b*) for an evaluation of the income variable in the NLSY for entrepreneurship research. In Chapter 3 we checked the presence and effect of several of the potential problems of the income measure for entrepreneurs that are mentioned by Fairlie (2005*b*). The results in Chapter 3 which are based on the same data and income measure as the current chapter, indicate that the use of entrepreneur income is valid for a comparison of the effect of specific covariates on the income of entrepreneurs and employees.

purpose of recruiting military personnel and is still used for this purpose. In 1980 the ASVAB intelligence test was added to the NLSY questionnaire with the purpose of generating a benchmark intelligence measure for the military that is representative of the total USA population and not only of the military. Like most other intelligence tests, the ASVAB is built up of several components. It correlates strongly with other intelligence tests that are frequently used, such as the Otis-Lennon Mental Ability test and the Lorge-Thorndike Intelligence Test (Frey and Detterman, 2003).

The ASVAB as recorded in the NLSY can be distinguished from other measures of intelligence in three respects. First, the ASVAB is included in a large, rich and representative panel dataset (NLSY). Second, unlike most intelligence tests, the ASVAB is administered at a relatively young age of the respondents, i.e. between 14 and 23 years old, such that the test outcomes are (almost) not affected by (future) labor market choices.[19] Hence, this favors a causal interpretation of the effect of intelligence on labor market outcomes. Third, the ASVAB consists of a broad array of specific abilities. It includes ten different components: (1) General science, (2) Arithmetic reasoning, (3) Word knowledge, (4) Paragraph comprehension, (5) Numerical operations, (6) Coding speed, (7) Auto and shop information, (8) Mathematics knowledge, (9) Mechanical comprehension, and (10) Electronic information. The broad array of specific abilities allows us to investigate the impact on performance of various specific abilities (relative to each other) as well as of the balance in the scores on these specific abilities.

The ASVAB scores have been measured when respondents were of different ages. As age is likely to be correlated with the scores, the measures are incomparable across individuals of different ages. There is also evidence that an individual's education level at the time of the ability test influences the scores on the different specific abilities (Murnane et al., 1995; Hansen, Heckman and Mullen, 2004). To remove the age and education effects from the ASVAB component scores we apply a simple but approved method (Blackburn and Neumark, 1993; Fairlie, 2005*a*). We regress each normalized test score on a set of age and education dummies.[20] The individuals' residuals are used as the corrected test scores. In this way the ability scores, that are scaled from zero to ten, are purged from any education or age effects. This method does not clear the ten ASVAB components from all factors that could potentially affect the measurement of someone's ability endowment. Differences in, for example, child rearing might still be picked up by the ASVAB components. These factors are however less likely to be driven by an investment decision, such as education is. We

[19]For a panel dataset which also includes intelligence measured at a young age see Zetterberg (2005).

[20]Each age dummy is equal to one for a specific age in 1980 and zero otherwise. Each education dummy is equal to one for a specific number of years of schooling in 1980 and zero otherwise.

therefore think that these other effects will not bias our results.

Unfortunately, the ASVAB does not include a measure of social ability. However, the NLSY includes a measure of how outgoing individuals were as a child. We will use a rescaled version of this measure (from 0 to 10) such that the scale is comparable to the scale used for the other abilities as a proxy for social ability.[21] The other four specific abilities are constructed on the basis of the 10 ASVAB components, such that the five measures of specific ability (including social ability) are as orthogonal to each other as possible. To determine orthogonality we computed the correlation matrix between the ten components and social ability. From this correlation matrix we excluded those components that correlated highly (correlation $> .6$) with more than one other component.[22] This results in the following five fairly independent specific abilities (1.) Language ability, corresponding to 'Paragraph comprehension' (2.) Mathematical ability, corresponding to 'Mathematics knowledge' (3.) Technical ability, corresponding to 'Mechanical comprehension' (4.) Clerical ability, corresponding to 'Coding Speed' and (5.) Social ability. The correlations between these five specific abilities do not exceed 0.57 and are shown in Appendix 5.A, Table 5.A-1. Using factor analysis we constructed one general ability measure from a combination of the 10 ASVAB scores[23] and the social intelligence test score.

To test Lazear's JAT theory we need to measure the balance in the specific ability levels. The coefficient of variation measured across the individual scores on the five types of specific ability included in our study will serve as an inverse measure of the extent of balance in the various areas of intelligence, i.e. as a measure of spread or 'ability variation'.[24] Unlike Lazear, we do not use the variance as a measure of spread since it is a function of the mean of the specific abilities.

Control variables

We include several control variables into our analysis. Schooling is measured as the total years of education completed. Scores on this variable have a maximum of 20 years of schooling. Since schooling is endogenous in an earnings equation we use instrumental variables to obtain an unbiased estimate of the return to education in exactly the same fashion as in Chapter 3. The approach and choice of instruments will be discussed briefly when dealing with the

[21]The original social ability scale ranged from 1 to 4.

[22]None of the ten components were highly correlated with just one other component.

[23]After corrections for age and education, i.e we use the residuals.

[24]It might be possible that individuals who took the ability test at a later age had more time to develop an unbalanced set of abilities. A simple check of the correlation between the respondent's age and the coefficient of variation reveals that this is not the case. The correlation coefficient is only -0.014 and insignificant.

Table 5.1: Summary statistics

	Employees		Entrepreneurs		**Difference***
Variable	**Mean**	**SD**	**Mean**	**SD**	**in mean**
Hourly pay	9.94	15.84	13.23	23.31	yes
General ability	5.06	1.05	5.16	1.00	no
Verbal ability	6.18	1.04	6.19	1.05	no
Math ability	5.41	1.19	5.39	1.14	no
Technical ability	5.08	1.73	5.54	1.72	yes
Clerical ability	5.20	1.35	5.17	1.28	no
Social ability	4.54	2.97	4.89	2.97	no
Ability variation	0.29	0.14	0.27	0.14	no
Age	27.45	5.13	29.18	4.79	no
Year of birth	1960.46	2.19	1960.09	2.17	no
Male	0.51	0.50	0.63	0.48	yes
Married	0.50	0.50	0.63	0.48	no
Not healthy	0.02	0.15	0.03	0.17	no
Live outside city	0.24	0.43	0.24	0.43	no
Live in South	0.32	0.46	0.26	0.44	yes
Hispanic	0.04	0.20	0.03	0.16	yes
Black	0.09	0.29	0.04	0.19	yes
Education	13.03	2.31	13.10	2.42	no
Education mother	11.70	2.43	12.10	2.27	yes
Education father	11.81	3.33	12.27	3.20	yes
N	42225		2601		

* The numbers in this table are averages over the period 1979-1998. The significance of the difference between entrepreneurs and employees is assessed on the basis of one year only, i.e. 1998.

empirical methodology in the next subsection.

Furthermore several dummies are included, where 'male' is 1 if the respondent is male, 'marital status' is 1 for those who are married, 'Live in South' is equal to 1 if the person lives in the South of the USA, 'not healthy' is 1 if the person is not healthy, 'Live outside city' indicates that a person lives outside the city, 'Hispanic' is 1 if the person is Hispanic, and "Black" is 1 if the person is 'black'. The education levels of the respondent's mother and father are measured in years, with a maximum of 20 years of schooling.

Descriptive statistics[25]

Table 5.1 shows the means and standard deviations of all the variables that are directly or indirectly used in the analyses. The values in Table 5.1 represent the averages of the specific variable over the period 1979-2000, where each year-sample includes only entrepreneurs (right

[25]The description of the summary statistics in this section largely overlaps with that in Section 3.3 of Chapter 3 where we use the same data.

hand column) or employees (left hand column). We highlight the statistics of the variables that are of main interest. First, the average percentage of entrepreneurs in the labor force is six. This percentage is lower than average due to the relatively young cohort studied. We observe at least one spell of entrepreneurship in the period 1979-2000 for twenty four percent of the sample (not shown in the table). Moreover, those individuals who have been entrepreneurs in the observed period, have been so for 3.3 years on average (not shown in the table). Second, both the mean and the standard deviation of the distribution of hourly incomes are higher for entrepreneurs than for employees. The higher standard deviation is (partly) explained by the absence of a 'minimum wage' and the absence of preset salary scales for entrepreneurs. Third, the average general ability level is equal for entrepreneurs and employees. The same is true for the five ability measures and the ability variation, except for technical ability which is higher for entrepreneurs.[26]

5.3.2 Empirical methodology

Our aim is to consistently estimate the returns to both general ability, specific abilities and the ability variation for entrepreneurs and compare these to employees. To this end we estimate the income equation under (5.1) by means of a random effects (RE) model. The RE-model does not assume independence of observations over time.[27]

$$W_{it} = \beta A_i + \gamma E_{it} + \delta AE_{it} + \eta X'_{it} + \zeta S_{it} + \theta ES_{it} + E_{it+1} + c_i + u_{it} \tag{5.1}$$

In equation (5.1), W_{it} represents the log hourly earnings for individual i in year t, A_i is a vector including the ability measures, i.e. general ability, five specific abilities and the variance of abilities (where general ability and the five measures of specific abilities are not simultaneously included into the model), E_{it} is a dummy indicating whether person i is an entrepreneur in year t, and AE_{it} is an interaction of the dummy E_{it} and A_i, such that its coefficients, δ, indicate the magnitude of the difference in returns to ability for entrepreneurs and employees. Furthermore, X'_{it} is a vector including the control variables of Table 5.1, as well as dummies controlling for cohort effects, age effects and macroeconomic shocks using the method developed by Deaton (2000).[28] This method transforms the year dummies, age dummies and birth year dummies such that the year effects add to zero, and are orthogonal

[26]A discussion of some of the control variables tabulated is provided when discussing their use.

[27]A fixed-effects model, which is a neat way of dealing with unobserved time-varying heterogeneity, cannot be used for this purpose, because the measure of ability we use is time-invariant.

[28]We use age instead of experience in the earnings equation, as in Harmon and Walker (1995). Experience is a negative function of education, and is therefore endogenous.

to a time trend (see also Chapter 3 where we use the same method).[29]

Equation (5.1) furthermore includes formal education S_{it} as a control variable, since ability might otherwise pick up part of the schooling effect. However, due to unobserved heterogeneity and the fact that the selection of a particular education level is dependent on the likely returns, education is likely to be an endogenous variable in an income equation (Card, 1999 and Chapters 3 and 4). We try to get rid of the endogenous nature of education by using instrumental variables. The identifying instruments for education are several family background variables measured at the age of 14 of each respondent as in Chapter 3. More in particular, a set of four identifying instruments for education is extracted from the NLSY data: (1) "Magazines present in the household at age 14", (2) "Library card present in the household at age 14", (3) "The presence of a stepparent in the household", and (4) "Number of siblings in the household".[30] The identifying instruments turn out to have a significant effect on the number of years of education attained and no direct effect on incomes.[31] Therefore, the separate effects of ability and education on income can be estimated consistently.

Moreover, ES'_{it}, the interaction of the instrumented schooling variable S_{it} and E_{it}, is included in the equation, since recent evidence indicates that entrepreneurs and employees differ substantially with respect to the effect of schooling on income (See Chapters 3 and 4). Finally, in equation (5.1) c_i is the unobserved individual-specific random effect, and u_{it} a white noise error term.

As was discussed in Chapter 3, a potential problem when estimating equation (5.1) is that E_{it}, i.e. the choice between entrepreneurship and salaried employment, might be endogenous. Individuals might decide to become entrepreneurs because their ability and some of their unobserved characteristics have higher value as an entrepreneur than as an employee. In this chapter we use a different approach than in Chapter 3 to test whether selectivity might be a problem for consistent estimation. We use the panel structure of the data to perform a selectivity test, which is an adapted version of work by Nijman and Verbeek (1992). The test comes down to including a lag of the choice for entrepreneurship (the selection indicator) in the income equation. The underlying assumption here is that sample selection is only related to the idiosyncratic errors u_{it}. Under the null hypothesis, u_{it} should not be correlated with any other entrepreneurship spell than the current spell, i.e. not to previous or future choices. Previous and future choices relate to entrepreneurial types (who have a higher probability

[29] These transformed dummies are included in all regression models, but the coefficients will be omitted from the tables reporting the estimation results.

[30] In Chapter 3 we motivate the choice of this set of identifying instruments for years of education. We also consider its drawbacks empirically.

[31] See Chapter 3 for the test results concerning the quality and validity of this set of instruments.

of being an entrepreneur in any period), whereas the current choice relates to the effect of being an occupational entrepreneur.

A disadvantage of introducing a lag of entrepreneurial status in the income regression is that its coefficient would also measure the human capital effect on income of previous entrepreneurship experience. This could obscure the selection test. A solution to this problem is to include a lead instead of a lag of the entrepreneurship selection decision into the wage equation: Current wages are less likely to be affected by the future decision to be an entrepreneur. Thus, we include E_{it+1} in all estimations.[32]

As will be shown below, all estimation results point at the same conclusion: Future entrepreneurship choices, i.e. E_{it+1}, do not affect current income. Thus, self-selection does not bias our estimation results. Therefore, a correction is not required. Hence, the effect of being an entrepreneur on the returns to ability can be interpreted as an effect of an individual's current status rather than as an effect of being an entrepreneurial type of person.

In the next section we will present the estimation results of equation (5.1). All results are presented with and without education controls to see whether ability picks up any education effects when education is omitted from the income equation. Comparing the two sets of results will also be useful to give insight into the types of ability that are obtained through schooling. The effects of the control variables as well as the results from the selectivity test are discussed following the discussion of all main estimation results.

5.4 Results

5.4.1 Total active workforce

Table 5.2 shows the regression results of income on general ability as well as on the five specific abilities for the total active workforce, i.e. entrepreneurs and employees. In particular, Model 1 in this table shows the effect of general ability without education, whereas Model 2 shows the results upon inclusion of the instrumented years of education regressor. Models 3 and 4 omit the measure of general ability but include the five measures of specific abilities.[33] Model 3 shows the results without education, whereas Model 4 shows the results when the instrumented years of education regressor is included in the regression. Table 5.4 shows the results when estimating Models 1-4 upon inclusion of the measure of the variance of abilities. The results in Tables 5.2 and 5.4 serve as a benchmark for the rest of our estimations, where

[32] The correlation between the current entrepreneurship status and its lead is 0.593.

[33] We never include general ability and specific abilities simultaneously, as the former is an almost perfect linear combination of the latter.

we distinguish between entrepreneurs and employees.

Table 5.2: Returns to general ability and specific abilities: Total active workforce

	Model 1		Model 2		Model 3		Model 4	
Variable	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**
General ability	0.113**	(0.006)	0.043**	(0.011)				
Verbal ability					0.007	(0.007)	0.000	(0.007)
Math ability					0.073**	(0.006)	0.001	(0.011)
Technical ability					0.003	(0.004)	0.012*	(0.005)
Clerical ability					0.029**	(0.005)	0.027**	(0.005)
Social ability					0.006**	(0.002)	0.005**	(0.002)
Education			0.110**	(0.014)			0.111**	(0.014)
E	-0.092**	(0.011)	-0.091**	(0.011)	-0.092**	(0.011)	-0.090**	(0.011)
Male	0.225**	(0.011)	0.255**	(0.012)	0.249**	(0.013)	0.263**	(0.013)
Married	0.048**	(0.005)	0.046**	(0.005)	0.048**	(0.005)	0.046**	(0.005)
Not healthy	-0.040**	(0.014)	-0.038**	(0.014)	-0.040**	(0.014)	-0.038**	(0.014)
Living outside city	-0.098**	(0.008)	-0.090**	(0.008)	-0.097**	(0.008)	-0.089**	(0.008)
Living in South	-0.043**	(0.012)	-0.049**	(0.012)	-0.047**	(0.012)	-0.050**	(0.012)
Hispanic	0.127**	(0.029)	0.021	(0.032)	0.107**	(0.029)	0.014	(0.031)
Black	0.000	(0.021)	-0.072**	(0.022)	-0.030	(0.020)	-0.075**	(0.021)
Education mother	0.018**	(0.003)	-0.007	(0.004)	0.019**	(0.003)	-0.007	(0.004)
Education father	0.017**	(0.002)	-0.001	(0.003)	0.016**	(0.002)	-0.001	(0.003)
E t+1	-0.013	(0.010)	-0.013	(0.010)	-0.012	(0.010)	-0.012	(0.010)
Intercept	0.935**	(0.191)	1.298**	(0.195)	0.316†	(0.190)	1.050**	(0.210)
N	44826		44826		44826		44826	
R^2 Within	0.42		0.42		0.42		0.42	
R^2 Between	0.41		0.42		0.42		0.43	
R^2 Overall	0.40		0.40		0.40		0.41	
χ^2	$_{df(61)}$ 31656		$_{df(62)}$ 31756		$_{df(65)}$ 31719		$_{df(66)}$ 31829	

Significance levels : † : 10% $*$: 5% $**$: 1%

Note: The dependent variable of the regression is the log of the average hourly income, i.e. the total annual income divided by the number of actual hours worked. Control variables that are included in the regression but are not reported in this table are year, age and birth year dummies that have been transformed according to the Deaton method (2000) as described in the text to control for cohort effects, age effects and macroeconomic shocks. Estimates are obtained by using a Random Effects model.

Levels of (specific areas of) intelligence

As expected, Model 1 in Table 5.2 shows that general ability has a significant positive effect on income. Upon including education (Model 2) into the regression, the effect of general ability has become almost three times smaller and is still significant. General ability thus indeed picks up part of the education effect when education is not included.[34] An increase in one's level of general intelligence of one standard deviation, increases individual earnings by approximately four percent.

Models 3 and 4 show the effects of the five specific abilities on performance. For verbal ability, we expected a small positive effect on income, but the results from both Models 3 and 4 seem to indicate that there is no effect. However, according to Dougherty (2000) and McIntosh and Vignoles (2001) this zero effect might be due to the fact that the returns to verbal ability are non-linear. We explore this suggestion by estimating quadratic effects for all specific abilities and general ability. Table 5.3 summarizes the significant results. The returns to verbal ability are indeed significantly non-linear in the form of an inverted u-shape. They are only significantly positive for the lower 30 percent of the verbal ability distribution of the sample. Hence, only at the basic level of verbal ability does an increase in the level increase an individual's income.

Model 3 further shows that mathematical ability has a strong and positive effect on income. This is according to our expectation based on previous studies. However, when we include the instrumented education variable into the regression (Model 4), the effect of math ability on income diminishes and turns out insignificant.[35] A further inspection of the data though, reveals that the effect of mathematical ability on income is non-linear as well. Table 5.3 shows that the returns to math ability (Column 1) are U-shaped. Only half of the individuals in the sample from the total active workforce has a level of mathematical ability above the level corresponding with the minimum of the curve, implying that the returns to mathematical ability are positive for (this upper) half of the population.

Moreover, the results from estimating Model 3 show that technical ability has no effect on income. The effect becomes marginally significant when we control for education (Model 4). Table 5.3 shows that the returns to technical ability are significantly non-linear in the form of an inverted U-shape. Only one quarter of the population has an ability level in the range of negative returns to technical ability, whereas the lower three quarter of the distribution experiences positive returns.

[34]The difference in the effect of general intelligence between Models 1 and 2 is illuminated in Appendix 5.A Table 5.A-3. The first column of this table shows the ability related part of the results of the first stage equation in which education is explained by all the independent variables included in Model 1 along with the four identifying instruments discussed in the previous section. The table indicates that the instrumented education variable is strongly determined by general ability.

[35]The difference in the effect of mathematical intelligence between Models 3 and 4 becomes clear when looking at Model 2 of Appendix 5.A, Table 5.A-3. The table indicates that the instrumented education variable is most strongly determined by mathematical ability.

Table 5.3: Nonlinear returns to ability

	Total active workforce		Employees		Entrepreneurs	
	Shape	Sample % to the LHS of minimum/ maximum	Shape	Sample % to the LHS of minimum/ maximum	shape	Sample % to the LHS of minimum/ maximum
General ability	Linear	n.a.	Linear	n.a.	linear	n.a.
Verbal ability	Inverted U	29.7	Inverted U	31.4	Inverted U	9.7
Math ability	U-shaped	53.7	U-shaped	52.0	U-shaped	68.7
Technical ability	Inverted U	74.9	Inverted U	68.5	Inverted U	78.9
Clerical ability	U-shaped	7.2	U-shaped	5.5	Linear	n.a.
Social ability	Inverted U	60.2	Inverted U	60.5	Linear	n.a.
Ability variation	Linear	n.a.	Linear	n.a.	Linear	n.a.

LHS = Left Hand Side.

Note: Results in the first column of this table are based on Models 2 and 4 in Table 5.2, and Models 2 and 4 in Table 5.4. The only difference is that the quadratic terms of all different areas and forms of ability have now been included into the regression equations. Results in the second and third column are based on Models 2 and 4 in Table 5.5 and Models 2 and 4 in Table 5.6. Again, the only difference is that the quadratic terms of all different areas and forms of ability and their interactions with E have now been included.

The final two specific abilities included in Models 3 and 4, clerical ability and social ability, show significantly positive returns. Clerical ability has a strong and positive linear effect on performance, whereas the effect of social ability is significantly positive as well, but smaller. Including or excluding education into the analysis does not change these results. Additionally, Table 5.3 shows that the returns to social ability are non-linear (inverted U-shape). The lowest sixty percent of the social ability distribution of the active workforce experience positive returns to this ability.

We conclude that, in general, when also taking non-linear effects into account, the returns to general ability and all specific abilities are significant. Returns to general ability are positive, whereas returns to specific abilities may be negative for some parts of the specific ability distributions and positive for others. Negative returns are not uncommon in this type of analysis. As ability cannot be marketed separately, one cannot expect the positive linear pricing model to be generally valid. Specific abilities may be correlated to unobserved heterogeneity or tastes for specific professions (that have specific income distributions).

Table 5.4: Returns to ability variation: Total active workforce

	Model 1		**Model 2**		**Model 3**		**Model 4**	
Variable	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**
General ability	0.108**	(0.006)	0.039**	(0.011)				
Verbal ability					0.007	(0.007)	0.000	(0.007)
Math ability					0.072**	(0.006)	0.000	(0.011)
Technical ability					0.002	(0.005)	0.011*	(0.005)
Clerical ability					0.028**	(0.005)	0.026**	(0.005)
Social ability					0.005*	(0.002)	0.004†	(0.002)
Ability variation	-0.102**	(0.041)	-0.100*	(0.040)	-0.077†	(0.046)	-0.046	(0.046)
Education			0.109**	(0.014)			0.111**	(0.014)
E	-0.092**	(0.011)	-0.091**	(0.011)	-0.092**	(0.011)	-0.090**	(0.011)
Male	0.222**	(0.011)	0.252**	(0.012)	0.248**	(0.013)	0.262**	(0.013)
Married	0.048**	(0.005)	0.046**	(0.005)	0.048**	(0.005)	0.046**	(0.005)
Not healthy	-0.040**	(0.014)	-0.038**	(0.014)	-0.040**	(0.014)	-0.038**	(0.014)
Living outside city	-0.098**	(0.008)	-0.091**	(0.008)	-0.097**	(0.008)	-0.089**	(0.008)
Living in South	-0.043**	(0.012)	-0.049**	(0.012)	-0.047**	(0.012)	-0.050**	(0.012)
Hispanic	0.127**	(0.029)	0.022	(0.032)	0.108**	(0.029)	0.015	(0.031)
Black	0.002	(0.021)	-0.070**	(0.022)	-0.029	(0.020)	-0.074**	(0.021)
Education mother	0.018**	(0.003)	-0.007	(0.004)	0.019**	(0.003)	-0.007	(0.004)
Education father	0.017**	(0.002)	-0.001	(0.003)	0.016**	(0.002)	-0.001	(0.003)
E t+1	-0.013	(0.010)	-0.013	(0.010)	-0.012	(0.010)	-0.012	(0.010)
Intercept	0.968**	(0.191)	1.327**	(0.196)	0.360†	(0.192)	1.076**	(0.211)
N	44826		44826		44826		44826	
R^2 Within	0.42		0.42		0.42		0.42	
R^2 Between	0.41		0.42		0.42		0.43	
R^2 Overall	0.40		0.40		0.40		0.41	
χ^2	$_{df(62)}$ 31665		$_{df(63)}$ 31765		$_{df(66)}$ 31723		$_{df(67)}$ 31833	

Significance levels : † : 10% $*$: 5% $**$: 1%

Note: The dependent variable of the regression is the log of the average hourly income, i.e. the total annual income divided by the number of actual hours worked. Control variables that are included in the regression but are not reported in this table are year, age and birth year dummies that have been transformed according to the Deaton method (2000) as described in the text to control for cohort effects, age effects and macroeconomic shocks. Estimates are obtained by using a Random Effects model.

Variance of specific areas of intelligence

Models 1-4 in Table 5.4 show the effect of the variation in abilities on performance. The results in Model 1 and 2 show that ability variation is negatively related to the income of the average labor force participant. In other words, people benefit from a more balanced set of abilities. Controlling for education reduces the significance and magnitude of this effect slightly. Moreover, controlling for the five specific abilities instead of only general ability (in Models 3 and 4) eliminates the effect of ability variation.

Control variables

The effects of the control variables are discussed based on Tables 5.2 and 5.4. These specifications are comparable to income equations presented in the numerous previous studies.[36]

The results show that the returns to education are on average 11 percent. This is consistent with Chapter 3. The estimates further show that entrepreneurs earn, on average and ceteris paribus, nine percent lower incomes than employees. Males earn 25 percent more than their female counterparts. Married individuals earn almost five percent more than others on average. Bad health conditions decrease earnings by four percent. Not living in a city or in the South of the USA decreases earnings by nine and five percent, respectively. Being Hispanic has no effect on income whereas blacks earn, on average, seven percent lower incomes than whites. Parental education levels are not related to the income of their offspring. All these results are equal in Tables 5.2 and 5.4. Adding ability inequality has no effect. It is also worth noting that the effects of gender, marital and health status and residential area are independent of education, whereas effects of race and parental background are determined through the schooling channel. Finally, as was discussed earlier, the selection effect -indicated by future entrepreneurship status- turns out insignificant (in all models in all tables). This means that a correction is not required.

5.4.2 Entrepreneurs versus employees

In Tables 5.5 and 5.6 a distinction is made between entrepreneurs (E=1) and employees (E=0), by adding interactions of entrepreneurial status on the one hand, and general ability, the five specific abilities and the balance in abilities on the other hand. Furthermore, an interaction of the instrumented education variable and entrepreneurial status is included where education is included as a regressor. Again, we present the results with and without controlling for education. The discussion pertains to the results based on the equations that include education.[37]

[36]The results pertaining to the control variables are very similar across all models and tables.

[37]This choice is made since the difference between the results including or excluding education are similar for entrepreneurs and employees and thus to what we already described above.

Levels of (specific areas of) intelligence

Model 2 in Table 5.5 shows the difference in returns to general ability between entrepreneurs and employees. The difference is insignificant: Entrepreneurs benefit as much from their general ability as do employees. Table 5.3 shows that the effect of general intelligence on income is linear for both entrepreneurs and employees.

Model 4 shows the differences in the returns to the five specific abilities between employees and entrepreneurs. As it turns out, there are substantial differences between entrepreneurs and employees for all five specific abilities. Again, Table 5.3 provides a summary of the results when quadratic (cross-) terms are included into the analysis. This gives additional insights, since Table 5.5 assumes all effects to be linear.

The effect of verbal ability on income is insignificant for employees and negative for entrepreneurs if it is assumed linear. Table 5.3 shows that the effect is significantly non-linear though for both groups of labor force participants. For both sub-samples, the relationship between verbal ability and income follows an inverted U-shaped curve. Out of the employee population, 31 percent of the verbal ability distribution is located to the left hand side of the maximum, whereas for entrepreneurs this percentage is only ten. This implies that verbal ability is productive for both groups, but only at the (for entrepreneurs very) low end of the distribution. Hence, as was the case for the aggregate active workforce, only a (very) basic level of verbal ability seems to be productive for each group.

The effects of mathematical ability on income shown in Model 4 of Table 5.5 depict the same, but inverse, pattern as the effect of verbal ability. The effect is zero for employees as well as entrepreneurs. Table 5.3 shows the patterns that emerge when quadratic (cross-) terms are included into the equations. Mathematical ability turns out to be productive only at the high end of the sample distribution of mathematical ability levels. For employees the minimum of the U-shaped curve lies around the median of the mathematical ability distribution, whereas for entrepreneurs the minimum is observed around the upper third of the distribution. Hence, also in this case, a smaller proportion of entrepreneurs relative to employees benefits in terms of their incomes from mathematical ability.[38]

[38] The distribution of mathematical ability is approximately the same for entrepreneurs and employees.

Table 5.5: Returns to general ability and specific abilities: Entrepreneurs versus employees

	Model 1		Model 2		Model 3		Model 4	
Variable	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**
General ability	0.110^{**}	(0.006)	0.042^{**}	(0.011)				
General ability*E	0.043^{**}	(0.010)	0.002	(0.013)				
Verbal ability					0.009	(0.007)	0.002	(0.007)
Verbal ability*E					-0.029^{*}	(0.012)	-0.041^{**}	(0.012)
Math ability					0.071^{**}	(0.006)	0.003	(0.011)
Math ability*E					0.032^{**}	(0.012)	-0.024	(0.015)
Technical ability					-0.001	(0.004)	0.007	(0.005)
Technical ability*E					0.064^{**}	(0.007)	0.070^{**}	(0.007)
Clerical ability					0.034^{**}	(0.005)	0.032^{**}	(0.005)
Clerical ability*E					-0.075^{**}	(0.009)	-0.074^{**}	(0.009)
Social ability					0.006^{**}	(0.002)	0.004^{*}	(0.002)
Social ability*E					0.013^{**}	(0.003)	0.010^{**}	(0.003)
Education			0.108^{**}	(0.014)			0.107^{**}	(0.014)
Education*E			0.047^{**}	(0.010)			0.066^{**}	(0.010)
E	-0.098^{**}	(0.011)	-0.120^{**}	(0.012)	$-0.107^{\dagger}$	(0.064)	0.195^{*}	(0.080)
Male	0.225^{**}	(0.011)	0.255^{**}	(0.012)	0.248^{**}	(0.013)	0.262^{**}	(0.013)
Married	0.048^{**}	(0.005)	0.046^{**}	(0.005)	0.050^{**}	(0.005)	0.048^{**}	(0.005)
Not healthy	-0.040^{**}	(0.014)	-0.038^{**}	(0.014)	-0.041^{**}	(0.014)	-0.039^{**}	(0.014)
Living outside city	-0.098^{**}	(0.008)	-0.091^{**}	(0.008)	-0.097^{**}	(0.008)	-0.090^{**}	(0.008)
Living in South	-0.043^{**}	(0.012)	-0.050^{**}	(0.012)	-0.049^{**}	(0.012)	-0.052^{**}	(0.012)
Hispanic	0.127^{**}	(0.029)	0.019	(0.032)	0.106^{**}	(0.029)	0.012	(0.031)
Black	-0.002	(0.021)	-0.074^{**}	(0.022)	-0.031	(0.020)	-0.074^{**}	(0.021)
Education mother	0.018^{**}	(0.003)	$-0.007^{\dagger}$	(0.004)	0.019^{**}	(0.003)	-0.007	(0.004)
Education father	0.017^{**}	(0.002)	-0.001	(0.003)	0.016^{**}	(0.002)	-0.001	(0.003)
E t+1	-0.015	(0.010)	-0.016	(0.010)	-0.014	(0.010)	-0.017	(0.010)
Intercept	0.942^{**}	(0.191)	1.307^{**}	(0.195)	$0.328^{\dagger}$	(0.190)	1.039^{**}	(0.210)
N	44826		44826		44826		44826	
R^2 Within	0.42		0.42		0.42		0.42	
R^2 Between	0.41		0.42		0.42		0.43	
R^2 Overall	0.40		0.40		0.40		0.41	
χ^2	$_{df(62)}$ 31688		$_{df(64)}$ 31828		$_{df(70)}$ 32161		$_{df(72)}$ 32262	

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Note: The dependent variable of the regression is the log of the average hourly income, i.e. the total annual income divided by the number of actual hours worked. Control variables that are included in the regression but are not reported in this table are year, age and birth year dummies that have been transformed according to the Deaton method (2000) as described in the text to control for cohort effects, age effects and macroeconomic shocks. Estimates are obtained by using a Random Effects model.

Moreover, the returns to technical ability are large, positive and significant for entrepreneurs and zero for employees according to Model 4 of Table 5.5. Table 5.3 shows that the returns to technical intelligence are non-linear both for employees and entrepreneurs. The far majority of entrepreneurs has positive returns. The zero returns to technical ability we found for employees in Model 4 of Table 5.5 disappear when taking account of non-linearity. Although a lower proportion than in the case of entrepreneurs, two thirds of employees have positive returns to technical ability.

As expected, clerical ability has positive returns for employees but not for entrepreneurs. For entrepreneurs, the total linear effect is even significantly negative. Table 5.3 provides little additional insight into the result observed in Table 5.5: Non-linearities play no role.

The final result of Model 4 in Table 5.5 to be discussed pertains to social ability. Some researchers have argued that social ability is especially valuable for entrepreneurs: For instance, they have to build relations with customers, providers of capital and possibly manage employees (Baron and Markman, 2003). The results in both Table 5.5 and Table 5.3 indeed show that social ability renders much higher returns for entrepreneurs than for employees.

Variance of specific areas of intelligence

Models 2 and 4 in table 5.6 show the effects of ability variation on income. The JAT theory states that individuals who are JATs select themselves more often into entrepreneurship since they have a comparative advantage when doing so. We expect this comparative advantage to show up in the form of a higher return to being a JAT as an entrepreneur than as an employee.

The results in Table 5.6 support this expectation. Entrepreneurs have a very large and significant return to a balanced set of abilities (i.e. low variation) in all models estimated. Employees, on the other hand, have a much smaller return to a balanced skill set according to Models 1 and 2 which even disappears if we use controls for the levels of the five specific areas of intelligence (Models 3 and 4).[39] We thus find support for the fact that being a JAT is valuable for entrepreneurs but hardly so for employees. Our result is consistent with Lazear's JAT theory: JAT's have a comparative advantage in entrepreneurship.[40]

[39] Appendix 5.A, Table 5.A-2 (bottom row) shows the returns to variation in ability upon inclusion of quadratic terms of the non-linear specific abilities. The differences between entrepreneurs and employees turn up even more pronounced. Being a specialist has even a positive association with employee performance (.28 in Table 5.A-2 as compared to -.025 in Table 5.6).

[40] Based on information about the high school curriculum of individuals included in the NLSY sample, we could replicate Lazear's course-related measure of JAT. Including this JAT-measure in a regression in which the dependent variable is one if individuals have ever been entrepreneur and zero otherwise, leads to the closest replication of Lazear's original test as we can get. Our result is consistent with his result: A higher degree of JAT leads to a higher probability of entrepreneurship. However, including measures of intelligence and years of education into the regression as control variables renders the effect insignificant. This is one of the reasons why we use the coefficient of variation instead of the variance of the scores across the five specific abilities. Both Lazear's JAT measure, and such a variance measure are functions of the mean of a weighed combination of the specific abilities and this is likely to affect the results.

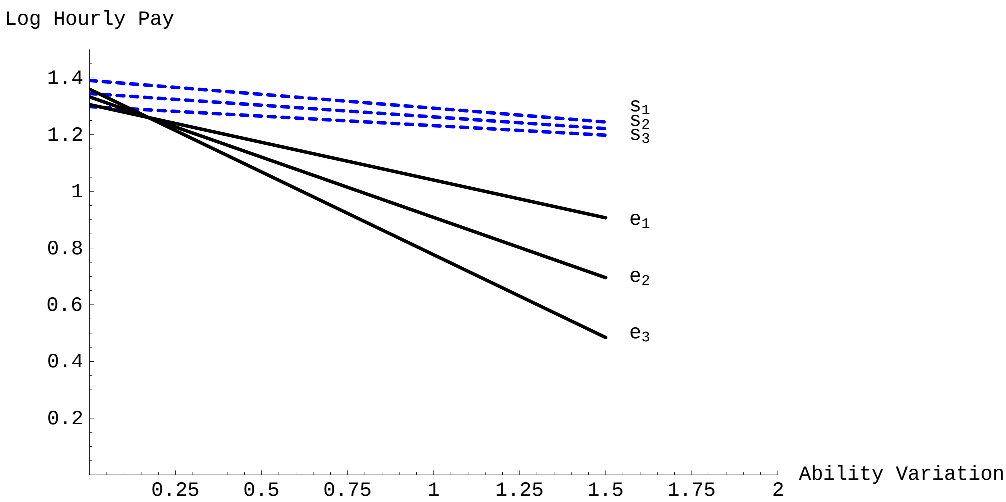
Table 5.6: Returns to variation in abilities: Entrepreneurs versus employees

	Model 1		**Model 2**		**Model 3**		**Model 4**	
Variable	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**
General ability	0.107**	(0.006)	0.039**	(0.011)				
General ability*E	0.022*	(0.011)	-0.019	(0.014)				
Verbal ability					0.009	(0.007)	0.002	(0.007)
Verbal ability*E					-0.024†	(0.012)	-0.036**	(0.012)
Math ability					0.071**	(0.006)	0.003	(0.011)
Math ability*E					0.025*	(0.012)	-0.031*	(0.015)
Technical ability					-0.002	(0.005)	0.007	(0.005)
Technical ability*E					0.057**	(0.007)	0.063**	(0.007)
Clerical ability					0.032**	(0.005)	0.031**	(0.005)
Clerical ability*E					-0.085**	(0.009)	-0.083**	(0.009)
Social ability					0.005*	(0.002)	0.004†	(0.002)
Social ability*E					0.006	(0.004)	0.003	(0.004)
Ability variation	-0.080*	(0.041)	-0.077†	(0.041)	-0.056	(0.046)	-0.025	(0.046)
Ability variation*E	-0.389**	(0.076)	-0.395**	(0.076)	-0.395**	(0.085)	-0.388**	(0.085)
Education			0.107**	(0.014)			0.107**	(0.014)
Education*E			0.047**	(0.010)			0.066**	(0.010)
E	0.014	(0.024)	-0.006	(0.025)	0.129	(0.081)	0.431**	(0.094)
Male	0.222**	(0.011)	0.253**	(0.012)	0.247**	(0.013)	0.261**	(0.013)
Married	0.048**	(0.005)	0.046**	(0.005)	0.050**	(0.005)	0.048**	(0.005)
Not healthy	-0.041**	(0.014)	-0.039**	(0.014)	-0.041**	(0.014)	-0.039**	(0.014)
Living outside city	-0.099**	(0.008)	-0.091**	(0.008)	-0.097**	(0.008)	-0.091**	(0.008)
Living in South	-0.043**	(0.012)	-0.050**	(0.012)	-0.049**	(0.012)	-0.052**	(0.012)
Hispanic	0.126**	(0.029)	0.019	(0.032)	0.106**	(0.029)	0.011	(0.031)
Black	-0.001	(0.021)	-0.072**	(0.022)	-0.030	(0.020)	-0.074**	(0.021)
Education mother	0.018**	(0.003)	-0.007	(0.004)	0.019**	(0.003)	-0.007	(0.004)
Education father	0.017**	(0.002)	-0.001	(0.003)	0.016**	(0.002)	-0.001	(0.003)
E t+1	-0.015	(0.010)	-0.017	(0.010)	-0.015	(0.010)	-0.017†	(0.010)
Intercept	0.977**	(0.191)	1.339**	(0.196)	0.368†	(0.192)	1.061**	(0.210)
N	44826		44826		44826		44826	
R^2 Within	0.42		0.42		0.42		0.42	
R^2 Between	0.41		0.42		0.42		0.43	
R^2 Overall	0.40		0.40		0.40		0.41	
χ^2	$_{df(64)}$ 31742		$_{df(66)}$ 31881		$_{df(72)}$ 32125		$_{df(74)}$ 32305	

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Note: The dependent variable of the regression is the log of the average hourly income, i.e. the total annual income divided by the number of actual hours worked. Control variables that are included in the regression but are not reported in this table are year, age and birth year dummies that have been transformed according to the Deaton method (2000) as described in the text to control for cohort effects, age effects and macroeconomic shocks. Estimates are obtained by using a Random Effects model.

Finally, we checked whether the JAT effect differs between people of high and low general ability levels. Is the value of being a JAT higher for more intelligent people than for less intelligent people? To answer this question, we re-estimate Model 2 in Table 5.6 and include the three-way interaction of general ability, the variation in abilities and E.[41] The results of this analysis are depicted in Figure 5.1. At higher ability levels the JAT effect for entrepreneurs is more pronounced than at lower levels of general ability. For employees there are no effects at all.



Note. s denotes salaried employees and e entrepreneurs. Subscript 1, 2 and 3 are low, average and high levels of general ability, respectively (1 sd difference between 1, 2 and 3)

Figure 5.1: Returns to ability variation evaluated at three levels of general ability

5.5 Discussion and conclusion

In this chapter we have measured the returns to intelligence for entrepreneurs vis-à-vis employees. In particular, we have assessed the labor market value of three different but related aspects or forms of ability, i.e., the level of general ability, the levels of specific abilities and ability variation. We now conclude the chapter with some final observations.

For the whole active workforce, we find that both general ability and all forms of specific ability contribute significantly to increasing individual income levels. In particular, whereas general ability, technical, clerical and social ability have the expected positive effect for the

[41] All other (partial) interaction effect of these three variables are also included. Estimation results for this three way interaction effects are available upon request.

entire sample, the effects of the verbal and mathematical abilities are not positive for the entire sample. Verbal ability has only a positive return at the lower end of the distribution, whereas mathematical ability has only a positive effect on income at the higher end of the distribution of this ability. Thus, the effects are non-linear. These results largely support the hypotheses formulated in Section 5.2.

Returns to general ability are as high for entrepreneurs as for employees. The effects of the five specific areas of intelligence are not the same for entrepreneurs and employees. As expected, the effect of technical and social intelligence is larger for entrepreneurs than for employees. The effects of verbal and mathematical ability are non-linear and distinct for entrepreneurs versus employees. For both segments of the labor market the returns to verbal intelligence are only positive at the lower end of their distribution, whereas the effect of mathematical ability is only positive at the higher end of the distribution. However, for entrepreneurs, the percentage of the population for which the effects of verbal and mathematical abilities are positive is much smaller than for employees. Hence, the total value of these two specific abilities is lower for entrepreneurs than for employees. This rejects the hypotheses we had formulated, basically based on our ignorance, that the returns to verbal and math ability would be similar for entrepreneurs and employees.

Based on Lazear's Jack-of-All-Trades theory, we also hypothesized that the return to being a JAT would be higher for entrepreneurs than for employees. Our results strongly support this hypothesis. The JAT measure that we use, i.e. the coefficient of variation of an individual's scores on the five specific abilities, can be considered a measure of JAT-endowment, unlike Lazear's own measures. We conclude that being a JAT increases the likelihood of successful entrepreneurship. Moreover, being a JAT is more beneficial for entrepreneurs with high levels of intelligence than for those with low levels of intelligence. All in all, our empirical results are supportive of the (largely explorative) hypotheses that we based on the somewhat scarce and scattered psychological and economic literature in this area.

Although we have an understanding of why entrepreneurs would collect higher returns than employees from their social and technical intelligence and their balanced skill sets, we have some difficulties in understanding our results pertaining to the three remaining specific types of ability, i.e. language, mathematical ability, and clerical ability. We do however have some intuitions. It is indeed conceivable that clerical and language abilities are less important for entrepreneurs. It could be the case that the skills that are least productive for entrepreneurs are the ones that are required for tasks that entrepreneurs can more easily delegate or outsource. But actually, in order to understand these differences, research is needed that links the choice for entrepreneurship versus employment with details of the specific tasks that each has to perform in the professions that they select.

Another remaining question is the relationship between team entrepreneurship (or start-ups with personnel) on the one hand, and the strength of the effect on entrepreneurs' incomes of being a JAT on the other hand. One could expect that being a JAT is less valuable when an enterprise is started up with a (multidisciplinary and/or complementary) team of entrepreneurs. In the USA there is a high incidence of team entrepreneurship among existing firms, i.e. 52 percent (Aldrich, Carter and Ruef, 2004). This high percentage makes it likely that a substantial fraction of the incumbent entrepreneurs have also started as a team, also in our (and Lazear's) sample. This might imply that the measured effects of being a JAT are underestimates of the real effect of JAT on "solo" entrepreneurship.

We conclude that the returns to technical and social intelligence, as well as to a balanced ability set are all higher for entrepreneurs than for employees. The returns to language, mathematical and clerical intelligence are higher for employees. Hence, our analysis shows that there are clear differences between entrepreneurs and employees with respect to the endowed characteristics that increase labor market performance.

These findings give a quite clear handle as to who should be stimulated to become an entrepreneur and who should not. Especially, since entrepreneurs render large returns to society in terms of growth, innovation and the creation of labor, which often surpass their private returns, it is clear that attracting people with balanced skill sets, who are both technically and socially sophisticated, would be beneficial for the economy (as well as for the individuals themselves). Providers of equity, loans, subsidies, and licenses -whether private or public authorities-, might use these insights to stimulate a larger and more successful entrepreneurial economy.

Appendix 5.A

Table 5.A-1: Correlation between the five specific abilities, general ability and ability variation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Verbal Ability	1						
(2) Math ability	0.5702	1					
(3) Technical ability	0.4318	0.5172	1				
(4) Clerical ability	0.4305	0.4185	0.193	1			
(5) Social ability	0.0658	0.0428	0.0266	0.0248	1		
(6) Ability variation	-0.1904	-0.2644	-0.307	-0.2613	-0.4421	1	
(7) General ability	0.7916	0.8311	0.6915	0.5738	0.1458	-0.3562	1

Table 5.A-2: Nonlinear returns in ability while including variation in abilities

	Total active workforce		Employees		Entrepreneurs	
	Shape	Sample % to the LHS of minimum/ maximum	Shape	Sample % to the LHS of minimum/ maximum	shape	Sample % to the LHS of minimum/ maximum
General ability	Linear	n.a.	linear	n.a.	linear	n.a.
Verbal ability	Inverted U	28.2	Inverted U	29.3	Inverted U	14.4
Math ability	U-shaped	50.7	U-shaped	48.6	U-shaped	80.1
Technical ability	Inverted U	73.0	Inverted U	70.7	Inverted U	95.3
Clerical ability	U-shaped	1.9	U-shaped	0.03	U-shaped	98.6
Social ability	Inverted U	60.2	Inverted U	60.5	U-shaped	55.1
Ability variation*	Linear	n.a.	Linear	.28 (.10)	Linear	-.84 (.20)

LHS = Left Hand Side.

Note: Results in the first column of this table are based on Models 2 and 4 in Table 5.2, and Models 2 and 4 in Table 5.4. The difference is that the regressions reported in this table include quadratic terms of the various areas and forms of ability. Results in the second and third column are based on Models 2 and 4 in Table 5.5 and Models 2 and 4 in Table 5.6. The difference is that the regressions reported in this table include quadratic terms of the various areas and forms of ability and their interactions with E. Reported numbers are based on the coefficient estimates (all significant).

*This row reports the coefficients for the linear effect of ability variation on performance for the subsamples of entrepreneurs and employees. Standard errors are in parentheses.

Table 5.A-3: First stage regression results: Explaining education

	Model 1		**Model 2**		**Model 3**		**Model 4**	
Variable	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**	**Coeff**	**(SE)**
General ability	0.581**	(0.030)			0.583**	(0.032)		
Verbal ability			0.038	(0.035)			0.041	(0.035)
Math ability			0.633**	(0.031)			0.631**	(0.031)
Technical ability			-0.080**	(0.022)			-0.084**	(0.023)
Clerical ability			0.010	(0.024)			0.004	(0.025)
Social ability			0.003	(0.009)			-0.002	(0.010)
Ability variation					0.053	(0.207)	-0.253	(0.231)
N	44826		44826		44826		44826	
χ^2	$_{df(64)}$ 8540		$_{df(68)}$ 8996		$_{df(65)}$ 8540		$_{df(69)}$ 8998	
R^2 Within	0.13		0.13		0.13		0.13	
R^2 Between	0.35		0.39		0.35		0.39	
R^2 Overall	0.31		0.35		0.31		0.35	

Significance levels : † : 10% $*$: 5% $**$: 1%

Note: The dependent variable of the regression is the respondent's number of years of education. Control variables that are included in the regression but are not reported in this table are for Model 1: all the Model 1 independent variables of Table 5.2; Model 2: all the Model 3 independent variables of Table 5.2 ; Model 3: all the Model 1 independent variables of Table 5.4; Model 4: all the Model 3 independent variables of Table 5.4. Moreover, all regressions include year, age and birth year dummies that have been transformed according to the Deaton method (2000) as described in the text to control for cohort effects, age effects and macroeconomic shocks. Finally, each of the first stage equations includes the four parental background related identifying instruments as described in the text. Just as in Chapter 3 Feasible Generalized Least Squares (FGLS) is used to estimate the first stage equation. Using FGLS makes sure the correct error structure is estimated even in the case of small or no variation over time in the dependent variable (education in our case).

Chapter 6

Summary, policy implications and reflection

6.1 Summary and policy implications

This chapter summarizes the main findings of the previous chapters and combines these results into several policy implications and suggestions for further research.

Chapter 1 sets out the social benefits of entrepreneurship. Entrepreneurship is associated with job creation, innovation, R&D etc.. However, as the benefits of entrepreneurship are realized only when an entrepreneur is successful, knowledge of the determinants of entrepreneurial success is imperative. The main determinant addressed in this book is formal education. This focus is motivated as follows. First, formal education is a factor that possibly affects entrepreneurship performance and that can be shaped to some extent by public policy. Second, the effect of formal education on the labor market performance of employees has been measured extensively. Employee literature devotes a great deal of attention to the role of education relative to ability or intelligence. Doing the same for the effect of intelligence on entrepreneur performance, we defined the following three research questions:

1. What is the current state of research on the relationship between formal education and entrepreneurship?

2. How high are the returns to formal education for entrepreneurs as compared to employees?

3. What are the returns to (different types of) intelligence, on top of the effect of formal education?

6.1.1 What is the current state of research on the relationship between formal education and entrepreneurship?'

Chapter 2 provides a review of empirical studies on the impact of formal education on entrepreneurship selection and performance in industrial countries. After examining almost a hundred studies measuring the returns to education for entrepreneurs, this chapter summarizes the main effects found in literature and explains the variation in results. The overview formulates the current state of affairs of this research. Five general conclusions result from this meta-analysis.

First, education has no significant effect on selection into entrepreneurship. Second, the effect of education on entrepreneurial performance is positive and significant. Third, based on the measurement of conditional correlations, the returns to a marginal year of schooling in the USA are 6.1% for entrepreneurs.[1] Fourth, the difference in returns between entrepreneurs and employees is slightly positive in the USA and slightly negative in Europe. Fifth, the returns to schooling for entrepreneurs seem to be higher in the USA than in Europe.

Almost one hundred studies are included in the meta-analysis. Still, compared with common practice in estimating returns to education for employees, a number of issues remain that have not been addressed at all or have not been addressed in a satisfactory manner.

First, all studies estimate the "effect" of schooling as a (conditional) correlation. It is quite plausible that this correlation does not measure the causal effect. Schooling is endogenous to one's potential (or expected) performance in the labor market. Although future earnings are not the only reason to pursue an education, the prospect of higher incomes entices many students to stay in school longer. In the established returns to education literature that focuses on employees, researchers attempt to correct for the endogeneity of education by using instrumental variables of one type or another, by running controlled experiments or by studying twins. Studies that have used these estimation strategies have found that the returns to education for employees are severely biased downwards when not taking account of the endogeneity of schooling. A major challenge for entrepreneurship research is to apply these estimation strategies and find out whether the returns to education for entrepreneurs as observed in the meta-analysis are also biased downwards.

Another important issue is that the choice between entrepreneurship and salaried employment might be endogenous. Individuals might decide to become entrepreneurs because their education and some of their unobserved characteristics are more valuable as an entrepreneur than as an employee. This could bias the estimate of the returns to education. About one

[1]The average of 6.1% is made up of a small number of studies that analyze an other country than the USA i.e., 19% of the studies do not analyze the USA.

fifth of the studies in the meta-analysis try to correct for this possible bias. The results suggest that the returns to education are not influenced by the choice for entrepreneurship or salaried employment. The question is whether this conclusion also holds when the endogenous nature of both education and entrepreneurship are taken into account at the same time.

6.1.2 How high are the returns to formal education for entrepreneurs as compared to employees?

Using family background variables as instruments for education and entrepreneurship

Chapter 3 makes a first attempt to apply estimation techniques that deal with endogeneity. We take account of the endogenous nature of education as well as the endogenous selection into entrepreneurship/salaried employment. We use family background variables as instruments for education and the probability that the father was an entrepreneur or the religious background of the individual as instrument for the selection into entrepreneurship. We use 19 waves of the USA NLSY database to compare the returns to education for entrepreneurs and employees.

Our benchmark estimate based on (conditional) correlations indicates that the returns to education are slightly higher for entrepreneurs (6.9 percent) than for employees (6.0 percent). This is in line with the estimates found in the meta-analysis. However, when we control for ability and take the endogenous nature of schooling and selection in an income equation into account, the returns to education jump to 9.9 percent for employees and 18.3 percent for entrepreneurs. The IV estimate for employees is comparable to previous findings using various identification strategies. The second estimate, which is larger and more novel, leads to the remarkable result that entrepreneurs' returns to education are not slightly higher, but an impressive 85 percent higher than the returns to education for employees. The increase in returns to education when applying IV is caused by the endogenous nature of education, as we find that selection has no influence. In Chapter 5 we find additional support that selection plays no role.

So why is education more valuable for entrepreneurs? Our research findings of Chapter 3 suggest the following explanation. Entrepreneurs have more freedom to optimize their use of education. Entrepreneurs are not constrained by rules imposed by superiors and can decide themselves how to put their education to its most productive use. The difference in opportunity to optimize the productivity of education for entrepreneurs and employees might therefore explain the higher returns to education for entrepreneurs.

Using changes in compulsory schooling laws as instruments

In Chapter 4 we also estimate the returns to education for USA entrepreneurs in comparison to employees. However, in this chapter we use different instruments for education, namely changes in compulsory schooling laws that vary by state and by year. Due to the lack of sound instruments, we are unable to check for any selection effects into entrepreneurship within this dataset. We rely on the results found in Chapters 3 and 5, where we find that selection is not an issue. We use US Census data from 1950 to 2000 to estimate the returns to education.

We find that the returns to education are 26.4 percent for entrepreneurs and 13.2 percent for employees. For both groups this is substantially higher than the OLS benchmark of 8.3 and 7.3 percent, respectively. The results confirm the difference in returns to education between entrepreneurs and employees we found in Chapter 3.

We also look at the sensitivity of the results to the fact that the dependent variable, i.e. log income, does not include negative incomes for entrepreneurs. The inclusion of negative incomes into the dependent variable shows that omitting these observations, as is common practice, biases the returns to education for entrepreneurs downwards by almost 5 percentage points. Hence the inclusion of negative incomes leads to an even higher estimate of the returns to education.

Comparing the results of both IV-studies

Both Chapters 3 and 4 find substantially higher returns to education for entrepreneurs than for employees. However, the returns to education are higher in Chapter 4 than Chapter 3. This is especially the case for entrepreneurs. The question is where this difference in results between the two studies comes from.

The main difference between Chapters 3 and 4 are the instruments used. However, it is not possible to directly attribute the difference in findings between the two studies to the difference in instruments used. Although both studies are estimated based on USA data, the average age of the individuals in the NLSY sample is much lower than the average age of the individuals in the US Census. Moreover, the time period covered by the NLSY is more recent and shorter. Another important difference is the dependent variable used: In Chapter 3 we use hourly income whereas in Chapter 4 we use weekly income. Finally, the US Census data used in Chapter 4 is made up of repeated cross-sections, whereas Chapter 3 is based on panel data. In cross-section data, as opposed to panel data, entrepreneurs longer in business, who are probably more successful, are over-represented in the sample.

An answer to why the studies differ could be provided by a within-sample comparison of both instruments. Unfortunately, both the NLSY and the US Census do not allow for such a comparison. We leave it to future research to answer this question.

Policy implications from Chapters 3 and 4

Before discussing policy implications, we elaborate on the assumptions we make to translate the estimation results into policy implications.

First, we assume that the development of more entrepreneurship is economically valuable. Second, we assume that the difference between the social and private benefits of entrepreneurial activity is at least as large as for employees. For instance, a successful entrepreneur is more likely to influence market competition positively than an employee. Also, entrepreneurs can introduce new and innovative ideas onto the market more easily than employees. Third, we assume that individuals invest in schooling at a stage in their lives at which they do not yet know, in general, whether they will become entrepreneurs or employees, or a (sequential) combination of both. As a consequence, investment in schooling is not motivated by the specific anticipated returns for entrepreneurs, but by the anticipated returns for both employment modes. Our fourth assumption is that individuals, as well as policy makers, bankers and other parties involved, do not have more insight in the returns to education than we as researchers do. This implies that individuals and policy makers share the common opinion that the returns to education are similar or slightly different, at most, for entrepreneurs and for employees.

Clearly, our finding that entrepreneurial returns to education are high in the USA and that education is therefore a key success factor for starting an enterprise, is relevant for individual labor market decisions, for the development of educational policies, and for bankers' and capital suppliers' strategies with respect to (selecting) starters.

Our finding could motivate the USA government to stimulate higher education for (prospective) entrepreneurs. Alternatively, policy makers could stimulate higher educated individuals to opt for an entrepreneurial career. The first route would increase the likelihood that entrepreneurs will perform better, and that they will generate more benefits that will not only accrue to the entrepreneurs themselves, but to society as a whole also. The second route appeals to the fact that entrepreneurship seems not to be the favored option among highly educated individuals. Both the meta-analysis and the results from the previous chapters indicate an insignificant relation between the choice for entrepreneurship and education level.

We strongly believe in the benefits of governmental programs that stimulate awareness of the attractions of entrepreneurship among college and university students. Future research into the returns to education for entrepreneurs may further increase the effectiveness of such policies.

6.1.3 What are the returns to (different types of) intelligence, on top of the effect of formal education?

Chapter 5 does not look at the relationship between acquired human capital and entrepreneur performance, but deals with the returns to a more innate form of human capital, i.e. intelligence. In particular, the following three questions are answered empirically for both entrepreneurs and employees; (1) To what extent does an individual's general intelligence level affect productivity? (2) Do different areas of intelligence (such as math, clerical, language, technical and social intelligence) affect productivity differently?, and (3) To what extent does the balance in an individual's scores in these areas of intelligence affect an individual's income? The latter question is related to and extends Lazear's Jack-of-all-Trades (JAT) theory pertaining to entrepreneurs. A JAT is someone who is not exceptionally gifted in terms of one specific ability, but someone who has a small dispersion on the test scores for the different abilities.

When comparing entrepreneurs and employees on the basis of their returns to general ability we find that the returns are as high for entrepreneurs as for employees in the USA based on the NLSY data. The effects of the five specific areas of intelligence are not the same for entrepreneurs and employees. As expected, the effect of technical and social intelligence is larger for entrepreneurs than for employees. The effects of verbal and mathematical ability are non-linear and distinct for entrepreneurs versus employees. For both segments of the labor market the returns to verbal intelligence are only positive at the lower end of their distribution, whereas the effect of mathematical ability is only positive at the higher end of the distribution. However, the total value of these two specific abilities is lower for entrepreneurs than for employees.

Based on Lazear's Jack-of-All-Trades theory, we find that the returns to being a JAT are higher for entrepreneurs than for employees. The JAT measure that we use, i.e. the coefficient of variation of an individual's scores on the five specific abilities can be considered a measure of JAT endowment, unlike Lazear's own measures.

Policy implications from Chapter 5

The findings from Chapter 5 present quite a clear handle as to who should and who should not be stimulated to become an entrepreneur in the USA. Attracting people with balanced skill sets, who score higher on both technical and social intelligence, to entrepreneurial positions, would be beneficial for the economy (as well as for the individuals themselves). Providers of equity, loans, subsidies, and licenses -whether private or public authorities-, might use these insights to stimulate a larger and more successful entrepreneurial economy.

6.2 Reflection

Our focus on applying methodological advances in employee literature to entrepreneur literature has certainly generated new insights with respect to the endogeneity of education in an income equation for entrepreneurs. We found that previous estimates of the returns to education can be substantially biased, in particular for entrepreneurs. Moreover, we found that this bias was totally due to the endogeneity of education and not to the selection into entrepreneurship or salaried employment. Although this thesis includes two chapters that deal with this issue, more studies using different instruments for both education and entrepreneurship selection are needed to verify the results.

In this thesis we only use USA data. This is not because we find the USA intrinsically more interesting, but simply because of data availability. Since we used only USA data, the results from this thesis cannot be easily transferred to other countries. Different education systems and institutional settings might have different effects on the performance of entrepreneurs. The meta-analysis indicates that the returns to education for entrepreneurs are higher than for employees in the USA. In Europe the opposite is found. However, as we discussed, these results are based exclusively on the measurement of (conditional) correlations. Hence, (preferably improved) replications of our USA studies on European data should demonstrate the relative magnitude of the returns to education for entrepreneurs vis-à-vis employees in European countries.

Besides issues directly related to the content of this thesis, there are some issues that could broaden the scope of entrepreneurship and educational research. One of these issues is the fact that the returns to education are not guaranteed returns. For each entrepreneur there is uncertainty as to the outcomes associated with his or her education level. Research in the field of entrepreneurship often deals with risk or risk taking as a determinant for the choice to become an entrepreneur. However, the risk associated with the returns to education has not received much attention as yet.

Another issue concerns the effect of different types of education on entrepreneur performance. No study has yet estimated the returns to specific types of education while at the same time taking account of the endogenous nature of these different types of education. Moreover, few researchers have studied the effects of entrepreneurship education. Knowledge about which type of education improves entrepreneur performance the most could help schools design better education programs.

These suggestions for further research by no means present a complete research agenda and many challenges remain in this policy-relevant field of research.

Bibliography

Abreu, M., H. De Groot and R. Florax (2005), 'A meta-analysis of beta-convergence: The legendary 2%', *Journal of Economic Surveys* **19**(3), 389–420.

Acemoglu, D. and J. Angrist (2001), 'How large are human capital externalities? Evidence from compulsory schooling laws', *NBER macroannual* pp. 9–59.

Aldrich, H., N. Carter and M. Ruef (2004), Teams, *in* W. Gartner, K. Shaver, N. Carter and P. Reynolds, eds, 'Handbook of Entrepreneurial Dynamics: The Process of Business Creation', Sage, Thousand Oakes, pp. 299–310.

Angrist, J. and A. Krueger (1991), 'Does compulsory school attendance affect schooling and earnings?', *Quarterly Journal of Economics* **106**(4), 979–1014.

Ashenfelter, O. and A. Krueger (1994), 'Estimates of the returns to schooling from a new sample of twins', *American Economic Review* **84**(5), 1157–1173.

Ashenfelter, O., C. Harmon and H. Oosterbeek (1999), 'A review of estimates of the schooling/earnings relationship, with tests for publication bias', *Labour Economics* **6**(4), 453–470.

Audretsch, D. and M. Keilbach (2004), 'Entrepreneurship and regional growth: An evolutionary interpretation', *Journal of Evolutionary Economics* **14**, 605–616.

Baron, R. (2000), 'Psychological perspectives on entrepreneurship: Cognitive and social factors in entrepreneurs' success', *Current Directions in Psychological Science* **9**(1), 15–18.

Baron, R. and G. Markman (2003), 'Beyond social capital: The role of entrepreneurs' social competence in their financial success', *Journal of Business Venturing* **18**(1), 41–60.

Binet, A. and T. Simon (1911), *A method of measuring the development of the intelligence og young children*, Lincon, II: Courier.

Bishop, J. (1991), Achievement, test scores, and relative wages, *in* M. Kosters, ed., 'Workers and their wages: Changing Patterns in the United States', Washington, D.C.: Brookings Institution.

Black, S., P. Devereux and K. Salvanes (2005), 'The more the merrier? The effect of family composition on children's education', *Quarterly Journal of Economics* **120**(2), 669–700.

Blackburn, M. and D. Neumark (1993), 'Omitted-ability bias and the increase in the return to schooling', *Journal of Labor Economics* **11**(3), 521–544.

Bonjour, D., L. Cherkas, J. Haskel, D. Hawkes and T. Spector (2003), 'Returns to education: Evidence from U.K. twins', *American Economic Review* **93**(5), 1799–1812.

Boone, C. and B. De Brabander (1993), 'Generalized versus specific locus of control expectancies of chief executive officers', *Strategic Manangement Journal* **14**, 619–626.

Boone, C., W. Van Olffen and A. Van Witteloostuijn (2005), 'Team locus-of-control composition, leadership structure, information acquisition, and financial performance: A business simulation study', *Acadamy of Management Journal* **48**, 889–909.

Borjas, G. and S. Bronars (1989), 'Consumer discrimination and self-employment', *Journal of Political Economy* **97**(3), 581–605.

Bound, J., D. Jaeger and R. Baker (1995), 'Problems with instrumental variables instrumentation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of American Statistical Association* **90**(430), 443–450.

Bound, J. and G. Solon (1999), 'Double trouble: On the value of twins-based estimation of the return to schooling', *Economics of Education Review* **18**(2), 169–182.

Brown, S. and J. Sessions (1998), 'Education, employment status and earnings: A comparative test of the strong screening hypothesis', *Scottisch Journal of Political Economy* **45**(5), 586–591.

Brown, S. and J. Sessions (1999), 'Education and employment status: A test of the strong screening hypothesis in Italy', *Economics of Education Review* **18**(4), 397–404.

Calvo, G. and S. Wellisz (1980), 'Technology, entrepreneurs and firm size', *Quarterly Journal of Economics* **95**, 663–678.

Card, D. (1998), 'The causal effect of education', *Working Paper, University of California, Berkley* .

Card, D. (1999), Causal effect of education on earnings, *in* O. Ashenfelter and D. Card, eds, '*Handbook of Labor Economics*', Vol. 3A, North-Holland, pp. 1801–1863.

Card, D. (2001), 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica* **69**(5), 1127–1160.

Card, D. and A. Krueger (1995), 'Time-series minimum-wage studies - a meta-analysis', *American Economic Review* **85**(2), 238–243.

Carrol, J. (1993), *Human cognitive abilities: A survey of factor-analytic studies*, Cambridge University Press.

Carson, C. (1984), 'The underground economy: An introduction', *Survey of Current Business* **64**, 21–37.

Clark, K. and S. Drinkwater (1998), 'Ethnicity and self-employment in Britain', *Oxford Bulletin of Economics and Statistics* **60**(3), 383–407.

Cruz, L. and M. Moreira (2005), 'On the validity of econometric techniques with weak instruments', *Journal of Human Resources* **40**(2), 393–410.

Darwin, C. (1859), *The origin of species, by means of natural selection*, John Murray, London.

De Wit, G. and F. Van Winden (1989), 'An empirical analysis of self-employment in the Netherlands', *Small Business Economics* **1**(4), 263–272.

Deaton, A. (2000), *The Analysis of Houshold Surveys*, Johns Hopkins University Press.

Devine, T. (1995), 'CPS earnings data for the self-employed: Words of caution for use and interpretation', *Journal of Economic and Social Measurement* **21**, 213–248.

Dolton, P. and G. Makepeace (1990), 'Self-employment among graduates', *Bulletin of Economic Research* **42**(1), 35–53.

Dougherty, C. (2000), 'Numeracy, literacy, and earnings: Evidence from the national longitudinal survey', *Centre for Economic Performance: London* .

Drakopoulou Dodd, S. and P. Seaman (1998), 'Religion and enterprise: An introductory exploration', *Entrepreneurship: Theory and Practice* **23**(1), 71–86.

English, H. and A. English (1958), *A comprehensive dictionary of psychological and psychoanalytical terms: A guide to usage*, New York: McKay.

Evans, D. and L. Leighton (1990), 'Small business formation by unemployed an employed workers', *Small Business Economics* **2**(4), 319–330.

Fairlie, R. (2005*a*), 'Entrepreneurship and earnings among young adults from disadvantaged families', *Small Business Economics* **25**(3), 223–236.

Fairlie, R. (2005*b*), 'Self-employment, entrepreneurship, and the NLSY79', *Monthly Labor Review* **128**(2), 40–47.

Fairlie, R. and A. Robb (2007), 'Families, human capital, and small business: Evidence from the Characteristics of Business Owners Survey', *Industrial and Labor Relations Review* **60**(2), Forthcoming.

Fairlie, R. and B. Meyer (1996), 'Ethnic and racial self-employment differences and possible explanations', *Journal of Human Resources* **31**(4), 757–793.

Ferris, G., W. Hochwachter and L. Witt (2001), 'Interaction of social skill and general mental ability on job performance and salary', *Journal of Applied Psychology* **86**(6), 1075–1082.

Fredland, J. and R. Little (1981), 'Self-employed workers: Returns to education and training', *Economics of Education Review* **1**(3), 315–337.

Frey, M. and D. Detterman (2003), 'Scholastic assesment or g?', *Psychological Science* **15**(6), 373–378.

Galton, F. (1869), *Hereditary genius: An inquiry into its laws and consequences*, Macmillan/Fontana, London.

Garcia Mainar, I. and V. Montuenga Gómez (2005), 'Education returns of wage earners and self-employed workers: Portugal vs. Spain', *Economics of Education Review* **24**(161-170).

Gill, A. (1988), 'Choice of employment status and the wages of employees and the self-employed: Some further evidence', *Journal of Applied Econometrics* **3**(3), 229–234.

Goldsmith, A., J. Veum and W. Darity Jr. (1997), 'The impact of psychological and human capital on wages', *Economic Inquiry* **35**, 815–829.

Griliches, Z. (1977), 'Estimating the returns to schooling: Some econometric problems', *Econometrica* **45**(1), 1–22.

Guion, R. (1998), *Assessment, measurement and prediction for personnel decisions*, Lawrence Erlbaum Associates, London.

Hamilton, B. (2000), 'Does entrepreneurship pay? An empirical analysis of the returns to self-employment', *Journal of Political Economy* **108**(3), 604–631.

Hansen, K., J. Heckman and K. Mullen (2004), 'The effect of schooling and ability on achievement test scores', *Journal of Econometrics* **121**(1-2), 39–98.

Harmon, C. and I. Walker (1995), 'Estimates of the economic returns to schooling for the United Kingdom', *American Economic Review* **85**(5), 1278–1286.

Hartog, J. (1980), 'Earnings and capability requirements', *Review of Economics and Statistics* **62**(2), 230–240.

Hartog, J. (1992), *Capabilities, allocation and earnings*, Dordrecht, Kluwer Academic Publishers.

Hartog, J. (2001), 'On human capital and individual capabilities', *Review of Income and Wealth* **47**(4), 516–540.

Hartog, J., J. Van der Sluis and M. Van Praag (2006), 'Returns to intelligence: Entrepreneurs versus employees'. *Working Paper, University of Amsterdam.*

Hause, J. (1972), 'Earnings profile; ability and schooling', *Journal of Political Economy* **80**(3), S108–S138.

Hausman, J. (1978), 'Specification tests in econometrics', *Econometrica* **46**(6), 1251–1271.

Hausman, J. and W. Taylor (1981), 'Panel data and unobservable individual effects', *Econometrica* (49), 1377–1398.

Hedges, L. (1992), 'Modeling publication effects in meta-analysis', *Statistical Science* **7**(2), 246–255.

Hernstein, R. and C. Murray (1994), *The Bell Curve*, New York: The Free Press.

Hmieleski, K. and M. Ensley (2004), 'An investigation of the linkage between entrepreneur intelligence and new venture performance', *Working paper, Rensselaer Polythechnic Institute* .

Iversen, J. (2004), 'Returns to schooling for wage-employed and self-employed: Is there a difference?', *CEBR 2004-04* .

Laferrère, A. (2001), 'Self-employment and intergenerational transfers', *International Journal of Sociology* **31**(1), 3–26.

Lazear, E. (2002), 'Entrepreneurship', *NBER Working paper 9109* .

Lazear, E. (2004), 'Balanced skills and entrepreneurship', *American Economic Review* **94**(2), 208–211.

Lazear, E. (2005), 'Entrepreneurship', *Journal of Labor Economics* **23**(4), 649–680.

Le, A. (1999), 'Empirical studies of self-employment', *Journal of Economic Surveys* **13**(4), 381–416.

Lee, S. (2005), 'Generalists versus specialists, ability and earnings', *Working Paper 200502, University of Hawaii at Manoa* .

Lleras-Muney, A. (2002), 'Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939', *Journal of Law and Economics* **45**, 401–435.

Lochner, L. and E. Moretti (2004), 'The effect of education on crime: Evidence from prison inmates, arrests, and self-reports', *American Economic Review* **94**(1), 155–198.

Lofstrom, M. (2002), 'Labor market assimilation and the self-employment decision of immigrant entrepreneurs', *Journal of Population Economics* **15**(1), 83–114.

Lombard, K. (2001), 'Female self-employment and demand for flexible, nonstandard work schedules', *Economic Inquiry* **39**(2), 214–237.

Longhi, S., P. Nijkamp and J. Poot (2005), 'A meta-analytical assessment of the effect of immigration on wages', *Journal of Economic Surveys* **19**(3), 451–477.

Lucas, R. (1978), 'On the size distribution of business firms', *Bell Journal of Economics* **9**(2), 508–523.

Lyssiotou, P., P. Pashardes and T. Stengos (2004), 'Estimates of the black economy based on consumer demand approaches', *Economic Journal* **114**(497), 622–640.

MacPherson, D. (1988), 'Self-employment and married women', *Economic Letters* **28**(3), 281–284.

Marshall, A. (1890;1930), *Principles of economics*, Macmillan and co. London (first edition 1890).

McIntosh, S. and A. Vignoles (2001), 'Measuring and assessing the impact of basic skills on labour market outcomes', *Oxford Economic Papers* **53**(3), 453–481.

Mincer, J. (1970), 'The distribution of labor income', *Journal of Economic Literature* **8**(1), 1–26.

Moore, R. (1983), 'Employer discrimination: Evidence from self-employed workers', *Review of Economics and Statistics* **65**(3), 496–501.

Murnane, R., J. Willett and F. Levy (1995), 'The growing importance of cognitive skills in wage determination', *Review of Economics and Statistics* **77**(2), 251–266.

Murnane, R., J. Willett, M. Braatz and Y. Duhaldeborde (2001), 'Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY', *Education Economic Review* **20**(4), 311–320.

Neal, D. and W. Johnson (1996), 'The role of premarket factors in black-white wage differences', *Journal of Political Economy* **104**(5), 869–895.

Nijkamp, P. and J. Poot (2004), 'Meta-analysis of the impact of fiscal policies on long-run growth', *European Journal of Political Economy* **20**(1), 91–124.

Nijman, T. and M. Verbeek (1992), 'Nonresponse in panel data: The impact on estimates of a life cycle consumption function', *Journal of Applied Econometrics* **7**(3), 243–257.

Oreopoulos, P. (2003), 'Do dropouts dropout too soon? International evidence from changes in school-leaving laws', *NBER Working paper 10155* .

Oreopoulos, P. (2005), 'Do dropouts drop out to soon? Wealth, health and happiness from compulsory schooling', *Working paper, University of Toronto* .

Oreopoulos, P. (2006), 'Estimating average and local average treatment effects of education when compulsory schooling laws really matter', *American Economic Review* **96**(1), 152–175.

Paglin, M. and A. Rufolo (1990), 'Heterogeneous human capital, occupational choice, and male-female earnings differences', *Journal of Labor Economics* **8**(1), 123–144.

Park, J. (1994), 'Estimation of sheepskin effects and returns to schooling using the old and new CPS measures of educational attainment'. *Working Paper, Princeton Department of Economics - Industrial Relations Sections no. 338*.

Parker, S. (2004), *The economics of self-employement and entrepreneurship*, Cambridge, UK.

Parker, S. and M. Van Praag (2006), 'Schooling, capital constraints and entrepreneurial performance: The endogenous triangle', *Journal of Business and Economic Statistics* **24**(4), 416–431.

Penrose, E. (1959), *The theory of the growth of the firm*, Oxford Basil Blackwell.

Phillips, J. (1994), 'Farmer education and farm efficiency', *Economic Development and Change* **43**(1), 149–165.

Pischke, B. and T. Von Wachter (2005), 'Zero returns to compulsory schooling in Germany: Evidence and interpretation'. *IZA Discussion Paper no. 1645*.

Plug, E. and J. Levin (1999), 'Instrumenting education and the returns to schooling in the Netherlands', *Labour Economics* **6**(4), 521–534.

Psacharopoulos, G. (1994), 'Returns to investment in education: A global update', *World Development* **22**, 1325–1343.

Ree, M. and J. Earles (1992), 'Intelligence is the best predictor of job performance', *Current Directions in Psychological Science* **1**(3), 86–89.

Rees, H. and A. Shah (1986), 'An empirical analysis of self-employment in the U.K.', *Journal of Applied Econometrics* **1**(1), 95–108.

Riley, J. (2002), 'Weak and strong signals', *Scandinavian Journal of Economics* **104**(2), 213–236.

Roberts, C. (2005), 'Issues in meta-regression analysis: An overview', *Journal of Economic Surveys* **19**(3), 295–298.

Robinson, P. and E. Sexton (1994), 'The effect of education and experience on self-employment success', *Journal of Business Venturing* **9**(2), 141–156.

Rodgers, J., H. Cleveland, E. Van den Oord and D. Rowe (2000), 'Resolving the debate over birth order, family size, and intelligence', *American Psychologist* **55**(6), 599–612.

Rotter, J. (1966), 'Generalized expectancies for internal versus external control of reinforcement', *Psychological Monographs* **80**(1), 1–28.

Ruggles, S., M. Sobek, T. Alexander, C. Fitch, R. Goeken, P. Kelly Hall, M. King and C. Ronnander (2004), *Integrated Public Use Microdata Series: Version 3.0*, Minnesota Population Center [Producer and Distributer], Minneapolis, MN.

Sargan, J. (1988), Testing for misspecification after estimating using instrumental variables, *in* E. Maasoumi, ed., 'Contributions to econometrics', Vol. 1 and 2, Cambridge: Cambridge University Press.

Schmidt, F. and J. Hunter (1992), 'Development of a causal model of process determining job performance', *Current Directions in Psychological Science* **1**(3), 89–92.

Schmidt, F. and J. Hunter (1998), 'The validity and utility of selection methods in personel psychology: Practical and theoretical implications of 85 years of research findings', *Psychological Bulletin* **124**(2), 262–274.

Schmidt, F. and J. Hunter (2004), 'General mental ability in the world of work: Occupational attainment and job performance', *Journal of Personality and Social Psychology* **86**(1), 162–173.

Schmidt, F., J. Hunter and A. Outerbridge (1986), 'The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance', *Journal of Applied Psychology* **71**, 432–439.

Schmidt, S. (1996), 'School quality, compulsory education laws, and the growth of American high school attendance 1915-1935', *Ph.D. dissertation. Cambridge: Massachusetts Institute of Technology* .

Shane, S. (2003), *A general theory of entrepreneurship: The individual-opportunity nexus*, Cheltenham: UK, Edward Elgar.

Silva, O. (2007), 'The Jack-of-All-Trades entrepreneur: Innate talent or acquired skill?'. *Economics Letters, Forthcomming.*

Spearman, C. (1904), 'General intelligence, objectively determined and measured', *American Journal of Psychology* **15**, 201–293.

Spence, A. (1973), *Market Signaling: Information Transfer in Hiring and Related Processes*, Cambridge, MA.

Stanley, T. (2005), 'Beyond publication bias', *Journal of Economic Surveys* **19**(3), 309–345.

Stanley, T. and S. Jarrell (2005), 'Meta-regression analysis: A quantitative method of literature surveys', *Journal of Economic Surveys* **19**(3), 299–308.

Taubman, P. and T. Wales (1974), 'Higher education and earnings: College education as an investment and a screening device'. *NBER General Series 101, New York: McGraw-Hill.*

Taylor, M. (1996), 'Earnings, independence, or unemployment: Why become self-employed?', *Oxford Bulletin of Economics and Statistics* **58**(2), 253–266.

Taylor, M. (1999), 'Survival of the fittest? An analysis of self-employment duration in Britain', *Economic Journal* **109**(454), c140–c155.

Taylor, M. (2001), 'Self-employment and windfall gains in Britain: Evidence from panel data', *Economica* **68**, 539–565.

Thorndike, E. (1904), *An introduction to the theory of mental and social measurements*, New York: Teachers College, Columbia University.

Thurstone, L. and T. Thurstone (1941), *Factorial Studies of Intelligence*, Vol. Psychometric Monographs No.2, Chicago: University of Chicago Press.

Tucker, I. (1985), 'Use of the decompensation technique to test the educational screening hypothesis', *Economics of Education Review* **4**(4), 321–326.

Tucker, I. (1987), 'The impact of consumer credentialism on employee and entrepreneur returns to higher education', *Economics of Education Review* **6**(1), 35–40.

Van der Sluis, J. and M. Van Praag (2004), 'Economic returns to education for entrepreneurs: The development of a neglected child in the economics of education', *Swedish Economic Policy Review* **11**(2), 183–225.

Van der Sluis, J. and M. Van Praag (2006), 'Returns to education for entrepreneurs and employees: Identification by means of changes in compulsory schooling laws'. *Working Paper, University of Amsterdam.*

Van der Sluis, J. and M. Van Praag (2007), Human capital and entrepreneurs, *in* J. Hartog and H. Maassen van den Brink, eds, 'Human Capital: Theory and Evidence', Cambridge University Press, Forthcomming.

Van der Sluis, J., M. Van Praag and A. Van Witteloostuijn (2005), 'Why are the returns to education higher for entrepreneurs than for employees?'. *Working Paper, University of Amsterdam.*

Van der Sluis, J., M. Van Praag and W. Vijverberg (2003), 'Entrepreneurship, selection and performance: A meta-analysis of the role of education'. *Working Paper, University of Amsterdam.*

Van der Sluis, J., M. Van Praag and W. Vijverberg (2005), 'Entrepreneurship selection and performance: A meta-analysis of the impact of education in developing economies', *World Bank Economic Review* **19**(2), 225–261.

Van Praag, M. (1999), 'Some classic views on entrepreneurship', *De Economist* **147**(3), 311–335.

Van Praag, M. (2003), 'Business survival and success of young small business owners: An empirical analysis', *Small Business Economics* **21**(1), 1–17.

Van Praag, M. (2006), Venture performance and venture input: Human and financial capital, *in* S. Parker, ed., '*The Life Cycle of Entrepreneurial Ventures, The International Handbook of Entrepreneurship*', Springer Academic Publishers, pp. 507–533.

Van Praag, M. and H. Van Ophem (1995), 'Determinants of willingness and opportunity to start as an entrepreneur', *Kyklos* **48**(4), 513–540.

Van Praag, M. and J. Cramer (2001), 'The roots of entrepreneurship and labour demand: Individual ability and low risk aversion', *Economica* **269**, 45–62.

Van Stel, A. and D. Storey (2004), 'The link between firm births an job creation: Is there a Upas Tree effect?', *Regional Studies* **38**(8), 893–909.

Wagner, J. (2003), 'Testing Lazear's jack-of-all-trades view of entrepreneurship with German micro data', *Applied Economics Letters* **10**(11), 687–689.

Webbink, D. (2005), 'Causal effects in education', *Journal of Economic Surveys* **19**(4), 535–560.

Willis, R. and S. Rosen (1979), 'Education and self-selection', *Journal of Political Economy* **87**(5), S7–S36.

Wolpin, K. (1977), 'Education and screening', *American Economic Review* **67**(5), 949–958.

Wong, C. and K. Law (2002), 'the effect of leader and follower emotional intelligence on performance and attitude: An exploratory study', *Leadership Quarterly* **13**(3), 243–274.

Zax, J. and D. Rees (2002), 'IQ, academic performance, environment, and earnings', *Review of Economics and Statistics* **84**(4), 600–616.

Zetterberg, J. (2005), 'Swedish evidence on the impact of cognitive and non-cognitive ability on earnings: An extended pre-market factor approach'. *IFAU working paper 2005:16.*

Samenvatting

(Summary in Dutch)

De bijdrage van ondernemers aan de economie is substantieel. Onderzoek toont aan dat ondernemerschap relatief veel bijdraagt aan economische groei en dynamiek en aan de creatie van werkgelegenheid. Ondernemers leveren ook een grote bijdrage aan de innovatiekracht van de maatschappij. Het EU Lissabon akkoord van 2000 roept op tot een innovatiever Europa. In veel landen wordt meer ondernemerschap dan ook gezien als de weg om dit doel te bereiken.

Het is echter alleen mogelijk om te profiteren van de voordelen van ondernemerschap als ondernemers succesvol zijn. Maar welke factoren bepalen het succes van een ondernemer? Dit boek richt zich op één mogelijke determinant van dit succes, namelijk menselijk kapitaal. Om precies te zijn, meten we de procentuele meeropbrengsten in termen van inkomen van een extra jaar formele scholing voor ondernemers. Kortweg, de rendementen op scholing voor ondernemers. De motivatie hiervoor is als volgt. Ten eerste lijkt scholing op basis van eerder onderzoek één van de factoren te zijn die de prestaties van ondernemers kan beïnvloeden. Ten tweede, mocht scholing op een bepaalde manier bijdragen aan succesvol ondernemerschap dan kunnen op basis daarvan ook instrumenten worden ontworpen en beleidsaanbevelingen worden gedaan. De hoeveelheid en soort opleiding die iemand volgt is immers een individuele keuze en het onderwijsaanbod wordt voor een groot deel bepaald door overheidsbeleid.

Het meten van het effect van scholing op de prestaties van *werknemers* is onderwerp van vele studies geweest. Vooral het meten van het causale effect van scholing op inkomen kreeg in deze studies aandacht. Er zijn veel factoren die onderzoekers niet kunnen observeren, zoals bijvoorbeeld motivatie, die én het scholingsniveau van een individu bepalen én het inkomen van een individu. Ook zullen personen die hogere opbrengsten denken te kunnen genereren op hun investering in onderwijs, meer investeren dan anderen. Het aantal jaren scholing dat iemand volgt wordt niet door een dobbelsteen bepaald en de (voor de onderzoeker vaak ongeobserveerde) factoren die erop van invloed zijn, beïnvloeden ook vaak weer iemands inkomsten. Kortom, scholing is een endogene factor die van invloed is op de hoogte van inkomens. Om die reden leveren eenvoudige econometrische analyses ('kleinste kwadraten-schatters') een onzuivere schatting op van de omvang van het causale effect van scholing op inkomen. Dergelijke correlaties kunnen dan ook niet als basis worden gebruikt voor beleidsaanbevelingen. Daarom zijn er verschillende schattingtechnieken ontwikkeld om causale effecten te meten. Het rendement op scholing van *werknemers* blijkt door het gebruik van deze schattingstechnieken rond de 10 procent te liggen. Het niet gebruiken van deze schattingstechnieken leidt tot een geschat rendement van rond de 6 procent, dat wil zeggen een onderschatting van het werkelijke rendement.

Voor ondernemers is het rendement op scholing helaas niet zo goed gemeten. De methoden die zijn gebruikt voor ondernemers geven geen causaal verband aan, maar alleen een correlatie. De schattingstechnieken die gebruikt zijn voor het meten van scholingsrendementen voor werknemers kunnen natuurlijk ook gebruikt worden voor het meten van het rendement op scholing voor ondernemers. Dat is nog niet eerder gebeurd. Pas als vergelijkbare metingen zijn verricht kunnen ook de (beleidsrelevante) verschillen tussen het rendement op scholing voor ondernemers en werknemers in kaart worden gebracht. Veranderingen in het scholingssysteem beïnvloeden de keuzes en rendementen van beide groepen, en wellicht op verschillende manieren.

In dit boek wordt antwoord gegeven op de volgende drie onderzoeksvragen:

1. Wat is de huidige stand van zaken op het gebied van het meten van de rendementen op (jaren formele) scholing voor ondernemers?

2. Hoe hoog zijn de rendementen op scholing voor ondernemers in vergelijking tot werknemers?

3. Wat zijn de rendementen op (verschillende typen) intelligentie, bovenop de rendementen op formele scholing?

Wat is de huidige stand van zaken op het gebied van het meten van de rendementen op (jaren formele) scholing voor ondernemers?

Hoofdstuk 2 geeft een overzicht van empirisch onderzoek dat met name in de laatste twee decennia is verricht naar de relatie tussen scholing enerzijds en de keuze om ondernemer te worden en de prestaties van ondernemers anderzijds. Op basis van bijna 100 studies vat dit hoofdstuk de gevonden resultaten samen. Deze meta-analyse levert vijf conclusies op.

Allereerst blijkt dat iemands keuze om wel of niet ondernemer te worden geen verband houdt met het aantal jaren opleiding van de persoon. Een tweede resultaat is dat er een positief verband wordt gevonden tussen de scholing en de prestaties van ondernemers. Dit geldt voor verschillende typen prestaties: niet alleen het inkomen van de ondernemer, maar ook diens overlevingskansen in de markt, groei en satisfactie zijn alle gemiddeld hoger voor ondernemers met een hogere opleiding. Ten derde blijkt dat de als correlatie gemeten rendementen op scholing voor ondernemers (die dus niet een causaal verband weergeven) gemiddeld 6.1% bedragen. Ten vierde blijkt dat voor ondernemers in de Verenigde Staten het verband tussen scholing en inkomsten sterker is dan voor werknemers, terwijl dit in Europa precies andersom is. Ten vijfde, lijkt er een indicatie te zijn dat het verband tussen scholing en inkomsten van een ondernemer sterker is in de Verenigde Staten dan in Europa.

Wat opvalt is dat alle studies die de basis vormen voor de meta-analyse het rendement op scholing meten als correlatie en niet als causaal effect zoals veel gebruikelijker is in de studies die de rendementen op onderwijs voor *werknemers* meten. Zoals reeds gezegd, is het goed mogelijk dat deze correlatie niet een causaal verband tussen scholing en ondernemersprestaties weergeeft. Mogelijk wordt het rendement op scholing zelfs onderschat, zoals ook bij werknemers het geval was. Een andere mogelijke methodologische verbetering van bestaande studies is gebaseerd op het gegeven dat maar in enkele studies die onderdeel uitmaken van de meta-analyse rekening gehouden wordt met het feit dat de keuze om ondernemer te worden ook niet door een dobbelsteen wordt bepaald (en dus ook endogeen is). Het kan zijn dat iemand ondernemer wordt omdat zijn scholingsniveau en een van zijn niet observeerbare kenmerken hogere opbrengsten heeft als ondernemer dan als werknemer. Het effect van de keuze voor ondernemerschap zal dus ook meegenomen moeten worden in de berekening van het rendement op scholing. In hoofdstukken 3 en 4 wordt geprobeerd deze meetproblemen zoveel mogelijk op te lossen door het toepassen van methodieken gebruikt door arbeidseconomen die de rendementen op scholing voor werknemers meten.

Hoe hoog zijn de rendementen op scholing voor ondernemers in vergelijking tot werknemers?

Kenmerken ouderlijk huis als instrumenten voor scholing

In hoofdstuk 3 wordt een eerste poging gedaan om rekening te houden met de endogeniteit van scholing en de keuze voor het ondernemerschap of werknemerschap bij het verklaren van inkomens. Hiervoor gebruiken we instrumentele variabelen technieken. Als instrument voor scholing gebruiken we verschillende kenmerken van iemands ouderlijk huis. Als instrument voor de keuze voor ondernemerschap proberen we twee instrumenten uit, d.w.z. de kans dat de iemands vader ondernemer was en iemands religie. We gebruiken een panel-dataset uit de Verenigde Staten (NLSY-dataset) om de rendementen op scholing te schatten.

Als referentiemodel schatten we het rendement op scholing als conditionele correlatie (met de gewone 'kleinste kwadraten'-schattingsmethode). We vinden dat de rendementen van scholing net wat hoger zijn voor ondernemers dan voor werknemers, respectievelijk 7 procent en 6 procent. Dit is in lijn met de bevindingen van de meta-analyse. Maar, als we rekening houden met endogeniteit gaan de schattingen omhoog naar 10 procent voor werknemers en 18 procent voor ondernemers. Voor werknemers is dit resultaat niet verrassend en al eerder gevonden, maar voor ondernemers is dit resultaat wel nieuw. De rendementen op scholing zijn 85 procent hoger voor ondernemers dan voor werknemers. Deze stijging ten opzichte van de conditionele correlatie komt geheel voor rekening van de endogeniteit van scholing, want uit de analyse blijkt dat de endogeniteit van de keuze voor ondernemerschap geen rol speelt.

De vraag is waarom ondernemers meer hebben aan hun opleiding dan werknemers. Op basis van aanvullende analyses die besproken worden in hoofdstuk 3 komt een mogelijke verklaring naar voren. Ondernemers hebben meer vrijheid om hun scholing optimaal in te zetten. Ondernemers worden namelijk niet beperkt door regels van hun superieuren en kunnen zelf uitmaken hoe zij de hoogste productiviteit met hun scholing kunnen behalen. Werknemers hebben deze vrijheid minder. Het verschil tussen ondernemers en werknemers om hun scholing vrijelijk en daardoor rendabeler in te zetten zou een verklaring kunnen zijn voor het verschil in de rendementen op scholing tussen ondernemers en werknemers.

Veranderingen in de leerplichtwet als instrument voor scholing

In hoofdstuk 4 berekenen we net als in hoofdstuk 3 de rendementen op scholing voor ondernemers en werknemers in de Verenigde Staten. Het verschil met hoofdstuk 3 is dat we nu veranderingen in de leerplichtwet als instrument voor scholing gebruiken. Helaas kunnen we in dit hoofdstuk niet voor selectie-effecten m.b.t. de keuze voor ondernemerschap controleren omdat we hier geen geschikte instrumenten voor hebben. We gaan er vanuit dat de resultaten uit 3 en 5, waar we geen bewijs voor de aanwezigheid van een selectie-effect vinden, ook van toepassing zijn voor de in dit hoofdstuk besproken studie. Om de rendementen op scholing te schatten gebruiken we de 'US Census' dataset van 1950 tot 2000.

We vinden dat de rendementen op scholing 26 procent zijn voor ondernemers en 13 procent voor werknemers. Dit is flink hoger dan de rendementen die we vinden als we geen rekening houden met endogeniteit, d.w.z. 8 procent voor ondernemers en 7 procent voor werknemers. De resultaten bevestigen het grote verschil in de rendementen van scholing tussen ondernemers en werknemers dat we vonden in hoofdstuk 3.

In dit hoofdstuk hebben we ook gekeken naar de gevoeligheid van onze resultaten voor veranderingen in de afhankelijke variabele. Om precies te zijn, hebben we gekeken naar de effecten van het wel of niet meenemen van negatieve inkomens voor ondernemers. Als we negatieve inkomens meenemen, blijken de rendementen op scholing voor ondernemers nog eens 5 procentpunten hoger te zijn dan als we negatieve inkomens niet meenemen in de schattingen.

Een vergelijking van de resultaten tussen de twee IV-studies

In zowel hoofdstuk 3 en hoofdstuk 4 vinden we dat de rendementen van scholing hoger zijn voor ondernemers dan voor werknemers. Maar in hoofdstuk 4 zijn de rendementen veel hoger dan in hoofdstuk 3. De vraag is hoe dit komt. Het grootste verschil tussen de twee hoofdstukken ligt in het gebruik van de verschillende instrumenten, d.w.z. kenmerken van iemands ouderlijke achtergrond in hoofdstuk 3 en veranderingen in de leerplichtwet in hoofdstuk 4. Helaas is het niet mogelijk de gevonden verschillen hieraan toe te schrijven. Ook andere aspecten van de twee studies zouden het verschil kunnen verklaren tussen de uitkomsten die beide betrekking hebben op de Verenigde Staten. De gemiddelde leeftijd van de respondenten in de NLSY dataset (hoofdstuk 3) is bijvoorbeeld een stuk lager dan die in de Census dataset (hoofdstuk 4). Ook de bestudeerde periode is veel korter in de NLSY. De Census dataset gaat een aantal decennia langer terug. Een ander verschil is de uitkomstmaat die is gebruikt in beide studies. In de NLSY is dat uurinkomen terwijl in de Census weekinkomen wordt gebruikt. Als laatste verschilt ook nog eens de opbouw van de twee datasets, de NLSY betreft panel data en meet de kenmerken van een bepaalde groep individuen over de tijd. Daarentegen bestaat de Census data uit herhaalde crosssecties (per decennium) van de populatie van de Verenigde Staten. In crosssectie data worden ondernemers die langer ondernemer zijn en waarschijnlijk succesvoller, oververtegenwoordigd. Al met al zijn er dus verschillende redenen aan te wijzen waarom de resultaten tussen de twee studies lastig te vergelijken zijn.

Beleidsimplicaties van hoofdstuk 3 en 4

Voordat we ingaan op de beleidsimplicaties leggen we eerst uit welke veronderstellingen we maken om de resultaten om te zetten in beleidsimplicaties.

Als eerste nemen we aan dat het creëren van meer ondernemerschap van economische waarde is. Ten tweede nemen we aan dat het verschil tussen de maatschappelijke en de privé rendementen op zijn minst even groot is voor ondernemers als voor werknemers. Een succesvolle ondernemer kan bijvoorbeeld veel makkelijker de dynamiek in een markt beïnvloeden dan een werknemer. Tevens kunnen ondernemers veel makkelijker nieuwe en innovatieve producten op de markt brengen. Ten derde nemen we aan dat de investeringsbeslissing voor scholing plaatsvindt op een moment dat een individu nog geen keuze gemaakt heeft voor ondernemerschap of werknemerschap. Ten vierde nemen we aan dat individuen en beleidsmakers er tot nu toe vanuit zijn gegaan dat de rendementen op scholing voor ondernemers niet of nauwelijks verschillen van die voor werknemers. Zoals ook bleek uit de meta-analyse, gaven de resultaten namelijk aan dat het verband tussen scholing en inkomsten voor ondernemers en werknemers ongeveer even groot was.

De in dit boek gepresenteerde bevinding dat de rendementen op scholing hoger zijn voor ondernemers dan voor werknemers in de Verenigde Staten heeft implicaties voor individuele arbeidsmarktbeslissingen, scholingsbeleid en voor het beleid van bankiers en kapitaalverstrekkers die kredieten of eigen vermogen verschaffen aan startende ondernemers.

De resultaten zouden de overheid van de Verenigde Staten kunnen aanmoedigen het volgen van hoger onderwijs onder (toekomstige) ondernemers te stimuleren. Ook zouden hoger opgeleiden kunnen worden aangemoedigd om te gaan ondernemen. Via de eerste route zou de kans dat ondernemers goed presteren kunnen worden vergroot en via de tweede route zou ondernemerschap een serieuze carrière optie voor hoogopgeleiden kunnen worden met de bijbehorende goede prestaties. Zowel de meta-analyse als de twee daarop volgende hoofdstukken laten namelijk een niet significante relatie zien tussen scholing en de keuze voor ondernemerschap, terwijl de rendementen wel drastisch lijken te verschillen. Een overheidsprogramma dat hoger opgeleide studenten bewuster maakt van de positieve kanten van ondernemerschap zou dus gunstige effecten kunnen hebben. Verder onderzoek naar de rendementen op scholing zou de effectiviteit van deze beleidsmaatregel nog eens kunnen vergroten.

Wat zijn de rendementen van (verschillende typen) intelligentie, boven op de rendementen op formele scholing?

Hoofdstuk 5 bestudeert niet, zoals in de vorige hoofdstukken, het effect van aangeleerd menselijk kapitaal op ondernemersprestaties, maar het effect van intelligentie in verschillende soorten en maten op de prestaties van ondernemers. De volgende drie vragen worden empirisch onderzocht voor ondernemers en werknemers op basis van de NLSY data uit de Verenigde Staten: (1) Wat zijn de rendementen op algemene intelligentie?, (2) Hebben verschillende soorten intelligentie (zoals wiskundige, taal, administratieve, technische en sociale intelligentie) verschillende rendementen?, en (3) in welke mate bepaalt de balans tussen de verschillende soorten intelligentie het inkomen van een individu? Het beantwoorden van de laatste vraag vormt een (verbeterde) empirische test van de Jack-of-all-Trades (JAT) theorie, oftewel de 'van alle markten thuis' theorie van Lazear. Een JAT is iemand die niet extreem hoog scoort op een specifiek type intelligentie, maar iemand die juist in veel dingen ongeveer even goed is. Volgends de JAT theorie moeten ondernemers, in tegenstelling tot werknemers, van alle markten thuis zijn en dus een breed scala aan bekwaamheden bezitten.

Voor algemene intelligentie zijn de rendementen voor ondernemers en werknemers niet verschillend. De vijf specifieke typen intelligentie blijken daarentegen wel verschillende rendementen te hebben voor ondernemers en werknemers. Technische en sociale intelligentie leveren meer op voor ondernemers. De rendementen op taal en wiskundige intelligentie zijn niet-lineair en verschillend voor ondernemers en werknemers. Voor beide groepen zijn de rendementen op wiskundige intelligentie alleen positief aan de bovenkant van de verdeling terwijl de rendementen op taal-intelligentie alleen positief zijn aan de onderkant van de verdeling. Voor beide typen intelligentie geldt dat de rendementen lager zijn voor ondernemers dan voor werknemers.

Volgens de JAT theorie hebben JATs een grotere kans om ondernemer te worden omdat zij een breed scala aan bekwaamheden bezitten die nodig zijn voor een goed presterende ondernemer. Een goede manier om te testen of de JAT theorie opgaat in de praktijk, is door te kijken naar het verschil in de rendementen tussen ondernemers en werknemers op het zijn van een JAT. Als (inverse) maat voor het zijn van een JAT gebruiken we de variatiecoëffici ënt van de vijf typen intelligentie. Uit de resultaten blijkt dat de rendementen op het zijn van een JAT significant groter zijn voor ondernemers dan voor werknemers. De resultaten ondersteunen dus de JAT theorie.

Beleidsimplicaties van hoofdstuk 5

Het stimuleren van JATs die hoog scoren op technische en sociale intelligentie om ondernemer te worden, lijkt door de resultaten gerechtvaardigd. Verstrekkers van kapitaal, leningen, subsidies en vergunningen zouden deze inzichten kunnen gebruiken om een grote en meer succesvolle ondernemerseconomie te creëren.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

352. A. GALEOTTI, *On social and economic networks.*
353. Y.C. CHEUNG, *Essays on European bond markets.*
354. A. ULE, *Exclusion and cooperation in networks.*
355. I.S. SCHINDELE, *Three essays on venture capital contracting.*
356. C.M. VAN DER HEIDE, *An economic analysis of nature policy.*
357. Y. HU, *Essays on labour economics: Empirical studies on wage differentials across categories of working hours, employment contracts, gender and cohorts.*
358. S. LONGHI, *Open regional labour markets and socio-economic developments: Studies on adjustment and spatial interaction.*
359. K.J. BENIERS, *The quality of political decision making: Information and motivation.*
360. R.J.A. LAEVEN, *Essays on risk measures and stochastic dependence: With applications to insurance and finance.*
361. N. VAN HOREN, *Economic effects of financial integration for developing countries.*
362. J.J.A. KAMPHORST, *Networks and learning.*
363. E. PORRAS MUSALEM, *Inventory theory in practice: Joint replenishments and spare parts control.*
364. M. ABREU, *Spatial determinants of economic growth and technology diffusion.*
365. S.M. BAJDECHI-RAITA, *The risk of investment in human capital.*
366. A.P.C. VAN DER PLOEG, *Stochastic volatility and the pricing of financial derivatives.*
367. R. VAN DER KRUK, *Hedonic valuation of Dutch Wetlands.*
368. P. WRASAI, *Agency problems in political decision making.*
369. B.K. BIERUT, *Essays on the making and implementation of monetary policy decisions.*
370. E. REUBEN, *Fairness in the lab: The effects of norm enforcement in economic decisions.*
371. G.J.M. LINDERS, *Intangible barriers to trade: The impact of institutions, culture, and distance on patterns of trade.*
372. A. HOPFENSITZ, *The role of affect in reciprocity and risk taking: Experimental studies of economic behavior.*
373. R.A. SPARROW, *Health, education and economic crisis: Protecting the poor in Indonesia.*
374. M.J. KOETSE, *Determinants of investment behaviour: Methods and applications of meta-analysis.*

375. G. MÜLLER, *On the role of personality traits and social skills in adult economic attainment.*
376. E.H.B. FEIJEN, *The influence of powerful firms on financial markets.*
377. J.W. GROSSER, *Voting in the laboratory.*
378. M.R.E. BRONS, *Meta-analytical studies in transport economics: Methodology and applications.*
379. L.F. HOOGERHEIDE, *Essays on neural network sampling methods and instrumental variables.*
380. M. DE GRAAF-ZIJL, *Economic and social consequences of temporary employment.*
381. O.A.C. VAN HEMERT, *Dynamic investor decisions.*
382. Z. ŠAŠOVOVÁ, *Liking and disliking: The dynamic effects of social networks during a large-scale information system implementation.*
383. P. RODENBURG, *The construction of instruments for measuring unemployment.*
384. M.J. VAN DER LEIJ, *The economics of networks: Theory and empirics.*
385. R. VAN DER NOLL, *Essays on internet and information economics.*
386. V. PANCHENKO; *Nonparametric methods in economics and finance: dependence, causality and prediction.*
387. C.A.S.P. SÁ, *Higher education choice in The Netherlands: The economics of where to go.*
388. J. DELFGAAUW, *Wonderful and woeful work: Incentives, selection, turnover, and workers' motivation.*
389. G. DEBREZION, *Railway impacts on real estate prices.*
390. A.V. HARDIYANTO, *Time series studies on Indonesian rupiah/USD rate 1995 - 2005.*
391. M.I.S.H. MUNANDAR, *Essays on economic integration.*
392. K.G. BERDEN, *On technology, uncertainty and economic growth.*
393. G. VAN DE KUILEN, *The economic measurement of psychological risk attitudes.*
394. E.A. MOOI, *Inter-organizational cooperation, conflict, and change.*
395. A. LLENA NOZAL, *On the dynamics of health, work and socioeconomic status.*
396. P.D.E. DINDO, *Bounded rationality and heterogeneity in economic dynamic models.*
397. D.F. SCHRAGER, *Essays on asset liability modeling.*
398. R. HUANG, *Three essays on the effects of banking regulations.*
399. C.M. VAN MOURIK, *Globalisation and the role of financial accounting information in Japan.*
400. S.M.S.N. MAXIMIANO, *Essays in organizational economics.*
401. W. JANSSENS, *Social capital and cooperation: An impact evaluation of a women's empowerment programme in rural India.*