TI 2004-104/3
Tinbergen Institute Discussion Paper

Comparing the Returns to Education for Entrepreneurs and Employees

Justin van der Sluis[1]
Mirjam van Praag[1, 2]
Arjen van Witteloostuijn[3]

[1] *Faculty of Economics and Econometrics, University of Amsterdam, and SCHOLAR,*
[2] *Tinbergen Institute,*
[3] *University of Groningen, University of Durham.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Amsterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Comparing the returns to education for entrepreneurs and employees*

Justin van der Sluis†
University of Amsterdam
Mirjam van Praag
University of Amsterdam
Arjen van Witteloostuijn
University of Groningen & University of Durham

Abstract

The returns to education (RTE) for entrepreneurs, unlike for employees, have not yet been estimated by methods coping with the potential endogeneity problem. We estimate the RTE for entrepreneurs and employees while testing for and coping with this problem. Our results derived from a large US sample strongly indicate that OLS renders downward biased RTE's for entrepreneurs and employees. The bias is larger for entrepreneurs. This leads to an interesting and policy relevant result: The RTE are higher for entrepreneurs than for employees (14% and 10%, respectively), whereas previous estimates suggested similar RTE's. Tests indicate that this result is robust.

*We would like to give special thanks to Rob Alessie for some excellent suggestions. The paper has further benefited from discussions at the Annual Meeting of the Economic Council of Sweden, the 2004 Babson Kaufman conference in Glasgow, as well as from a seminar at ESADE Business School, Barcelona.

†Correspondence author: University of Amsterdam / Scholar, Roetersstraat 11, 1018 WB Amsterdam, room E11.07, j.vandersluis@uva.nl.

1 Introduction

Policy makers and academic researchers are increasingly aware of the importance of entrepreneurship in our society. Entrepreneurs are often seen as the engine of the economy, responsible for sustained levels of competition, the creation of jobs, and new innovating processes and products. These benefits, that accrue to society at large, justify public expenditures in the development of entrepreneurship. To stimulate the further development of entrepreneurs, policy makers have implemented entrepreneurship stimulation programs on schools, and have made several subsidies and support services available for start-ups and small firms.

However, the question remains in which way optimal stimulation of entrepreneurial performance can be achieved. There is one factor that both academic scholars and policy makers see as an important determinant of entrepreneurial performance, namely human capital. We think that the correct measurement of the magnitude of the returns to human capital is therefore of utmost importance to device and implement effective policies. However, a recent meta-analysis (Van der Sluis, Van Praag & Vijverberg 2003) reveals that the effect of formal schooling, one of the most prominent manifestations of human capital, on entrepreneurial performance has not yet been measured consistently. This is due to shortcomings in the empirical strategies applied so far. Previous studies measuring the relationship between education and entrepreneurial performance merely measure (conditional) correlations rather than causal effects. No attempt has yet been made to apply identification strategies such as Instrumental Variables, twins studies and the like in order to estimate causal effects that are not biased due to the neglect of unob-

served heterogeneity and the endogenous nature of the decision to invest in schooling.

Our aim is to measure the returns to education for entrepreneurs and relate its magnitude to the returns to education for employees. To this end, we will estimate the returns to education for both entrepreneurs and employees. Using the same methodology for both samples will allow us to compare the size of the omitted ability biases, the magnitude of the returns to education, and the importance of sample selection for both groups. To investigate the effects of omitted ability and endogenous education, we include a set of detailed ability proxies into our regression equations that we estimate by means of a random-effects model applying an Instrumental Variable (IV) approach. We use the US National Longitudinal Survey Youth cohort (NLSY) as our sample.

Our results reveal that the returns to education are higher for entrepreneurs than for employees, and that applying an Instrumental Variables approach widens the difference between the two. The rather large difference we find can not be attributed to sample selection or any of the other alternative explanations we address when performing robustness checks. We therefore arrive at interesting policy recommendations that follow from the estimation results under quite broad assumptions.

The paper is organized in the following way. Section 2 provides an overview of the literature of the returns to education for entrepreneurs. It also describes the methods used for obtaining unbiased estimates of the returns to education for employees. Section 3 describes the sample used and the methodology applied. In Section 4, we estimate the effect of education

on the incomes of both entrepreneurs and employees, using the same variables and the same estimation techniques. Moreover, we perform several robustness checks. Finally, results are discussed, and conclusions and policy recommendations are provided in Section 5.

2 Empirical literature

The relationship between schooling and entrepreneurship entry and performance has been measured in various empirical studies. Van der Sluis, Van Praag and Vijverberg (2003) ,VVV 2003 hereafter) provide an overview of such empirical studies into the impact of schooling on entrepreneurship selection and performance. They perform a meta-analysis in order to assess whether there are any consistent findings from the vast empirical entrepreneurship and economic literatures with respect to the impact of education on performance in and choice of entrepreneurship.

For the sake of their meta-analysis, VVV have first gathered all relevant studies pertaining to industrialized countries that meet certain reasonable criteria. Each of the 94 resulting studies measures, among others, the impact of schooling on entrepreneurship entry, performance, or both for a specific sample. This results in 299 observations in their database. Almost 50% of these observations pertain to the relationship between education and entrepreneur performance, the topic of this paper, whereas the remainder measures the relationship between education and entrepreneurship entry (or the more hybrid measure of whether or not an individual is an entrepreneur in a certain year). The 145 studies measuring the relationship between performance and

education use various performance measures, but the majority (i.e., 58%) focusses on the entrepreneur's earnings, which is consistent with the focus of the current study. Furthermore, out of the set of performance studies, 38% uses the most commonly used measure of educational achievement that we are interested in, i.e., years of schooling. Taking the intersection of these two subsets of the sample, only 34 observations appear to be measuring the returns to education, i.e., the effect of years of education on earnings, in the sense that we are interested in. Most of these studies involve the United States. An additional condition should be met for a comparable measurement of the rate of return to education: earnings should be measured in logarithms. This additional requirement is met by 21 observations only, all pertaining to the US. Apart from concluding from this that the use of definitions of performance and education has been fairly scattered, we wish to pay attention to four important outcomes from this meta-analysis.

First, the impact of education on selection into an entrepreneurial position is mostly, i.e., in 75% of cases, insignificant. The impact of schooling on performance, however, is unambiguous and significantly positive for 67% of the observations. For the sub-sample of observations that consider earnings as their performance measure, this percentage amounts to even more than 80. We conclude that entrepreneurship performance is significantly affected by schooling.

Second, the meta-analysis gives insight into the level of the returns to education for entrepreneurs. This insight, though, can only be based on the small sub-sample of 21 US observations that use similar measures for education and for earnings. The return to a marginal year of schooling in

terms of the income it generates turns out to be 6.1%. on average.

Third, the meta-analysis allows a comparison of the rate of return to education for entrepreneurs to the returns to education for employees. This comparison is based on the results from studies that compare the rates of returns of these two groups of labor market participants using one dataset and thereby one set of definitions, time periods, countries, etcetera. Approximately twenty papers have actually measured the returns to education for entrepreneurs and employees in a comparable fashion. From these studies the third result is obtained: the returns to education are at least as high for employees as they are 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.

A smaller part 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 (WSH) is mixed. Studies based on US data reject the WSH (Fredland & Little 1981, Tucker 1985, Tucker 1987, Evans & Leighton 1990, Robinson & Sexton 1994), 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 Nether-

lands) support the WSH (Rees & Shah 1986, De Wit & Van Winden 1989, Brown & Sessions 1998, Brown & Sessions 1999). The latter result implies that the returns to education are lower for entrepreneurs than for employees in Europe.

The majority of the twenty papers 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 (Moore 1982, Gill 1988, MacPherson 1988, Borjas & Bronars 1989, Fairlie & Meyer 1996, Lombard 2001, Lofstrom 2002). The results from these (exclusively US) studies are consistent with the results obtained in the screening literature: the estimated returns to education for entrepreneurs are at least as high, and usually higher, than for employees.

The fourth conclusion from the meta-analysis is quite striking: all results obtained so far are potentially biased. Estimation and identification strategies used to identify the effect of education on performance have merely measured the (conditional) correlation between education and performance rather than the causal effect, which is the estimate of interest. There are at least two possible sources of inconsistency when OLS is used to estimate this relationship. First, there may be unobserved individual characteristics, such as ability and motivation, that affect both the schooling level attained and subsequent entrepreneurial performance. The omission of these unobserved characteristics from a performance equation would also serve to bias OLS estimates, where the direction and magnitude of the bias depends on the correlation between these characteristics and the schooling level attained.

Second, the schooling decision is probably endogenous in a performance equation because individuals are likely to base their schooling investment decision, at least in part, on their perceptions of the expected payoffs to their investment.

The literature studying the returns to education for employees has acknowledged this, and assessed the extent of these sources of bias. Theory predicts that omitting ability in the wage equation causes OLS-estimates to be upward biased (Griliches 1977, Harmon & Walker 1995, Ashenfelter, Harmon & Oosterbeek 1999). Empirical studies demonstrate mixed evidence on the existence and even sign of such an ability bias. These studies compare the estimated returns to education obtained with and without using a set of ability proxies. Using such a set of ability proxies might resolve the problem of unobserved heterogeneity. Several methods to cope with the endogeneity problem have been applied recently to estimate the returns to education for employees. The general conclusion from these studies is that OLS-estimates of the returns to education for employees are biased downwards (Ashenfelter et al. 1999).

The potential bias also makes the comparisons of returns to education for entrepreneurs and employees suspicious. Following Griliches (1977), the neglect of unobserved influential characteristics and not dealing with the endogenous nature of the education decision can have a different impact on the estimate of the returns to education for entrepreneurs and employees. As a result, the conclusion that the returns to education for employees is higher in Europe and lower or equal to the returns to education for entrepreneurs in the United States should be re-evaluated. It can only be maintained

when it would be supported by additional research that uses more recently developed estimation methods that account for endogeneity and unobserved heterogeneity.

The remainder of this section is devoted to a short presentation of possible methods to obtain more consistent estimates. There are basically four methods to account for the potential problems of endogeneity and/or unobserved heterogeneity when estimating the returns to education. All four have been applied to the estimation of the returns to education for employees (Ashenfelter et al. 1999).

The first strategy to cope with unobserved ability is trying to make the unobservable observable. To this end, various proxies of intelligence and test scores have been added to equations from which estimates of returns to education result. The effects so far of adding ability controls on the estimated returns to education are negative for the United States, positive elsewhere and (hence) ambiguous in total (see Ashenfelter et al., 1999, Table 3).

However, there are many reasons why the number of years of schooling could be correlated with the disturbance term and one of these is unobserved ability. In addition, inclusion of particular ability proxies on the right-hand side of the earnings function does neither completely purge the estimated returns from ability bias due to an imperfect correlation between such proxies with on-the-job ability nor is it sufficient to control for endogeneity since ability is not necessarily perfectly correlated with time-discounting behavior and/or with other factors such as the degree of risk aversion. Additional approaches are thus required and used.

The second strategy to identify causal effects is setting up a randomized

experiment (Leuven, Oosterbeek & van der Klaauw 2003). The proper design of an experiment requires a random assignment of individuals into a treatment group (participating) and a control group (not participating). In this manner, endogeneity does not play any role since the choice to follow education is forced. The problem is that setting up an experiment where some people do not get (higher) education but others do, is often ethically not feasible. Therefore, this identification strategy is seldom used.

The third strategy to identify the causal effect of education on performance makes use of the specific characteristics of monozygotic twins (Ashenfelter & Kruger 1994, Behrman & Rosenzweig 1999, Rouse 1999, Bonjour, Cherkas, Haskel, Hawkes & Spector 2003). The basic idea is that monozygotic twins share the exact same genetic endowment and usually experience even more similar environments than non-twin siblings or dizygotic twins do. Identification comes from those twins who differ in their amount of schooling obtained and their earnings, assuming that all unobserved factors are approximately equal. There has however been some critique by Bound and Solon (1999) on this seemingly flawless identification technique: most twin studies rely on small samples that usually describe twins who volunteer to participate, twin strategies are very sensitive to measurement errors, they do not really cope with the endogeneity of the schooling decision, and it is not clear why twins who are genetically identical end up with different outcomes.

The fourth identification strategy used is the instrumental variable (IV) approach. The idea is to find instruments that explain a substantial proportion of the variance of the endogenous variable, education in this case, while at the same time these instruments are not allowed to be related to the

dependent variable - i.e., earnings. In this way, the instrumented endogenous variable is not related to the error term anymore. This is the strategy that we will apply. Although using IV seems a natural way to estimate the effect of education on earnings, it is not without critique. The most important issue in using IV is the type of instruments used. Often family background variables are used in this respect. According to Card (1999), however, the use of family background variables as instruments in IV (especially parental education) can be problematic. He states that the ability of the parents is possibly transferred to their children by inheritance. Family background variables would then have an additional effect on performance. If this is the case, then the family background variables would not be valid IVs. Luckily, the validity of instruments can be empirically determined (Trostel, Walker & Woolley 2002), and the set of instruments that we propose has been empirically validated in earlier work (Blackburn & Neumark 1993). In general, IV-estimates of the returns to an employee's education are higher than estimates obtained by means of OLS (Ashenfelter et al., 1999), no matter what set of identifying instruments is used.

We wish to identify the causal effect of education on incomes for entrepreneurs and employees in exactly the same manner. This will allow us to compare the estimates and biases between the two labor market segments. From the above mentioned techniques, we use ability proxies and IV-estimates, where the instrumentation is based on detailed family background information. Additionally, we will empirically determine the validity of our instruments.

3 Data

3.1 Data description

We estimate the effect of education for both entrepreneurs and employees on a sample drawn from the National Longitudinal Survey of Youth (NLSY) in the US. We replicate several aspects of an earlier study by Blackburn and Neumark (1993: BN hereafter) that estimated the rate of return to education for employees based on the same sample. The nationally representative part of the NLSY consists of 6,111 individuals aged between 14 and 22 years in 1979. They have been interviewed annually up to 1994, and since then on a bi-annual basis. From these 6,111 persons we extract, per year observed, the hourly income,[1] the number of years of education attained, and a rich set of personal and family background variables. As a performance measure we use hourly earnings averaged over a year. We dropped from the sample all farmers, people working less than 300 hours a year, persons who report working while still in school, entrepreneurs who were also employed for a considerable amount of time, and those workers who earned less than one dollar an hour.[2]

Table 1 shows the means and standard deviations of the variables used in our wage equations based on the representative part of the sample. The values in Table 1 represent the averages of the specific variable over the pe-

[1] Created by the NLSY Research Bureau as an average over the past year.

[2] Farmers are excluded from the sample because their economics are very different from all other occupations. From 1979 to 2000, we left out 299 farmers. Most studies drop farmers from their samples or study them separately (cf. VVV). We wanted to make sure that no hobby entrepreneurs are included in our sample. Therefore, we used the lower boundary of 300 working hours as an entrepreneur per year and a lower hourly earnings threshold of one dollar. Including or excluding the lower hourly earnings threshold for both employees and entrepreneurs did not lead to different results.

riod 1979-2000. Table A in the Appendix gives an overview of the full names of the abbreviated variable names. We highlight two variables that are of main interest. First, we notice that for hourly pay the mean and especially the standard deviation are higher for entrepreneurs than for employees. This common observation can be explained by the fact that the incomes of entrepreneurs have no fixed minimum and do not fit into preformed salary scales. Second, the average level of education that individuals complete in the US is slightly more than high school, being equal for entrepreneurs and employees ($\approx$ 13 years).

-Insert Table 1 about here-

3.2 Data suitability

We will now address the suitability of the data to compare the consistent estimates of the returns to education for entrepreneurs with these of employees. First, the NLSY contains the Armed Services Vocational Aptitude Battery (ASVAB), which is an IQ-like test score that we can use as a proxy for ability. ASVAB (administered in 1979-1980)[3] 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.[4] Following BN, we combine component (1) to (4) and (8) into an "academic" test score, and component (6),

[3]Adminstration of the ASVAB as early as 1979-1980 allows us to treat this variable as exogenous.

[4]This decomposition into ten components is suggested by the NLSY Research Bureau.

(7), (9) and (10) into a "non-academic" one.[5] Splitting up the ASVAB into an academic and a non-academic component gives more insights than using one combined score, while at the same time it is still possible to make sensible interpretations. Following BN, we remove the age effects from the ASVAB, as respondents were of different ages when the test was administered, by regressing each normalized test score on a set of seven age dummies. We use the individuals' residuals as the new test scores. Including the ASVAB in the wage equation allows for a comparison between the returns to education estimate with and without control for "omitted" ability, as far as this is proxied correctly by the ASVAB.

A second relevant feature of the NLSY is the presence of detailed family background variables. These family background characteristics are possibly good predictors of the educational level of the respondent while otherwise independent of their future wage. 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). Following BN, we use those variables as identifying instruments that pass the criteria for quality and validity. The quality criterion comes down to requiring a sufficient correlation between the set of instruments and the endogenous regressor, education in this case. Instruments are valid if they affect performance via the education equation only. A set of instrument therefore passes the validity test if it is not correlated with the error term in the performance equation. Variables proposed by BN as components of the set of identifying instruments were dropped if they turned out invalid. The resulting set of

[5]Coding speed is not used (see Bishop 1991).

valid identifying instruments that is of sufficient quality differs only slightly between entrepreneurs and employees; it consists of "A Library Card present in the household at age 14", "Magazines present in the household at age 14", "Fathers education", "Mothers education", "A dummy for the presence of a male in the household", "Number of siblings", "Number of older siblings", "If there is a foreign language spoken in the household" and, finally, a dummy measuring "If both parents are present in the household" (see Table A in the Appendix).

However, critical evaluations of using family background as identifying instruments for education in an income equation have been expressed by Card (1999): he doubts the validity of instrument sets consisting of parental background variables. The idea is that family background variables are very likely to be correlated with a child's innate ability, and hence affect her or his educational attainment choices, as well as expected returns from school. Such a set of instruments would therefore not be valid. We acknowledge this drawback of our choice of instruments. At the same time we try to measure and minimize its potential negative effects by performing Sargan's validity test and by including indicators for the child's ability into the regression. Indeed, as we shall see below, the resulting set of identifying instruments for education, which also includes a measure of parental education, passes the tests of quality and validity.

A third relevant feature of the sample is that it includes both entrepreneurs and employees, and records individuals' switches between these states over time. All entrepreneurship spells, also short ones, are recorded. Therefore, the subsample of entrepreneurs does not suffer from survival bias, i.e.,

the returns to education will not pertain to surviving entrepreneurs only. Moreover, the incomes, education levels and all other relevant variables are measured in a comparable way for both groups. These features of the data enable estimating the returns to schooling for employees and entrepreneurs in a comparable fashion.

A fourth important feature of the NLSY is its panel character. Nineteen years of information on approximately 6,000 individuals results in a large number of data-points.[6] An advantage of panel data over cross-sectional ones is the possibility to control for time-varying individual characteristics and economic fluctuations. To ascertain that the year and age dummies are not correlated, and therefore not interpretable, we use a decomposition technique suggested by Deaton (2000). This decomposition ensures that cohort and time trends are orthogonal to each other. Moreover, panel data allow the estimation of a fixed-effects model if there is sufficient variation over time in the most important variables (education and income, in our case). This, in turn, would solve the omitted variable bias issue. Since education does not vary (enough) over time, we apply a random-effects model, and thus control for biases via IV and ability proxies.

4 Estimation results

In this section, we estimate the effect of education on income for both entrepreneurs and employees. As a benchmark, we start with estimating the wages of both groups as in Equation (1). W_{it} represents the log hourly wage,

[6] In general, this results in more degrees of freedom and less multicolinearity. These factors together, again, produce improved efficiency (Hsiao 1986).

S_{it} the years of schooling obtained, X'_{it} a vector containing control variables, c_i an unobserved individual-specific effect, and u_{it} a white noise error term. D'_{it} represents a vector of dummies controlling for cohort effects, age effects[7] and macroeconomic shocks using a method suggested by Deaton (2000). This method transforms the year dummies, age dummies and birth year dummies so that the year effects add to zero, and are orthogonal to a time trend.

$$W_{it} = \alpha S_{it} + \beta X'_{it} + \delta D'_{it} + c_i + u_{it} \tag{1}$$

4.1 Returns to education

We report the OLS-estimates resulting from estimating Equation (1) for both entrepreneurs and employees in Columns (1) and (4) of Table 2, respectively. The results show that without controlling for ability the OLS-estimated returns to education are .071 for entrepreneurs and .067 for employees. In accordance with previous studies using US data (Fredland & Little 1981, Tucker 1985, Tucker 1987, Evans & Leighton 1990, Robinson & Sexton 1994), the returns are slightly higher for entrepreneurs than for employees. They are of the same order of magnitude, though.

The first step in increasing the quality of our estimates is to control for the bias resulting from omitting ability. We include two compound ability proxies from the ASVAB test scores in the wage equation - that is, academic and non-academic ability. We see in Columns (2) and (5) of Table 2 that the estimates for the return to education drop for both entrepreneurs and employees to .067 and .059, respectively, while all other results stay approximately the same.

[7] We followed Harmon and Walker (1995) in using age instead of experience in the wage equation. Experience is a negative function of education, and is therefore endogenous.

These results are in accordance with the theory that the returns to education are upward biased when ability is omitted (Griliches 1977). A further effect is that the bias appears to be larger for employees than for entrepreneurs, as the drop in the schooling estimate is smaller for entrepreneurs. Contrary to expectations but in accordance with BN, almost none of the ability proxies is significant. However, if we use only one compound measure of ability, its coefficient is significantly positive for both entrepreneurs and employees (see Table B in the Appendix). Nonetheless, this finding indicates that it is worthwhile to experiment with further decompositions of ability to see which component explains most of the variance in both equations. Our next step is to apply IV-estimation.

-Insert Table 2 about here-

For the instrumentation of the possible endogenous education variable we use the discussed set of family background variables as identifying instruments. We check whether the proposed set of instruments is of sufficient quality, and whether the suggested IVs are valid. To test the first criterion we estimate the regression postulated in Equation (2). Schooling is represented by S_{it}, and family background variables by F_i'. X_{it}' is a vector of controls, and D_{it} a vector of dummies.

$$S_{it} = \gamma F_i' + \kappa X_{it}' + \varphi D_{it}' + \mu_i + \epsilon_{it} \tag{2}$$

The estimation results of Equation (2) presented in Table 3 show that all family background variables are significant. About 40% of the variation in education is explained. A Chi-square test supports the quality of the set of

identifying instruments.

-Insert Table 3 about here-

To test the second criterion - i.e., instrument validity and over-identification - we use the Sargan-test of the validity of instruments.[8] For both employees and entrepreneurs a set of seven identifying instruments results. The $\chi^2_{(df6)}$ test-statistic for entrepreneurs is 4.72 with a p-value of 0.58. This means that the null hypothesis, assuming exogeneity of the instruments, is not rejected. The $\chi^2_{(df6)}$ test-statistic for employees is 4.24 with a p-value of 0.64. So, also for employees the instruments for education are exogenous. Hence, the use of this particular set of family background variables as identifying instruments is valid.

There remains one important fact that we need to test. Is it justified to assume that education is endogenous? If not, implying that education is exogenous, OLS-estimates would not be biased due to endogeneity. We performed a Hausman-test to investigate this question. Indeed, we find that instrumentation of the education variable is necessary - i.e., education is endogenous.

Columns (3) and (6) of Table 2 show the IV-results estimated with 2SLS. Applying IV results in significantly higher estimates of the returns to education. The returns for employees jump from 6.7% to 10.7%. This is in line with previous research, using various sets of identifying instruments (Ashenfelter et al. 1999), and is comparable to the effect found by BN. A more revealing finding is the jump in the returns to education for entrepreneurs of 7.1 percent point to 14.2%. This leads to the remarkable result that the returns to

[8]Following BN, we removed those variables that did not pass this test.

education for entrepreneurs are 33% higher than the comparable returns for employees.

However, we have to perform several robustness checks to see whether the returns to education are really higher for entrepreneurs than for, employees and to be able to interpret the result appropriately. In what follows we address and test several explanations that would lead to very different interpretations.

4.2 Robustness checks

So far, we have implicitly assumed that individuals' choices for education do not affect their choice of employment status, i.e., for entrepreneur or for employee. However, every time a specific sample is selected from a larger population, there is the possibility of *sample-selection bias* (Wooldridge 2002). It may, for instance, be the case that individuals with high (or low) education are selecting into (or away from) self-employment. More specifically, there could be two alternative, but spurious causes for the finding that entrepreneurs have a higher return to education than employees. First, returns to education for entrepreneurs would be higher if, on average, entrepreneurs are lower educated and the returns to schooling feature decreasing returns. Second, if entrepreneurs are higher educated, on average, and the returns to education are increasing, then entrepreneurs would also get a higher return to their education than would employees. Such combination of findings would lead to the (misleading) observation that employees experience higher returns to education than do entrepreneurs. This example illustrates the importance

of dealing with such issues of selectivity.[9] Few previous studies have done this in satisfactory manners (Rees & Shah 1986, Gill 1988, MacPherson 1988, De Wit & Van Winden 1989, Dolton & Makepeace 1990, Taylor 1996, Clark, Drinkwater & Leslie 1998).

Hence, to investigate whether one of these alternative selectivity-related explanations is valid, we have to analyze: (1) whether the returns to education are increasing, decreasing or constant as a function of years of education; and (2) whether entrepreneurs have higher or lower education levels than their employed counterparts. We pursue two routes to assess the functional form of returns to education, where we distinguish between entrepreneurs and employees. First, we split both the sample of entrepreneurs and the one of employees into two equal parts: one with higher than median education levels, the other with below-median education levels. A comparison of the resulting (four) IV-estimates of the return to education shows that, if anything, the returns to education are increasing: they are higher for the better educated halves of both samples.

As to the second route, we re-estimated our wage equations while including the education squared as an additional regressor.[10] For both entrepreneurs and employees the returns to education follow a U-shaped distribution, with its minimum at 8.2 and 4.7 years of education completed for entrepreneurs and employees, respectively. By far the majority of en-

[9]Throughout, we assume that individuals decide upon their educational investments and then choose to become employees or entrepreneurs. In reality, these decisions might be made simultaneously or individuals might choose their employment status and then the most compatible educational stream.

[10]We account for the potential endogenous character of education squared by including the residuals of the first-stage education equation and the first-stage squared education equation.

trepreneurs - i.e., 98.5% - has completed more than 8.2 years of education, whereas even more than 99.9% of the employees have completed more than 4.7 years of schooling. Hence, for 99.8% of the total sample returns to education are increasing. We conclude that both routes lead to the same conclusion: if anything, returns to education are increasing, both for entrepreneurs and for employees. Selectivity would therefore only be a possible spurious explanation for the higher returns for entrepreneurs in case higher educated individuals are selected into entrepreneurship.

We use several methods to investigate the possibility that education determines the selection into self-employment positively. A first indication can be found in Table 1: the mean education levels for entrepreneurs and employees are almost equal. To check whether the education level has indeed no influence on the employment status, as the similar mean levels of education suggest, we estimate a random-effects probit. The results in Table 4 reveal that the effect of education on the time-varying employment status is insignificant. This result is consistent with previous research into the relationship between years of education and entrepreneurship selection.[11]

-Insert Table 4 about here-

An alternative method to test for sample-selection bias is based on work by Nijman and Verbeek (1992). They suggest to include a lag of the employment status in the wage regression. The underlying assumption to this approach is that sample selection is related to the idiosyncratic errors μ_{it} only. Under the null hypothesis, μ_{it} should not be correlated to any other

[11]The meta-analysis by VVV demonstrated that 75% of all such studies find an insignificant effect.

period than the current period. When performing this test, our results indicate that the lagged employment status variable is insignificant in both the income equation of entrepreneurs and employees. So, again, the sample-selection hypothesis is rejected.

Another, quite similar test is to include in the wage equation the fraction of total past labor market experience that the respondent has been an entrepreneur. In this way, we measure more precisely in what way state dependence influences our results. Performing this test reveals that the effect is insignificant for entrepreneurs and significant for wage workers. A problem attached to introducing a lag or the 'entrepreneurial experience fraction' into the wage regression is that it includes the human capital effect on income of previous entrepreneurship experience. This could obscure our selection test. A solution to this problem is to include a lead instead of a lag of employment status into the wage equation: current wages are unlikely to be affected by the future decision to be an entrepreneur. Incorporating a lead in the wage equation results in the rejection of the presence of selection for both entrepreneurs and employees.

All in all, the results lead to the conclusion that our estimates of the returns to education for entrepreneurs and employees are not biased by selectivity. Therefore, a correction is not required.

The second robustness check relates to the unit of measurement of the returns to education. Our estimations suggest that the *percentage* gain in terms of income of a marginal year of education for entrepreneurs is higher than for employees. The question is whether the returns to education are also higher for entrepreneurs in absolute terms: does a year of education

generate more dollars per hour for an entrepreneur than for an employee? For instance, if entrepreneurs earned much less per hour, on average, than employees, then the higher percentage gain could correspond with a lower dollar gain. To answer this question, we estimate the returns to education for entrepreneurs and employees using hourly pay as the dependent variable instead of log hourly pay, keeping all else equal. We present the results in Table C in the Appendix. The results still provide support for the finding in terms of percentages: the returns to education in dollars are higher for entrepreneurs than for employees.

A third robustness check concerns the assumed *log-normality of the distribution of hourly pay.* The estimates in Table 2 have been obtained under this assumption, both for entrepreneurs and for employees. Especially for entrepreneurs, this assumption might be questionable (Blanchflower & Meyer 1994) To this end, we re-calculated the percentage returns to education for both groups using the results from Table C in the Appendix. In this way, we circumvent using the results obtained under the assumption of log-normality of the hourly earnings distribution.[12] The outcome, again, supports our claim that the returns are higher for entrepreneurs than for employees (with returns to education of 0.141 and 0.127, respectively). This outcome is apparently invariant to the assumed log-normality of the hourly earnings distribution.

Our fourth 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* obtained by higher educated entrepreneurs. It could be

[12]We divide the IV-estimate of the education coefficient in table C by the average wage (for entrepreneurs and employees separately).

the case that more highly educated individuals require higher risk premia for being an entrepreneur: higher educated individuals might experience more additional income risk as an entrepreneur compared to an employee vis-à-vis lower educated individuals. We base this robustness check on three observations derived from our data. First, regressing the (time) variance of the incomes[13] of individual entrepreneurs on their education levels and some control variables renders no significant education effect. Hence, the variance of the entrepreneurial income 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. Based on this, we can safely assume that the higher returns to education for entrepreneurs are not a kind of risk premium.

Hence, our finding that the returns to education are higher for entrepreneurs than for employees is not due to selectivity. It is neither due to wrong assumptions with regard to the functional form of the income distribution of entrepreneurs, nor to percentage gains that would not translate into absolute gains. Moreover, the higher returns to education for entrepreneurs cannot be explained in terms of higher required risk premia for higher educated entrepreneurs.

So why is education more valuable for entrepreneurs? We propose a

[13]To be more precise, the variance of the residuals of the income equations as presented above is the dependent variable in this case.

simple explanation: entrepreneurs have more freedom to optimize their employment of education. Entrepreneurs are not constrained by rules from superiors and can decide on how to implement 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 difficult for the employee to optimize the productive effect of education. Besides, the organization cannot adapt its structure to every individual due to organizational inertia and individual incompatibilities. This difference in ability to optimize the productivity of education for entrepreneurs and employees might therefore be an explanation for the higher returns to education for entrepreneurs vis-à-vis employees.

4.3 Remaining effects

Some of the effects of the control variables presented in Table 1 are notable. Males earn statistically and economically significantly higher incomes than females. This confirms previous findings for both segments of the labor market. Moreover, the extent of the gender effect differs largely across labor market segments. Male wage employees earn 24% more than female wage employees. The comparable difference between male and female entrepreneurs is 67%. This large difference of the gender effect of entrepreneurs vis-à-vis employees is consistent with previous studies (Moore 1982, Tucker 1987, De Wit & Van Winden 1989, Dolton & Makepeace 1990, Robinson & Sexton 1994). Interestingly, the correlation between being married and income is higher for employees than for entrepreneurs: it is more than twice as large for employees, and insignificant for entrepreneurs. The income of married employees

is 6% higher than the income of single employees, whereas this difference is an insignificant 2.5% for entrepreneurs. Previous findings support this result (Moore 1982, Tucker 1987, Gill 1988, Dolton & Makepeace 1990, Evans & Leighton 1990). A striking result is that the effect of race (i.e., being black) is much larger for entrepreneurs (Blanchflower & Meyer 1994) than for employees: 24 versus 10%. The support for this difference in previous literature is less clear. Fairly and Meyer (1996) and Moore (1982) support this finding, while Fredland and Little (1981) and Rees and Shah (1986) report that the effect of race is smaller for entrepreneurs than for employees. Evans and Leighton (1990) and Dolton and Makepeace (1990) even find that ethnicity is positively related to the incomes of entrepreneurs. A side remark should be made about these other studies. We explicitly distinguish "blacks" as an ethnic group. Doing this ensures that the often very mixed effect of other ethnicities is taken out. Other studies reporting race effects have neglected this differentiation, and are therefore less clearly interpretable.

Comparing the explanatory power of both equations leads to the observation that a much larger proportion of the variance in earnings can be explained by the observed factors for the group of employees than for entrepreneurs. This result is not uncommon (Poutvaara & Tuomala 2003).

5 Conclusion

The aim of this paper was to estimate the effect of education on the performance of entrepreneurs. The performance measure we 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 is lacking in the comparable entrepreneurship literature. Previous studies have measured the returns to education for entrepreneurs by means of (conditional) correlations. The following observations had emerged from a meta-analysis based on that literature: The returns to education for entrepreneurs are positive and their order of magnitude is comparable to the returns to education for entrepreneurs. However, these estimates might be biased if problems relating to unobserved heterogeneity, endogeneity and selectivity are present. We address and, if required, solve these problems when estimating the returns to education for entrepreneurs and for employees in a comparable way. Via the usage of ability proxies we offer evidence for the fact that omission of ability results in an upward bias of the education coefficient in an OLS-equation. This bias turned out to be more pronounced for employees than for entrepreneurs. Consistent with the literature, our OLS-estimates also indicate that the return to education is slightly higher for entrepreneurs than for employees. However, when we apply IV to solve the endogenous schooling problem, the returns to education jump from 5.9% to 10.7% for employees and from 6.7% to 14.2% for entrepreneurs, where the first jump is comparable to previous findings in labor economics. Surprisingly, this indicates that the returns to education for entrepreneurs are not slightly higher, but an impressive 33% higher than the returns to education for employees. The absence of sample selection bias and the further robustness of this result adds to the credibility of this finding. All together, we think that our findings bear implications for researchers and

policy makers.

The observation that OLS-estimates are biased and that the extent of this bias differs per labor market group should be the starting point for researchers investigating returns to education for entrepreneurs. Studies on screening implicitly assume that the bias resulting from omitting ability and not treating schooling as an endogenous variable is similar for entrepreneurs and employees. Our results do not support this assumption, and show that this bias differs per population studied. We suggest that further research should aim at understanding the size of this bias.

Furthermore, replicating our study with different data, also for various countries and possibly with different sets of instruments, would be useful to confirm the findings that are now based on one study only. Alternatively, our findings could be validated (or not) by means of twins research. In any case, more insight into the causal relationship between education and the performance of entrepreneurs is required. 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. As we shall see below, such research forms the basis for several policy implications.

Before we discuss policy implications, it is important to elaborate on the remaining untested assumptions made in this context. First, we assume that the social return of entrepreneurial activity is larger than the private return that accrues to the entrepreneur her or himself. Second, we assume that the difference between the social and private benefits of entrepreneurial activity is larger than this difference is for employees. A successful entrepreneur is, for

example, more likely to influence competition in a market positively than is an employee. Also, new and innovative ideas can be brought into the market more easily by entrepreneurs than by employees.[14] 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 expected return when belonging to the group of entrepreneurs, but by some (weighted) average expected return of both employment modes. Our fourth assumption is that individuals, as well as policy makers, bankers and other parties involved, do have no 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 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 the development of educational policies, as well as for bankers' and capital suppliers' strategies with respect to (selecting) starters. Moreover, the adequate design of subsidy and tax measures that target both entrepreneurs (and starters) and their capital suppliers, often by the national Ministry of Economic Affairs, or similar authorities, might benefit from this key insight.

Policy makers should be aware of our result that the returns to education for entrepreneurs are substantially higher than those for employees. In this

[14] In addition to these positive external effects of entrepreneurial activity, entrepreneurs are often credited for their impact on labor demand. However, this is a short-term rather than a long-term effect: in the absence of new firms, their demand for labor would be effectuated by growing incumbent firms.

respect, policy-making authorities could launch two policies. First, they can invest in higher education for (prospective) entrepreneurs. Second, they can invest in stimulating higher educated individuals to opt for an entrepreneurial career. The first policy will 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 also to society as a whole. In so doing, the social costs of bankruptcies will decrease accordingly. The second strategy appeals to the fact that, at least in Europe, entrepreneurship seems not to be the favored option, or even to be part of the choice set, amongst young people with higher education degrees. They usually prefer working in large multinational enterprises, rather than to even think about self-employment alternatives. We strongly believe in the benefits of governmental programs to stimulate the awareness among college and university students of the high potential of the entrepreneurship route. For sure, future research into the entrepreneurial returns of specific types of education may further increase the effectiveness of such policies.

References

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

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

Behrman, J. & M. Rosenzweig (1999), 'Ability biases in schooling returns and twins: A test and new estimates', *Economics of Education Review* **18**(2), 159–167.

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

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

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

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

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

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

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

De Wit, G. & 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.

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

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

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

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

Gill, A. M. (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.

Griliches, Z. (1977), 'Estimating the returns to schooling: Some econometric problems', *Econometrica* **45**(1), 1–22.

Harmon, C. & I. Walker (1995), 'Estimates of the economic returns to schooling for the United Kingdom', *American Economic Review* **85**(5), 1278–1286.

Hsiao, C. (1986), *Analysis of Panel Data*, Econometric Society Publication, Cambridge University Press.

Leuven, E., H. Oosterbeek & B. van der Klaauw (2003), 'The effect of financial rewards on students' achievement: Evidence from a randomized experiment', *Tinbergen Institute Discussion Paper 38/03* .

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 felxible, nonstandard work schedules', *Economic Inquiry* **39**(2), 214–237.

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

Moore, R. L. (1982), 'Employer discrimination', *Review of Economics and Statistics* **65**(3), 496–501.

Nijman, T. & 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.

Poutvaara, P. & J. Tuomala (2003), 'Entrepreneurs and workers', *CBER 2003-14* .

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

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

Rouse, C. (1999), 'Further estimates of the economic return to schooling from a new sample of twins', *Economics of Education Review* **18**(2), 149–157.

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

Trostel, P., I. Walker & P. Woolley (2002), 'Estimates of the economic return to schooling for 28 countries', *Labour Economics* **9**(1), 1–16.

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

Tucker, I. B. (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., C.M. Van Praag & W. Vijverberg (2003), 'Entrepreneurship, selection and performance: A meta-analysis of the role of education', *Working Paper* . University of Amsterdam.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Table 1: Descriptive Statistics

	Employees			Entrepreneurs		
	n^*	Mean	SD	n^*	Mean	SD
Academic	62073	0.06	0.84	4005	0.07	0.89
Age	62073	28.17	5.66	4005	30.28	5.31
Being Religious	61902	0.18	0.38	3999	0.15	0.36
Being Shy	60349	0.61	0.49	3873	0.55	0.50
Birth Year	62073	1960	2.18	4005	1960	2.19
Black	62073	0.11	0.31	4005	0.05	0.21
Education	61912	12.89	2.39	4001	12.97	2.45
Fathers Education	56469	11.74	3.47	3695	12.17	3.44
Foreign Language	62054	0.14	0.35	4002	0.14	0.35
Gender	62073	0.52	0.50	4005	0.64	0.48
Hispanic	62073	0.06	0.24	4005	0.05	0.22
Library Card	61823	0.74	0.44	3991	0.77	0.42
Live Outside City	59589	0.22	0.42	3792	0.21	0.41
Live in South	59062	0.33	0.47	3807	0.26	0.44
Locus of Control	61667	49.94	28.56	3952	54.01	27.80
Log Hourly Wage	62073	10.32	15.85	4005	14.38	26.72
Magazines	61701	0.66	0.47	3990	0.74	0.44
Married	62072	0.50	0.50	4004	0.63	0.48
Mothers Education	58917	11.52	2.66	3829	12.04	2.45
No Male in Home	62073	0.11	0.31	4005	0.09	0.29
Not Academic	62073	0.08	0.80	4005	0.15	0.87
Not Healthy	61468	0.02	0.15	3955	0.04	0.19
Older Siblings #	57815	1.97	1.89	3724	1.87	1.88
Parents Present	62073	0.76	0.43	4005	0.77	0.42
Siblings #	62021	3.29	2.20	4000	3.17	2.08

**Note.* Observation number displayed is the aggregate observation number from 1979 till 2002.

Table 2: OLS and IV wage equation estimation results

	Employees			Entrepreneurs		
	(1. OLS)	(2. OLS)	(3. IV)	(4. OLS)	(5. OLS)	(6. IV)
Education	.067***	.059***	.107***	.071***	.067***	.142***
	.002	*.002*	*.007*	*.009*	*.010*	*.026*
Academic		.015	-.093***		.026	-.142
		.013	*.021*		*.065*	*.088*
Not Academic		.064***	.113***		-.002	.075
		.013	*.017*		*.064*	*.074*
Gender	.242***	.229***	.240***	.668***	.668***	.648***
	.009	*.009*	*.010*	*.044*	*.045*	*.050*
Marital Status	.058***	.058***	.058***	.025	.025	.026
	.004	*.004*	*.004*	*.032*	*.032*	*.036*
Not Healthy	-.048***	-.047***	-.050***	-.076	-.076	-.061
	.010	*.010*	*.011*	*.070*	*.070*	*.079*
Live Outside City	-.071***	-.071***	-.064***	-.126***	-.127***	-.150***
	.006	*.006*	*.006*	*.041*	*.041*	*.045*
Live in South	-.077***	-.068***	-.063***	.034	.034	.044
	.010	*.010*	*.012*	*.048*	*.048*	*.054*
Locus of Control	.001***	.001***	.001***	.003***	.003***	.002**
	.000	*.000*	*.000*	*.001*	*.001*	*.001*
Hispanic	-.005	.014	.018	-.123	-.113	-.094
	.022	*.022*	*.025*	*.117*	*.118*	*.133*
Black	-.136***	-.090***	-.103***	-.244***	-.231**	-.235***
	.014	*.015*	*.018*	*.088*	*.090*	*.114*
Constant	.048	.145	-.499***	-.962	-.926	-1.590*
	.092	*.092*	*.133*	*.848*	*.850*	*.876*
R^2	0.49	0.50	0.47	0.28	0.28	0.27
N	55769	55769	47152	3519	3519	2951

Note. Standard errors in italics. Each estimation controls for year effects, age effects and macroeconomic shocks (Deaton 2000). *p-value $\leq$ .10, **p-value $\leq$ .05 and ***p-value $\leq$ .01.

Table 3: OLS education equation

	Employees	Entrepreneurs
	(1)	(2)
Library Card	.383***	
	.065	
Magazines		.491***
		.149
Fathers Education	.191***	.129***
	.009	*.023*
Mothers Education		.205***
		.032
No Male in Home	.484***	
	.117	
Siblings #	-.141***	-.128***
	.019	*.046*
Older Siblings #	.086***	.097***
	.021	*.049*
Foreign Language	.198**	.405*
	.100	*.212*
Parents Present	.665***	.380**
	.084	*.152*
Constant	9.323***	4.646**
	.236	*2.042*
R^2	0.39	0.40
N	47040	2951

Note. Standard errors in italics. Each estimation controls for year effects, age effects and macroeconomic shocks (Deaton 2000). *p-value $\leq$ .10, **p-value $\leq$ .05 and ***p-value $\leq$ .01.

Table 4: Random-Effects Probit on the decision to become an entrepreneur

Education	-.013
	.013
Academic	-.219***
	.079
Not Academic	.148*
	.083
Gender	.375***
	.060
Black	-.829***
	.100
Hispanic	-.283**
	.117
Birth Year	.001
	.017
Being Shy	-.101**
	.057
Locus of Control	.013
	.012
Being Religious	-.148***
	.066
Constant	-5.245
	33.197
Loglikelihood	-10042.5
N	63475

Note. Standard errors in italics. Each estimation controls for year effects, age effects and macroeconomic shocks (Deaton 2000). *p-value $\leq$ .10, **p-value $\leq$ .05 and ***p-value $\leq$ .01.

Appendix

Table A: Codebook

Variables	
Age	In years
Age Squared	Age squared in years
Academic	An ASVAB test score indicating academic intelligence
Being Religious	A dummy that indicates 1 if the person is adhering to a strict religion
Being Shy	A dummy that indicates 1 if the person was shy as a child
Black	A dummy that indicates 1 if the person is black
Education	Total education completed (in years)
Entrepreneur	A dummy that indicates 1 if the person is an entrepreneur
Gender	A dummy that indicates 1 if the person is male
Hispanic	A dummy that indicates 1 if the person is Hispanic
Live in South	A dummy that indicates 1 if the person lives in the south of the US
Live Outside City	A dummy that indicates 1 if the person lives outside a city
Locus of Control	Locus of control, with a high value indicates high internal locus of control
Log Hourly Wage	Log income per hour (in dollars)
Marital Status	A dummy that indicates 1 if the person is married
Not Healthy	A dummy that indicates 1 if the person feels not healthy
Not Academic	An ASVAB test score indicating non academic intelligence
White	A dummy that indicates 1 if the person is white
Instruments	
Fathers Education	Education of the father (in years)
Foreign Language	A dummy that indicates 1 if foreign language spoken in household
Library Card	A dummy that indicates 1 if there was a library card in household at age 14
Magazines	A dummy that indicates 1 if there were magazines in household at age 14
Mothers Education	Education of the mother (in years)
No Male in Home	A dummy that indicates 1 if there was no man in the household at age 14
Older Siblings #	Number of older siblings
Parents Present	A dummy that indicates 1 if Father and mother in household at age 14
Siblings #	Number of siblings

Table B: OLS wage equation with composite ability measure

	Employees	Entrepreneurs
	(1)	(2)
Education	.059***	.068***
	.002	*.009*
IQ	.011	.003
	.001	*.004*
Gender	.228***	.664***
	.009	*.044*
Marital Status	.058***	.025
	.004	*.032*
Not Healthy	-.047***	-.076
	.010	*.070*
Live Outside City	-.071***	-.127***
	.006	*.041*
Live in South	-.069***	.034
	.010	*.048*
Locus of Control	.001***	.003***
	.000	*.001*
Hispanic	.017	-.116
	.022	*.118*
Black	-.092***	-.232
	.015	*.090*
Constant	.149	-.933
	.092	*.849*
R^2	.497	.277
N	55769	3519

Note. Standard errors in italics. Each estimation controls for year effects, age effects and macroeconomic shocks (Deaton 2000). *p-value $\leq$ .10, **p-value $\leq$ .05 and ***p-value $\leq$ .01.

Table C: OLS and IV wage equation where the dependent variable is hourly income (instead of log hourly income)

	Employees			Entrepreneurs		
	(1. OLS)	(2. OLS)	(3. IV)	(4. OLS)	(5. OLS)	(6. IV)
Education	.949***	.839***	1.328***	.984***	.799***	2.068***
	.039	*.044*	*.142*	*.203*	*.225*	*.634*
Academic		.397	-.816**		1.558	-1.214
		.270	*.427*		*1.435*	*2.123*
Not Academic		.434***	1.090***		-.407	1.078
		.270	*.341*		*1.423*	*1.782*
Gender	2.642***	2.545***	2.616***	6.417***	6.487***	6.623***
	.182	*.186*	*.217*	*1.000*	*1.040*	*1.213*
Marital Status	1.051***	1.031***	1.080***	.122	.158	.003
	.146	*.146*	*.166*	*1.029*	*1.030*	*1.215*
Not Healthy	-.654	-.613	-.612	-3.662	-3.700	-3.383
	.400	*.400*	*.455*	*2.500*	*2.500*	*2.932*
Live Outside City	-.991***	-.990***	-.936***	-2.493**	-2.521**	-3.172**
	.189	*.189*	*.216*	*1.150*	*1.150*	*1.349*
Live in South	-.907***	-.816***	-.781***	1.400	1.374	2.357
	.202	*.202*	*.240*	*1.076*	*1.077*	*1.283*
Locus of Control	.013***	.011***	.006	.049**	.045**	.032
	.003	*.003*	*.004*	*.017*	*.018*	*.021*
Hispanic	.487	.676	.890	3.247	3.712	4.414
	.441	*.441*	*.526*	*2.757*	*2.767*	*3.268*
Black	-1.169***	-.684***	-.738*	-4.437**	-3.830*	-5.344*
	.296	*.305*	*.387*	*2.245*	*2.284*	*3.173*
Constant	-11.251***	-9.869***	-17.545***	-26.568	-24.958	-39.750
	3.581	*3.589*	*4.573*	*27.907*	*27.914*	*30.295*
R^2	.100	.101	.09	.072	.073	.072
N	55769	55769	47152	3519	3519	2952

Note. Standard errors in italics. Each estimation controls for year effects, age effects and macroeconomic shocks (Deaton 2000). *p-value $\leq$ .10, **p-value $\leq$ .05 and ***p-value $\leq$ .01.