DISCUSSION PAPER SERIES

IZA DP No. 3058

Why Are the Returns to Education Higher for Entrepreneurs than for Employees?

Justin van der Sluis
Mirjam van Praag
Arjen van Witteloostuijn

September 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Why Are the Returns to Education Higher for Entrepreneurs than for Employees?

Justin van der Sluis
Amsterdam Center for Entrepreneurship and University of Amsterdam

Mirjam van Praag
Amsterdam Center for Entrepreneurship, Tinbergen Institute, University of Amsterdam and IZA

Arjen van Witteloostuijn
University of Groningen and University of Durham

Discussion Paper No. 3058
September 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3058
September 2007

ABSTRACT

Why Are the Returns to Education Higher for Entrepreneurs than for Employees?*

We compare the returns to education (RTE) for entrepreneurs and employees, based on 19 waves of the NLSY database. By using instrumental variable techniques (IV) and taking account of selectivity, we find that the RTE are significantly higher for entrepreneurs than for employees (18.3 percent and 9.9 percent, respectively). We perform various analyses in an attempt to explain the difference. We find (indirect) support for the argument that the higher RTE for entrepreneurs is due to fewer (organizational) constraints faced by entrepreneurs when optimizing the profitable employment of their education.

JEL Classification: J23, J24, J31, J44, M13

Keywords: entrepreneurship, self-employment, returns to education employment

Corresponding author:

C. Mirjam van Praag
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: c.m.vanpraag@uva.nl

* We are grateful to Robert Fairlie and Rob Alessie for their excellent suggestions. The paper has further benefited from discussions at the 2005 Babson Kaufman conference in Boston, the 2005 SOLE/EALE conference in San Francisco, seminars at ESADE Business School in Barcelona, Utrecht School of Economics, Utrecht and a presentation at the closed workshop "Entrepreneurship and Human Capital", organized by the Amsterdam Center for Entrepreneurship and the Gate2Growth Academic Network of Entrepreneurship. The usual disclaimer applies.

1 Introduction

Entrepreneurs are considered by policy makers 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. The Lisbon strategy dictates the development of higher rates of innovation in Europe. Most European governments emphasize the development of entrepreneurship to reach this goal.[1] The value of entrepreneurship for societies is supported by research, see Van Praag and Versloot (2007).

However, the question remains how entrepreneurship can be optimally stimulated. There is one factor that both academic scholars and policy makers see as an important determinant of entrepreneurial performance, namely human capital. This study focuses on the measurement of the returns to human capital, in particular formal education, for entrepreneurs relative to employees.

The contribution of this study 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 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. With a fixed-effects IV-approach we check and confirm that the difference in returns to education between entrepreneurs an employees is not due to unobserved heterogeneity. Section 5 is de-

[1] See, for instance, *Working together for growth and jobs: A new start for the Lisbon Strategy*, communication from President Barroso 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).

voted 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.

2 Empirical literature

Almost a hundred empirical studies have measured the relationship between schooling and entrepreneurship outcomes.Van der Sluis, Van Praag and Vijverberg (2003) provide an overview and perform 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 study.

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 common practice (Ashenfelter, Harmon and Oosterbeek, 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

[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).

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 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 study, 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 study 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 study differs from both previous studies. In addition, we estimate the difference in the returns to education for entrepreneurs and employees by means of a fixed-effects IV-model such that unobserved heterogeneity across individuals does not affect the result. Last but not least, we provide 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 study.

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 Data

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

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

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

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.

In the literature, three different measures of entrepreneurs' incomes are compared to employees' incomes: (i) net profit; (ii) a periodic wealth transfer from the firm to the entrepreneur, much like a regular wage, labeled 'draw', and (iii) draw plus changes in the firm's equity value (Hamilton, 2000; Parker, 2004). Our measure falls in the second category, i.e., 'draw'. 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-

[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 1 shows the means and standard deviations of all the variables that are directly or indirectly used in the analyses. The values in Table 1 represent the averages of the specific

[7]Recent evidence suggests that entrepreneurs underreport their income more than employees do, see Lyssiotou, Pashardes and Stengos (2004); Feldman and Slemrod (2007). 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 administration of the ASVAB allows us to treat this variable as exogenous in all equations.

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.

Table 1: Summary of descriptive statistics

	Total sample		
Variable	**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 3.2 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]

3.2 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 (1) by means of a random effects (RE) as well as a fixed-effects (FE) model, such the observations that are distributed over years and individuals are used in an appropriate manner, i.e., yearly observations for one individual are not treated as independent. The FE model deals with unobserved heterogeneity and can only identify the *difference* in the returns to education based on the variance over time in occupational status of individuals, i.e., whether they are entrepreneurs or wage employees. This is due to the fact that formal education is (almost) time-invariant for labor market participants.[11]

$$W_{it} = \beta S_{it} + \gamma E_{it} + \delta SE_{it} + \eta X'_{it} + \zeta EX'_{it} + c_i + u_{it} \tag{1}$$

In equation (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 (where the subscript t could actually be omitted due to a lack of time variation, 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 1, as well as dummies controlling for cohort effects, age effects[12] and

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

[11] 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 (1).

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. The effects of variables included in the vector X'_{it} that are time-invariant cannot be estimated in the FE-model.

However, as was pointed out already, the variable S_{it} – the number of years of schooling – is likely to be endogenous in the income equation. Therefore, equation (1) is estimated and used as first stage equation in both the RE-model and the FE-model, i.e., the predicted values resulting from this equation is used as the instrumented values of education in equation (1).

$$S_{it} = \gamma F'_i + \lambda E_{it} + \varpi EF'_{it} + \kappa X'_{it} + \varsigma EX'_{it} + \mu_i + \epsilon_{it} \tag{2}$$

In addition, in the RE-model, we employ the same methodology for the possibly endogeneous variable E_{it} – being an entrepreneur – by means of equation (3):

$$E_{it} = \eta G'_i + \theta X'_{it} + \omega S_{it} + \vartheta_i + \varepsilon_{it} \tag{3}$$

Equations (2) and (3) represent the first-stage equations. In (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), entrepreneurship status is denoted by E_{it}, and its instruments by G'_i. X'_{it} is a vector of control variables as in (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

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

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 income via the education (entrepreneurship) 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 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 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

[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.

[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.

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 follow 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 that persons who have entrepreneurial fathers have a higher probability of becoming an en-

[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 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.

trepreneur 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 an 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.

The second route we propose uses the religion of the individual as an 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

[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.

[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.

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, the relationship between adhering to a strict religion and earnings is insignificant (Drakopoulou Dodd and Seaman, 1998, Table 4), thereby supporting the validity of such an instrument. 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.

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 (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 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 2) and then by using IV (the last two columns of the table).

Table 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 (2) are presented in the first two columns of Table 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

[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.

necessary – i.e., education is endogenous ($\chi^2_{(df1)} = 14.31$, and $p = 0.0002$).

Table 3: First stage equations of education and selection into entrepreneurship

	Education		**Selection (father)**		**Selection (religion)**	
Variable	**Coeff.**	**(SE.)**	**Coeff.**	**(SE.)**	**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 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. 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

[23]Using RE, without instrumenting education, the returns are 15 percent higher for entrepreneurs than for employees.

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

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.

Table 4: Earnings equations where both education and entrepreneurship are instrumented

	IV-education-entrepreneur(father)		**IV-education-entrepreneur(religion)**	
Variable	**Coeff.**	**(SE.)**	**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 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

[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.

[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.

father (and its interaction with age) as instruments, whereas the next set of columns shows the results of using religion as an identifying instruments for entrepreneurship. The two columns show similar results and the returns to education for employees and entrepreneurs are of similar magnitude 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.

Table 5: Fixed-effects earnings equation where education is instrumented

Variable	**Coefficient**	**(Std. Err.)**
E*Education	0.098**	(0.028)
E*General Ability	-0.022	(0.016)
E	0.086	(0.503)
Married	0.053**	(0.005)
Not Healthy	-0.041**	(0.014)
Live outside city	-0.046**	(0.008)
E*Male	0.458**	(0.028)
E*Married	-0.070**	(0.020)
E*Not Healthy	0.062	(0.049)
E*Live outside city	-0.018	(0.011)
E*Live in South	0.104**	(0.022)
E*Hispanic	-0.164*	(0.074)
E*Black	0.005	(0.051)
E*Locus of Control	0.001	(0.007)
E*Mother Education	-0.023*	(0.009)
E*Father Education	0.004	(0.006)
Average fixed effect	1.068**	(0.120)
N	50268	
R^2 Within	0.45	
R^2 Between	0.17	
R^2 Overall	0.32	
Significance levels :	† : 10%	∗ : 5% ∗∗ : 1%

E denotes Entrepreneur.

Since the instruments for entrepreneurship are far from perfect, we perform an additional analysis to assess whether the difference in the returns to education between entrepreneurs

and employees is influenced by unobserved heterogeneity. We estimate a fixed effects model where the coefficient of interest pertains to the interacted variable SE_{it} which can be estimated in the fixed-effects framework due to the variation over time in each individual's entrepreneurship status. Hence, the coefficient shows to what extent the returns to education of a specific individual increase or decrease when this person becomes an entrepreneur instead of a wage employee. Since a fixed-effects model does not deal with the endogenous nature of education we use the same IV approach as before.

The results from estimating the fixed-effects model strongly support the results from the previous analysis, see Table 5. The interaction-effect between education and entrepreneurship is again significant and even slightly larger than before. The returns to education for entrepreneurs are estimated to be almost ten percent higher than for employees. This supports our conclusion that being an entrepreneur rather than being of an "entrepreneurial type" increases the returns to education.

We conclude, based on our selectivity tests and the fixed effects regression, that correcting for selectivity leaves the key results unchanged. For convenience, the study 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 2. The discussion is based on the results reported in the last two columns of the table.

Table 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]

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.

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 2.

[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.

[28]To be discussed in detail later on.

5.1 Risk premium

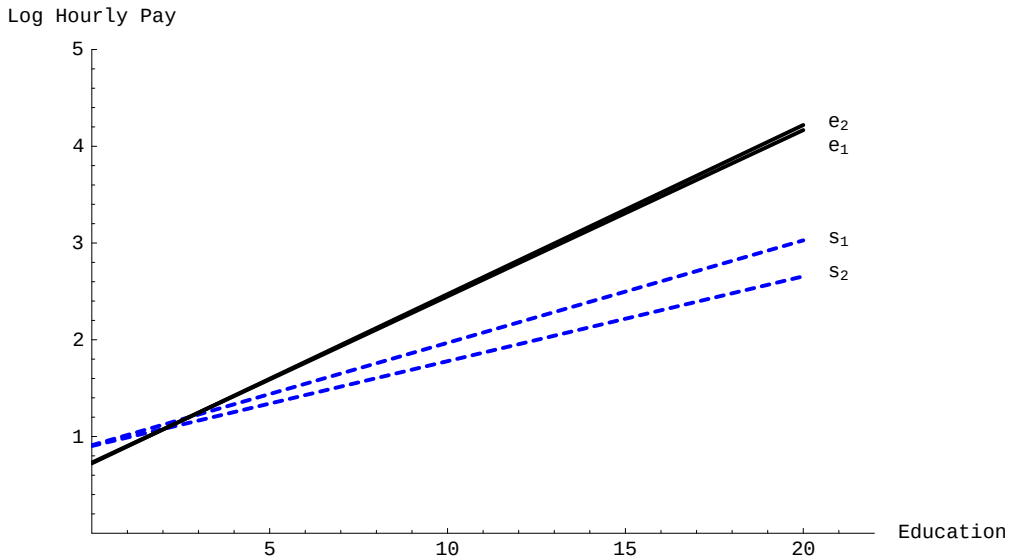
The first check relates to the question as to whether the difference in returns to education between entrepreneurs and employees can be attributed to a *risk premium* required by higher educated entrepreneurs. More highly educated individuals would perhaps require higher risk premia for being an entrepreneur 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.

5.2 Underreporting of earnings

The second check pertains to recent evidence that entrepreneurs underreport their incomes more than employees do (Feldman and Slemrod, 2007). 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 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 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. Only the *slopes* of the relationships shown are meaningful.

Figure 1: Returns to education: Blue- and white-collar employees and entrepreneurs

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 income. This could explain the result if more highly educated entrepreneurs have

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

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 a subtraction of five percent of their total business value from the business incomes of entrepreneurs with unincorporated businesses . As was indicated above, the variable 'total business value' has not been measured in every year, resulting in a smaller sample size. Therefore, Table A-1 in Appendix A shows the estimation results not only based on the adjusted entrepreneurial incomes, but also without this adjustment for the same sub-sample, allowing a proper comparison. The comparison indicates that the adjustment for capital returns does not at all reduce the difference in returns to education between entrepreneurs and employees.

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 A-2 in Appendix A.

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 A-2 in Appendix A shows, excluding

professional workers[30] from the sample does not decrease the estimated difference between entrepreneurs and employees.

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

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

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 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 6 and Figure 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

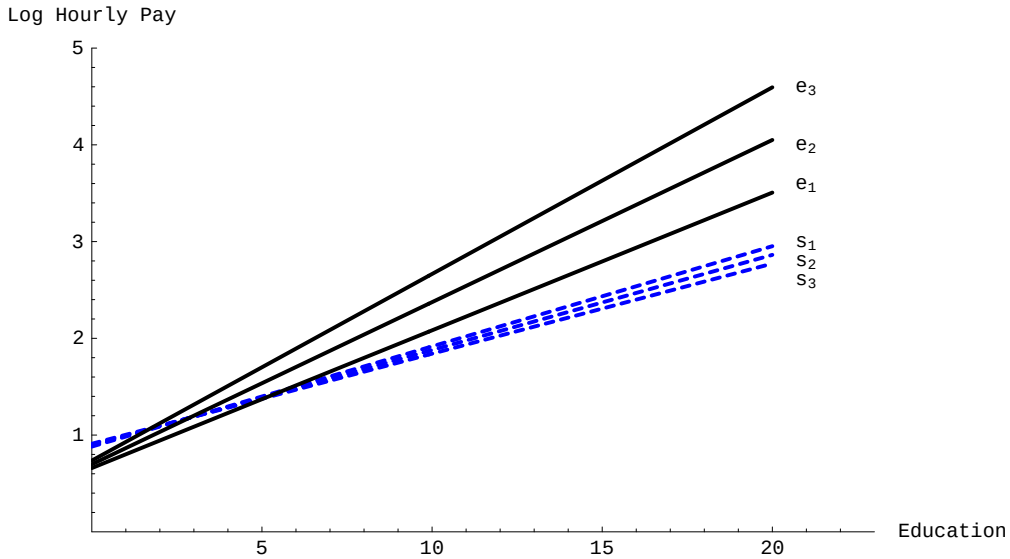
[31]The NLSY includes an abbreviated version of the original Rotter scale. The test was administrated in 1979, before the respondents had made any major decisions regarding their jobs or occupations. We re-scaled the Rotter score in such a way that 0 reflects high external locus of control and 10 indicates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

Table 6: 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.

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 2: Returns to education evaluated at three levels of locus of control

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. Altogether, 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 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

[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.

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 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 study 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.

References

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.

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.

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.

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

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.

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

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

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.

Feldman, N. and J. Slemrod (2007), 'Estimating tax non-compliance with evidence from unaudited tax returns', *Economic Journal* **117**(518), 327–352.

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

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.

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

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

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

Hausman, J. (1978), 'Specification tests in econometrics', *Econometrica* **46**(6), 1251–1271.

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

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.

Moore, R. (1983), 'Employer discrimination: Evidence from self-employed workers', *Review of Economics and Statistics* **65**(3), 496–501.

Oreopoulos, P. (2003), 'Do dropouts dropout too soon? International evidence from changes in school-leaving laws', *NBER Working paper 10155* .

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.

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.

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.

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.

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.

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., 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 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 P. Versloot (2007), 'What is the value of entrepreneurship? A review of recent research', *Small Business Economics* **29**(4), 351–.

Appendix A

Table 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 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.