TI 2009-088/4

Tinbergen Institute Discussion Paper

Education and Entrepreneurial Choice: An Instrumental Variables Analysis

Jörn H. Block[1,2,3,4]

Lennart Hoogerheide[1,4,6]

Roy Thurik[1,3,5,6,7]

[1] *Erasmus University Rotterdam, The Netherlands;*

2 Technische Universität München, Germany;

[3] *Centre for Advanced Small Business Economics; The Netherlands;*

[4] *Econometric Institute; The Netherlands;*

[5] *EIM Business and Policy Research, Zoetermeer, The Netherlands;*

[6] *Tinbergen Institute, The Netherlands;*

[7] *Max Planck Institute of Economics, Germany.*

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 Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Education and entrepreneurial choice: an instrumental variables analysis

Joern H. Block [a], Lennart Hoogerheide [b], Roy Thurik [c]

[a] Centre for Advanced Small Business Economics, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, block@ese.eur.nl; Technische Universität München, München, Germany.

[b] Econometric Institute and Tinbergen Institute, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, lhoogerheide@ese.eur.nl

[c] Centre for Advanced Small Business Economics, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands; EIM Business and Policy Research, P.O. Box 7001, 2701 AA Zoetermeer, the Netherlands and Max Planck Institute of Economics, Jena, Germany. thurik@ese.eur.nl.

Abstract: Education is argued to be an important driver of the decision to start a business. The measurement of its influence, however, is difficult since it is considered to be an endogenous variable. This study is the first to account for this endogeneity by using an instrumental variables approach. The effect of education on the decision to become self-employed is found to be strongly positive, much higher than the estimated effect in case no instrumental variables are used. That is, the higher the respondent's level of education, the greater the likelihood that he or she starts a business. Implications for method and practice are discussed.

Current version: October 2009

JEL codes: C35, I20, J24, L26

Keywords: Occupational choice, entrepreneurial choice, education, self-employment, endogeneity, instrumental variables, entrepreneurship

Acknowledgements: This paper has been written in cooperation with the research program SCALES, carried out by EIM and financed by the Dutch Ministry of Economic Affairs.

Corresponding author: Joern Block, Rotterdam School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, block@ese.eur.nl, (tel. +31104081004, fax. +31104089141).

Introduction

The effect of education on entrepreneurial choice and performance is widely researched (Van der Sluis et al., 2008). Policy-makers are particularly interested in its effect since education can be influenced by educational and other policy measures (European Commission, 2003; OECD, 2009). Establishing its effect, however, is difficult due to endogeneity (Griliches and Mason, 1972; Blackburn and Neumark, 1993; Ashenfelter et al., 1999). That is, education appears as a causal variable in an econometric model but is in fact correlated with the errors in the model. As a result, the effect of education may be under- or overestimated.

In these situations the use of instrumental variables regressions (IV regressions) is a solution to isolate the causality (Angrist and Krueger, 1991; Angrist et al., 1996). Using such IV regressions, Parker and Van Praag (2006) find that education is indeed endogenous to *entrepreneurial performance* and that it makes a difference whether or not IV methods are used. So far, however, there exists no study using IV regressions to analyze the effect of education on *entrepreneurial choice*. This is surprising, since entrepreneurial choice is widely examined in the literature (Evans and Jovanovic, 1989; Grilo and Thurik, 2008; Parker, 2009). This study is a first attempt.

There are two main groups of candidate instruments for education: family background variables and natural experiment variables such as changes or differences in compulsory schooling laws (Angrist and Krueger, 1991; Webbink, 2005; Hoogerheide et al., 2007). In the present study, we rely on the first category and use the social class of the parents as an instrument. Our data set comprises of more than ten thousand individuals from 27 European countries and the US, who are either self-employed or in a paid employment job. We obtain two important findings: *first*, the effect of education on the decision to become self-employed is found to be strongly positive. The higher the respondent's level of education, the greater the likelihood that (s)he starts a business. *Second*, our results show that a standard Probit or Logit model underestimates the effect of education on entrepreneurial choice and leads to biased results. We suggest that this is exactly the reason why many earlier studies have found weak or insignificant results (Van der Sluis et al., 2008). The underestimation (under the assumption of no endogeneity) of the effect of education on the choice to become self-employed is also in line with the underestimation of the OLS estimator for the effect of education on wage (Angrist and Krueger, 1991).

Data and Method

To analyze the effect of education on entrepreneurial choice, we use data from the 2007 Flash Eurobarometer Survey on Entrepreneurship. The dataset has been used in a number of published studies (Grilo and Irigoyen, 2006; Grilo and Thurik, 2008; Van der Zwan et al., 2009) and contains detailed information on the respondents' employment status. We restrict the sample to those participants who are either self-employed or in a paid employment job (10,962 obs.). We excluded respondents with solely domestic activities (1,678 obs.) or searching for a job (632 obs.), students (1,443 obs.), retirees (5,242 obs.), and respondents who refused to give an answer or do not fall in any of these categories (717 obs.). We lose some further observations due to missing values. The final dataset contains 10,397 observations.

Our dependent variable is a dummy variable, which indicates whether the participant is self-employed or not. Education is measured as the number of years which the participant spent on education. We include a number of commonly used control variables in the regression model such as gender or job experience (Grilo and Thurik, 2008). We also controlled for country effects. Table 1 (see Appendix A) describes the construction of the variables; Table 2 shows correlations and descriptive statistics.

To analyze the effect of education on the decision to become self-employed, we estimate both a standard Probit model and an IV Probit model. As instruments, we use the social class of the parents (e.g., blue collar vs. white collar). The IV model is estimated to account for the above discussed endogeneity issue associated with the education measure (Angrist and Krueger, 1991; Angrist et al., 1996). We test the validity of the instruments with the Amemiya-Lee-Newey minimum chi-square statistic (Amemiya, 1978; Newey, 1987; Lee, 1992). The null hypothesis of valid instruments is not rejected ($p=0.146$).

Results

Table 3 shows the regression results. The results regarding the effect of education on entrepreneurial choice are clear-cut: both in the standard Probit model and in the IV model, a positive effect can be recorded. There seems to be a positive influence of education regarding the decision to start a business. The IV model however shows a much stronger effect ($\beta=0.014$ in the standard Probit model[1]; $\beta=0.137$ in the IV model)[2]. This strong difference in the size of the effects is explained by the fact that education is endogenous to entrepreneurial choice: estimating a standard Probit model underestimates the 'true' effect. The Wald-test of exogeneity is highly significant. There are two possible reasons for the negative bias in the standard Probit model. *First*, there may exist omitted variables such as cognitive ability that have both a positive influence on education level and a negative effect on the decision to become self-employed. *Second*, years of education may be a poor proxy for the level of education[3]; then the measurement error drives the estimate for education in the standard Probit model towards zero or insignificance.

The results regarding the control variables are as expected (Grilo and Thurik, 2008). For example, male respondents have a higher likelihood of falling into the self-employment category (IV model: $\beta=0.388$, $p<0.001$). The effect of labor market experience is positive in its linear term and negative in its squared term. Country effects are important. An F-test on joint significance of the country variables shows a significant result.

Discussion

The advent of the knowledge driven economy together with the recognition that such an economy requires a prominent entrepreneurial sector (Audretsch and Thurik, 2001; Audretsch, 2007) produced many studies regarding the effect of education on entrepreneurial choice and performance (for a summary, see Van der Sluis et al., 2008). Moreover, of the many factors known to influence entrepreneurial choice and performance (Grilo and Thurik, 2008; Parker, 2009) education is popular among politicians since it can be influenced. Our study contributes to this literature by estimating an IV model to explain the causal effect of education on entrepreneurial choice. We show that education appears to be an endogenous variable regarding the decision to become self-employed, which is why an IV model is needed to estimate its effect. Using such a model, we then show that a higher level of education increases the likelihood of becoming self-employed.

---

1 Using a standard Logit model yields $\beta=0.024$ ($p<0.001$).

2 The respective marginal effects are $\beta=0.003$ in the standard Probit model and $\beta=0.023$ in the IV model. Hence, an additional year of education increases the probability of becoming self-employed by 0.3% in the standard Probit model and 2.3% in the IV model. For the calculation, all dummy variables are set at zero (the modal value) and all continuous variables are set at their sample mean.

3 For example, years of education as a measure does not account for the *quality* of education.

These two main results have a number of important implications for both method and practice: *first*, our results show that a standard Probit model should not be used to estimate the effect of education, since it tends to underestimate the effect of education. An IV approach is needed to find the 'true' effect. In that respect, entrepreneurial choice does not differ from other educational outcome variables such as wage (Angrist and Krueger, 1991; Card, 2001; Webbink, 2005). *Second*, the popularity among politicians to promote education as an important driver of economic growth is supported by the 'second order' effect that education promotes entrepreneurship which itself is a driver of economic growth (Audretsch et al., 2006).

The results of our paper offer several interesting avenues for further research. One avenue would be to analyze whether a higher level of education increases the preference for self-employment as a means to obtain non-monetary benefits (e.g., more flexibility or independence) or whether more education increases the economic returns from self-employment. Moreover, it would be interesting to analyze whether the positive effect of education on entrepreneurial choice holds for all types of entrepreneurs alike (e.g., necessity versus opportunity entrepreneurs).[4]

[4] See Block and Wagner (2010) or Block and Sandner (2009) for a discussion of necessity versus opportunity entrepreneurship.

References

Amemiya, T., 1978. The estimation of a simultaneous equation generalized probit model. Econometrica 46(5), 1193-1205.

Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Quarterly Journal of Economics 106(4): 979-1014.

Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91(434): 444-455.

Audretsch, D.B., 2007. Entrepreneurship capital and economic growth, Oxford Review of Economic Policy, 23(1): 63-78.

Audretsch, D.B., Keilbach, M.C., Lehmann E.E., 2006. Entrepreneurship and Economic Growth, New York: Oxford University Press.

Audretsch, D.B., Thurik, A.R., 2001. What is new about the new economy: sources of growth in the managed and entrepreneurial economies, Industrial and Corporate Change 10(1): 267-315.

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

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

Block, J., Sandner, P. 2009. Necessity and opportunity entrepreneurs and their duration in self-employment. Journal of Industry, Competition and Trade, 9(2), 117-137.

Block, J., Wagner, M. 2010. Necessity and opportunity entrepreneurs in Germany: characteristics and earnings differentials. Schmalenbach Business Review, forthcoming.

Card, D., 2001. Estimating the returns to schooling: progress on some persistent econometric problems. Econometrica 69(5): 1127-1160.

European Commission, 2003. Green Paper: Entrepreneurship in Europe. Brussels, Directorate General Enterprise and Industry.

Evans, D.S., Jovanovic, B., 1989. An estimated model of entrepreneurial choice under liquidity constraints. Journal of Political Economy 97(4): 808-827.

Griliches, Z., Mason, W.M., 1972. Education, income, and ability. Journal of Political Economy 80(3): S74-S103.

Grilo, I., Irigoyen, J.M., 2006. Entrepreneurship in the EU: to wish and not to be. Small Business Economics, 26(4): 305-318.

Grilo, I., Thurik, A.R., 2008. Determinants of entrepreneurial engagement levels in Europe and the US. Industrial and Corporate Change, 17(6): 1113-1145.

Hoogerheide, L.F., Kleibergen, F., Van Dijk, H.K., 2007. Natural conjugate priors for the instrumental variables regression model applied to the Angrist–Krueger data. Journal of Econometrics 138(1), 63-103.

Lee, L.-F., 1992. Amemiya's generalized least squares and tests of overidentification in simultaneous equation models with qualitative or limited dependent variables. Econometric Reviews 11(3), 319-328.

Newey, W.K., 1987. Efficient estimation of limited dependent variable models with endogeneous explanatory variables. Journal of Econometrics 36, 231–250.

OECD, 2009. Education at a Glance. Paris

Parker, S.C., 2009. The economics of entrepreneurship, second edition. Cambridge University Press, forthcoming.

Parker, S.C., Van Praag, C.M., 2006. Schooling, capital constraints and entrepreneurial performance: the endogenous triangle. Journal of Business and Economic Statistics, 24(4): 416-431.

Van der Zwan, P., Thurik, A.R., Grilo, I., 2009. The entrepreneurial ladder and its determinants. Applied Economics, forthcoming.

Van Praag, C.M., Vijverberg, W., Van der Sluis, J., 2008. Education and entrepreneurship selection and performance: a review of the empirical literature. Journal of Economic Surveys, 22(5): 795-841.

Webbink, D., 2005. Causal effects in education. Journal of Economic Surveys, 19(4): 535-560.

Appendix A: Tables

Table 1: Description of variables

Variable	Description
Dependent variable	
Self-employment	Dummy = 1 if respondent is self-employed
Variable of interest	
Education	Number of years the respondent has been in full-time education
Instruments [1]	
Father was/is white collar	Dummy = 1 if father of respondent had/has a white collar job
Father was/is blue collar	Dummy = 1 if father of respondent had/has a blue collar job
Father was/is civil servant	Dummy = 1 if father of respondent was/is civil servant
Father was/is without professional activity	Dummy = 1 if father of respondent was/is without professional activity
Mother was/is white collar	Dummy = 1 if mother of respondent had/has a white collar job
Mother was/is blue collar	Dummy = 1 if mother of respondent had/has a blue collar job
Mother was/is civil servant	Dummy = 1 if mother of respondent was/is civil servant
Mother was/is without professional activity	Dummy = 1 if mother of respondent was/is without professional activity
Control variables	
Labor market experience	Age of the respondent minus age when stopped full time education
Male	Dummy = 1 if respondent is male
Father was/is self-employed	Dummy = 1 if father of respondent was/is self-employed
Mother was/is self-employed	Dummy = 1 if mother of respondent was/is self-employed
Rural region	Dummy = 1 if respondent lives in a rural region
Metropolitan region	Dummy = 1 if respondent lives in a metropolitan region
Country dummies	28 Country indicator variables (Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, UK, US)

Note: [1] The instruments do not sum up to 1, since the response categories 'father/mother was/is self-employed' and 'don't know/no answer' are not used as instruments.

Table 2: Descriptive statistics and correlations

		Min.	Max	Mean	Median	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Self-employment	0	1	0.20	0																
2	Education	1	75	13.66	12	-0.01															
3	Labour market experience	0	81	23.02	23	0.11	-0.42														
4	Male	0	1	0.44	0	0.14	0.01	0.02													
5	Father was/is self-employed	0	1	0.24	0	0.16	0.02	0.04	0.03												
6	Mother was/is self-employed	0	1	0.10	0	0.10	-0.00	0.02	-0.01	0.41											
7	Rural region	0	1	0.35	0	0.04	-0.12	0.05	-0.01	0.05	0.05										
8	Metropolitan region	0	1	0.23	0	-0.01	0.14	-0.05	0.01	-0.04	-0.03	-0.40									
9	Country is US	0	1	0.05	0	0.05	0.12	0.00	0.05	0.00	-0.01	-0.02	0.09								
	Instruments																				
10	Father was/is white collar	0	1	0.17	0	-0.02	0.07	-0.08	-0.01	-0.25	-0.10	-0.03	0.05	0.05							
11	Father was/is blue collar	0	1	0.33	0	-0.09	-0.10	0.05	-0.00	-0.39	-0.16	0.04	-0.05	-0.01	-0.32						
12	Father was/is civil servant	0	1	0.18	0	-0.03	0.05	-0.02	-0.01	-0.26	-0.11	-0.09	0.05	-0.05	-0.21	-0.33					
13	Father was/is without professional activity	0	1	0.04	0	-0.01	-0.04	-0.02	-0.01	-0.12	-0.05	0.02	-0.02	-0.01	-0.10	-0.15	-0.10				
14	Mother was/is white collar	0	1	0.14	0	-0.02	0.09	-0.17	-0.03	-0.07	-0.14	-0.02	0.06	0.06	0.31	-0.10	-0.08	-0.06			
15	Mother was/is blue collar	0	1	0.19	0	-0.07	-0.07	0.01	-0.03	-0.16	-0.16	0.01	-0.01	0.00	-0.10	0.36	-0.12	-0.06	-0.19		
16	Mother was/is civil servant	0	1	0.14	0	-0.04	0.08	-0.11	-0.03	-0.11	-0.14	-0.08	0.07	-0.05	-0.06	-0.13	0.38	-0.06	-0.16	-0.19	
17	Mother was/is without professional activity	0	1	0.41	0	0.03	-0.06	0.17	0.06	0.02	-0.28	0.03	-0.06	-0.00	-0.03	-0.00	-0.04	0.16	-0.33	-0.40	-0.33

Notes: N=10,397 obs.; All correlations above r=0.02 have a p-value less than 0.05; We checked estimation results omitting 'outliers' (e.g. observations with education over 30 years): changes in results are minor.

Table 3: Results of standard Probit regression and instrumental variables Probit regression
Dependent variable: Individual is self-employed

	Standard Probit Regression		Instrumental Variables Probit Regression (two step) [1]	
Variables	**Coefficient (SE)**		**Coefficient (SE)**	
Education [a, b]	0.014 (0.003)	***	0.134 (0.030)	***
Labour market experience/10	0.137 (0.038)	***	0.475 (0.094)	***
(Labour market experience/10)^2	0.001 (0.010)		-0.023 (0.010)	*
Male	0.392 (0.030)	***	0.388 (0.032)	***
Father was/is self-employed	0.347 (0.037)	***	0.290 (0.042)	***
Mother was/is self-employed	0.178 (0.050)	***	0.203 (0.054)	***
Rural region [c]	0.169 (0.035)	***	0.279 (0.046)	***
Metropolitan region [c]	0.051 (0.040)		-0.042 (0.049)	
Country dummies [d]	27 categories (p<0.001)		27 categories (p<0.001)	
Intercept	-1.999 (0.151)	***	-3.945 (0.512)	***
N	10,397		10,397	
Minus Log pseudolikelihood	4656.61			
Pseudo R^2	0.083			
Wald Chi² (df)	765.11 (35) ***		683.44 (35) ***	

* p<0.05 ** p<0.01 *** p<0.001
SE=Robust standard errors (standard Probit regression)

Notes:
[a] Instruments for education: 'father was/is white collar', 'father was/is blue collar', 'father was/is civil servant', 'father was/is without professional activity', 'mother was/is white collar', 'mother was/is blue collar', 'mother was/is civil servant', 'mother was/is without professional activity' (F-test for significance of the instrument: F(8, 10,392)=39.56***; R^2=0.026
Wald-test of exogeneity: p<0.001
Validity of the instruments: Amemiya-Lee-Newey minimum chi² statistic: 10.837 (p=0.146)
[b] When excluding outliers (education is more than 30 years), the coefficients are β=0.131 *** (IV model) and β=0.014 *** (Standard Probit Model). We also tested for a non-linear effect of education on entrepreneurial choice but found no evidence of such an effect.
[c] Reference category is 'other town/urban centre'.
[d] Reference category is 'US'.