

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Tokila, Anu; Tervo, Hannu

Title: Regional differences in returns to education for entrepreneurs versus wage earners

Year: 2011

Version:

Please cite the original version:

Tokila, A., & Tervo, H. (2011). Regional differences in returns to education for entrepreneurs versus wage earners. *The Annals of Regional Science*, 47(3), 689-710.
<https://doi.org/10.1007/s00168-010-0374-8>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Regional differences in returns to education for entrepreneurs vs. wage earners

1 Introduction

Human capital theory predicts that education augments an individual's abilities in the labour market, thereby enhancing his/her productivity and earned income. Due to labour and product market conditions, returns to education may differ across regions. In the New Economic Geography literature, as well as in the earlier regional science literature, agglomeration economies are supposed to lead to higher productivity of workers and hence higher wages and higher returns to education (Fujita et al. 1999; see also de Blasio and Di Addario 2005; Dalmazzo and de Blasio 2007; Lehmer and Möller 2009). Marshall (1890) was the first to present that spatial agglomerations create location advantages in terms of spillovers, labour pooling and co-operation between firms. Agglomeration economies may also enhance the earnings and returns to education for entrepreneurs. In addition, variation in relative returns to education across regions may account for the prevailing regional differences in entrepreneurship, which are pronounced and persistent in many countries (e.g. Bernhardt 1994). On the other hand, labour and product market conditions possibly will also account for the rate entrepreneurship (Georgellis and Wall 2000; Parker 2004). Especially less-educated individuals in a small, dispersed labour market may be pushed into self-employment if they see no other realistic option (Carrasco and Eijnán 2003; Tervo 2008).

Apart from enhancing earnings, it is also generally supposed that human capital, including education, increases an individual's probability of becoming self-employed (e.g. Rees and Shah 1986). Thus, education can potentially be seen as a tool for supporting entrepreneurship. In order to be an attractive choice, an earnings premium may be necessary to compensate for the greater uncertainty and higher inherent risk of entrepreneurship (Storey 1994; Hamilton 2000). If educated entrepreneurs do not achieve an adequate premium for their risk in a region, entrepreneurship may not be an attractive occupational choice, whereas high expected earnings will pull educated individuals into entrepreneurship. In many countries and regions, however, individuals

with a higher level of education have a lower probability of being self-employed. This is also the case in Finland (Johansson 2000; Niittykangas and Tervo 2005), which has one of the most expensive education systems in the world. This prompts the question as to why highly-educated individuals do not choose entrepreneurship, though they have better abilities. Is it not remunerative for them? In order to be an attractive choice, entrepreneurship should yield an adequate earnings premium for the higher inherent risk it entails (Storey 1994).

The performance of educated entrepreneurs can be studied by comparing their returns to those of wage earners. No clear assumption, on either theoretical or empirical grounds, can be made about whether the returns to education for entrepreneurs will be higher or lower. Many previous studies suggest that the rates of return to schooling are lower for entrepreneurs than for wage earners (Brown and Sessions 1998 & 1999; Hamilton 2000; García-Mainar and Montuenga-Gómez 2004), although the opposite has been reported (Evans and Leighton 1990; Robinson and Sexton 1994; Alba-Ramírez and San Segundo 1995; van der Sluis et al. 2004). The impact of education on income varies according to the labour market and levels of education. Van der Sluis et al. (2004) observed that studies pertaining to Europe more often indicated lower returns for entrepreneurs, while the evidence from the United States was the opposite. The study by Iversen et al. (2006) found that higher levels of schooling resulted in larger returns for the self-employed, while lower levels of education indicated hardly any return in self-employment. In contrast, García-Mainar and Montuenga-Gómez (2004) concluded that secondary education is the most profitable choice for the self-employed.

In this paper, we are interested in regional variation in returns to education between entrepreneurs and wage earners. The role of regions for general wage differentials has been studied to some extent (e.g. Dumond et al. 1999; Duranton and Monastiriotis 2002; Bernard et al. 2003; Goetz and Rupasingha 2004), but regional comparison between entrepreneurs and wage earners has been largely ignored in the literature. We ask: does the education attained by an individual imply a higher rate of return for the entrepreneur than for the wage earner and are there regional differences in these returns? Typically, studies on regional wage differentials analyze specific regions. We, instead, focus on types of regions in order to provide more generalized results. In our analysis, Finnish regions are classified into urban and non-urban regions on the basis of

their number of inhabitants and population density. These regions show different rates of self-employment, the rate being higher in rural areas and lower in urban areas.

The analysis is based on a rich register-based data set which includes a 7 per cent random sample of all Finns. In the analysis, we apply Mincer-type income equations which are estimated separately for employees and the self-employed (cf. Parker 2004). Different estimation procedures are used to avoid potential sources of bias in the results. To avoid selection bias, we apply Heckman's (1979) method. For the instrumentation of the possible endogenous education variable, we also use a set of family background variables as identifying instruments. For the sake of comparison, the standard OLS-method is also applied.

Our estimation results suggest that returns to education are much the same for entrepreneurs and wage earners in the entire country. No clear-cut premium for entrepreneurs is discovered except in rural areas, where the return especially to higher education, is much higher for entrepreneurs. The estimated returns to higher education for entrepreneurs remain smaller in urban areas than in rural areas. Nevertheless, well-educated individuals in rural areas do not opt for entrepreneurship, although the general level of entrepreneurship is high. This result suggests that it is the push rather than the pull effect which accounts for regional differences in entrepreneurship. It is more likely that regional variation in entrepreneurship is due to weak employment conditions than to higher expected earnings.

The rest of the paper is organized as follows. Theoretical issues are discussed in Section 2. The model, data and variables are presented in Section 3 including basic definitions and information on the income and educational choices in the two different types of regions. In section 4 the estimation results for the entire data set are presented first, followed by the results for the regions. Section 5 concludes.

2 Theoretical issues

The well-known human capital theory holds that education augments an individual's abilities of in the labour market, thereby enhancing his/her productivity and earned

income (Mincer, 1974; Becker, 1974). Accordingly, for both entrepreneurs and wage earners, having more education should be rewarded in the labour market. Higher education improves several abilities needed in business such as risk-awareness and comprehension of market prospects (Kangasharju and Pekkala, 2002). However, the returns of the self-employed may be more dependent on factors other than formal education. Many entrepreneurial skills such as motivation and salesmanship are non-academic in nature (Parker, 2004). This may lead to lower returns to education for entrepreneurs. In the case of wage earners, part of their return to productivity may be exploited by the firm, thus lowering their returns to education (García-Mainar and Montuenga-Gómez, 2004).

Human capital can be divided into general and specific human capital (Becker, 1975). Education and work experience measured in years or levels represent general human capital. In the context of self-employment, specific human capital is distributed as industry-specific and entrepreneurship-specific human capital. Industry-specific experience in paid work increases the productivity of entrepreneurs as they already have familiarity with the main activities of the industry. In addition, it may yield knowledge about potential niches in business. Entrepreneur-specific human capital can best be obtained through self-employment, although some entrepreneurial skills may also be acquired through special entrepreneurial training (Brüderl, 1992; Firkin, 2003).

Spence (1973) was the first to claim that greater human capital is acquired only in order to signal of inherent productivity. The weak screening hypothesis concedes that the effect of education is manifested both through signalling and through increased productivity (Spence, 1973; Arrow, 1973). In the strong screening hypothesis education operates merely as a signal of inherent productivity, having no role in its enhancement (Psacharopoulos, 1979). In the case of the self-employed, the signalling role of education is not that clear, as they employ themselves. For wage earners as well, more evidence has been shown for the weak hypothesis than for the strong version (e.g. Brown and Sessions, 1998 & 1999). Thus, the idea that education has in an important role in the generation of earnings gains further support.

Earnings are an important incentive in occupational choice. Self-employment and paid work represent the two largest sources of income; however their generation,

characteristics and degree of inequality are fundamentally different (Parker, 1999). In most countries, the earnings of entrepreneurs range more widely than those of employees. This means that a relatively larger number of the self-employed are concentrated in the lower and upper tails of the income distribution compared to wage earners (Parker, 2004). This is also the case in Finland (Tervo and Haapanen, 2009). Comparisons between the mean income of entrepreneurs and paid workers may be strongly influenced by a few very high-income entrepreneurial superstars (Rosen, 1981; Hamilton, 2000). For this reason it might be better to use median income in the comparative analyses.

In some studies, rural areas have been found to suffer from substantially lower rates of returns to education than urban areas (Goetz and Rupasingha, 2004). People may accept a lower income in exchange for other amenities in a region, such as scenery (see e.g. Power and Barrett, 2001). On the other hand, for individuals with high qualification the gap in regional rates has been found to be relatively low (Bennett et al., 1995). In peripheral regions, jobs with higher educational demands are more comparable in rates of pay with jobs in central areas, while for lower-skill jobs rates are determined more by the local labour market. Yet, very little is known about regional variation in returns to education between self-employment and paid work. Studies on regional wage differentials are commonly based on aggregate level approaches, which are incapable of yielding explanations or sources for regional inequalities in returns to education (see e.g. Duranton and Monastiriotis, 2002).

3 Method, data and definitions

Methodology

It is important that the effect of education (formal schooling) on entrepreneurial performance is measured consistently. We apply the income equations first suggested by Mincer (1974) to estimate causal effects that are not biased due to the neglect of unobserved heterogeneity or the endogenous nature of the occupational choice (employment/self-employment). According to the Mincerian approach, the main

determinants of individual earnings are schooling and experience. The general specification for a Mincer-type earnings equation is:

$$(1) \quad \ln Y_i = f(\text{Edu}_i) + bX_i + u_i,$$

where Y_i is the annual earnings of individual i , $f(\text{edu}_i)$ is a function of the educational attainment of individual i , X_i contain labour market experience and other characteristics of the individual, and u_i is a random error. In the standard version, $f(\text{edu}_i)$ is years of education, i.e. it is assumed that the logarithm of earnings is a linear function of years of completed education. In addition to this linear approach, we consider a specification in which we use a dummy to compare the effect of having higher education with the effect of not having higher education.

Since we are estimating earnings functions separately for entrepreneurs and wage earners, it is necessary to take account of possible selection bias. Those who enter self-employment or paid employment may not be a random sample from the population. A common way of removing possible selection bias is Heckman's (1979) method in which the estimation is performed in two steps. In the first step, the selection (participation) equation is estimated:

$$(2) \quad z_i = w'W_i + v_i,$$

where z_i indicates whether individual i is an entrepreneur/wage earner, W_i is a set of explanatory variables and v_i is a disturbance term with unit variance. The selection equation should contain at least one variable that is not in the outcome equation. In our case, several variables describing family background and regional features are used. Computing fitted values for z_i yields the 'inverse Mills ratios':

$$(3) \quad I_i = -f(\hat{z}_i)/F(\hat{z}_i),$$

which are added to the earnings equation in the second step of the estimation.

Another bias may follow from the endogenous nature of education. For this reason, many previous studies on returns to education for entrepreneurs are potentially biased (van der Sluis et al., 2004). Their methodological approaches lag behind especially in terms of identification. A cross-sectional correlation between education and earnings may not reflect the true causal effect of education (Card, 1999). Generally, OLS estimates of the returns to education are biased downwards (Ashenfelter et al., 1999). These problems have long received attention in other types of studies on returns to education. A standard solution to the problem of identification is the instrumental variables (IV) estimation, where identifying instruments must satisfy two conditions: (i) the instrument should not be correlated with the error term and (ii) the instrument should be highly correlated with the endogenous variable of interest. The former relates to the validity of the instrument and the latter to its quality. Family background is a widely used instrument for education. In our case, a set of variables for parents' education are used to instrument the schooling choice of an individual. Estimations are performed using two-stage least squares (2SLS), in which the first stage is to estimate the structural form of the earnings equation and the second stage the reduced form of the equation with instrument variables. Two recent studies suggest that correcting for the endogeneity of schooling substantially increases the estimated rate of return for entrepreneurs (Parker and van Praag, 2004; van der Sluis et al., 2004). To confront both types of biases in the same model, we use a combined Heckman-IV approach in which the inverse Mills ratios are included in the IV regression.

The data set

The data set is based on the Longitudinal Census File and the Longitudinal Employment Statistics File constructed by Statistics Finland. These two register-based data sets have been updated annually since 1987. They, together with some other registers, provide panel data on each resident of Finland. A seven per cent random sample of all the individuals in this data set in 2001 was taken for this study. The data set includes very rich information on individuals' educational attainments, labour market performance, regional and family characteristics and many other variables from the period 1987-2002 and from the earlier years 1970, 1975, 1980 and 1985. We analyze the rates of return in the base year 2001, as this data set for this year contains the most comprehensive information on earnings.

In this paper, the sample is restricted to persons who were 18-64 years old, in paid- or self-employment and who has a net earned income in 2001. Persons employed in the agricultural sector are excluded from the sample due to the special nature of agricultural entrepreneurship (see e.g. Blanchflower, 2000; Parker, 2004). The total number of individuals in the sample is 148 861, of which 11 528 (7.7%) are self-employed and 137 333 (92.3%) wage earners.

Definitions of entrepreneur and income

The first task, often a problematic one, in studies of entrepreneurship is to define the concept of 'entrepreneur'. Self-employment and entrepreneurship are concepts used in widely diverse ways (see e.g. Parker, 2004). In this analysis, the information of employment status comes from the Longitudinal Census File. Thus, the concept of entrepreneurship directly follows from the statistical definitions used by Statistics Finland (see Statistics Finland, 2001). The data on employment status are based on the person's national insurance status and type of income and thus describe whether a person is wage earner or entrepreneur. The third category, the non-employed, is not considered in this study. Entrepreneurs are defined as persons who have a self-employed person's pension insurance during the last week of the year, and whose income from entrepreneurship exceeds a specified level of earnings. This threshold is set inferentially using data from the Labour Force Survey (for details, see Statistics Finland 2001). The category of entrepreneurs also includes unpaid family workers.

The data give the highest level of education completed by the individual. A shortcoming is that no previous same-level degrees or uncompleted degrees are recorded. Our data set shows that wage earners are more educated than entrepreneurs (Table 1). The share of persons with only basic education is statistically greater in the group of entrepreneurs, whereas wage earners more often have a level of education beyond secondary. In the empirical analysis, higher education (master, doctorate or equivalent level) is used as the baseline in the analysis of education levels.

Table 1. Education levels for entrepreneurs and wage earners

	Entrepreneurs (%)	Wage earners (%)
General education	29.5	20.2
Secondary education	44.0	43.7
Lower tertiary education	20.0	25.8
Higher education	6.5	10.3

Another problem concerns the definition and measurement of income, as these are not unambiguous for entrepreneurs. Analysis of self-employment income data should be performed with caution for several reasons (Parker, 2004). Entrepreneurs may underestimate their income for taxation reasons. Owners of incorporated firms may be classified in different ways – employees or employers. Survey-based data sets suffer from high non-response rates from entrepreneurs. Furthermore, measuring employee fringe benefits and elements of entrepreneur income is problematic.

Due to its official nature, the data used here are reliable. The measurement of income is based on the tax files of the National Board of the Inland Revenue, which offers an excellent starting point for an accurate analysis of earnings. Hence, problems related to survey data collection or classification are not notable in our analysis. Income underreporting and other kinds of undeclared income, however, remain unresolved problems.

The dependent variable is defined as the logarithm of the individual's annual gross income minus transfer payments. The transfer payments in question are unemployment benefit and daily, maternity, and home care allowances. Gross income includes both wage income and entrepreneurial income as well as social security benefits. In order to apply logarithmic transformation, some zero and negative observations were dropped from the data set. The number of these observations (75 in total, 63 entrepreneurs), however, has a negligible effect on the results. Due to data protection legislation, the highest percentile in income subject to state taxation is given as a mean in the data. Should this not have a significant influence on the results either.

Regional classification

The analysis is based on the comparison of urban and rural areas. The basic regional unit used is a statistical grouping of municipalities (location of workplace) into three groups. In 2001 the number of municipalities in Finland was 448. Urban municipalities are defined as those where at least 90 per cent of the population lives in an urban settlement, or where the population of the largest urban settlement is at least 15,000. All other municipalities are defined as rural, despite the fact that the original classification separates semi-urban municipalities from rural municipalities. Thus, the two groups of semi-urban and rural municipalities are combined in our analysis. For the purposes of international comparisons, Finnish semi-urban municipalities are more rural than urban. Also according to Labrianidis (2006), smaller size towns are more suited for inclusion in rural areas in the regional analysis.

Distinct differences are found in regional entrepreneurial activity by educational classes (Table 2). Overall intensity of entrepreneurship is more than double in rural areas compared to urban ones. This can be explained by the fact that they offer fewer wage earning opportunities than urban areas (Tervo, 2008). It is interesting to note that entrepreneurial activity among the highly educated shows much less variation. For highly educated individuals, self-employment may be a less attractive choice than employment (Kangasharju and Pekkala, 2002). There are several possible reasons for this. According to some studies, highly educated persons would earn more as paid workers than as entrepreneurs (e.g. Hamilton, 2000; Uusitalo, 2001). In addition, compared to employees entrepreneurs have a much less constant income stream and less security than do employees (Storey, 1994). Instead, the non-pecuniary benefits of the job are often more important than financial benefits to the self-employed (Hamilton, 2000).

Table 2. Self-employment rate in the regions

	General education	Secondary education	Lower tertiary education	Higher education	All
Urban areas	8.3	5.8	5.0	4.8	6.0
Rural areas	16.6	12.6	9.9	6.4	12.7
Total	10.9	7.8	6.1	5.0	7.7

Table 3 shows how median income varies according to the level of education and type of region in the groups of entrepreneurs and wage earners. In general, median income is

lower in rural areas than in urban areas, the higher the education, the higher is the individual's income. There are, however, a couple of exceptions, especially among entrepreneurs. In rural area, higher educated entrepreneurs earn significantly more than wage-earners. With respect to the level of education, entrepreneurs with secondary education have a lower median income than those with general education. In all, it seems that the income of entrepreneurs varies more irregularly than the income of wage earners. This may also suggest that the returns to education for entrepreneurs and wage earners vary between regions.

Table 3. Median income of entrepreneurs and wage earners according to education and type of region

<i>Region</i>	General education		Secondary education		Lower tertiary education		Higher education		All	
	E	W	E	W	E	W	E	W	E	W
Urban areas	20 400 (1 774)	21 400 (19 519)	18 350 (2 668)	21 300 (43 330)	26 291 (1 457)	26 500 (27 816)	39 900 (603)	38 500 (11 968)	22 200 (6 371)	23 600 (92 619)
Rural areas	18 100 (1 626)	19 500 (8 176)	17 100 (2 409)	20 300 (16 737)	22 800 (846)	24 400 (7 676)	40 900 (145)	33 989 (2 111)	18 800 (5 026)	21 300 (34 700)
Finland in aggregate	19 190 (3 400)	20 800 (27 695)	17 700 (5 077)	21 000 (60 067)	24 900 (2 303)	26 000 (35 492)	40 100 (748)	37 700 (14 079)	20 363 (11 528)	23 100 (137 333)

Notes: E = entrepreneurs, W = wage earners. Number of observations in parentheses.

Other variables

In addition to level of education, many other explanatory variables are used in the estimations. These include experience, field of education, age, gender and mother tongue. Experience is divided into experience in wage work and self-employment experience. In the IV estimations, the parental education is used as identifying instruments. In the Heckman estimations, parental self-employment history is used in the estimation of the selection equation in addition to other variables. Furthermore, the Heckman estimations include variables which describe regional characteristics as well a

variable which indicates whether an individual is residing in his/her region of birth. The descriptions and means of all the variables are given in Appendix 1.

4 Estimation results

Our estimation strategy is to apply the Mincerian approach first to the data on the entire country and then to the two different types of regions. The aim is to estimate causal effects which are not biased due to the endogenous nature of the decision to invest in schooling or to the problem of self-selection. We apply the Heckman and combined Heckman - instrumental variable (IV) approaches, and compare these results to the OLS estimates of the returns to education. In the estimations, the first specifications consider the impact of education on earnings in years, whereas the latter specifications compare the effect of higher education on earnings to that of lower levels of education. Years of schooling are a standard measure where information is derived from the level of education. A linear relationship is assumed to exist between years of education and log-income, and the results indicate the marginal effect of an extra year of schooling. An alternative is to use levels of education, regardless of how many years it takes an individual to attain. In this case, the annual marginal effect of schooling does not have to be constant, but may vary according to education level. The level obtained by an individual may be more informative than the number of years needed to attain it: a vocational qualification or university degree may matter more than years of schooling per se (Card, 1999; Garcia-Mainar and Montuenga-Gómez, 2005). As we were especially interested in the effect of higher education, the latter specifications use a dummy which separates the individuals with higher education from those without.

4.1 Results for the entire country

Table 4 displays the results obtained from the standard OLS method. The results of the Heckman and Heckman - IV estimations are reported in Table 5. In the OLS estimation, education has a highly statistically significant effect on earnings. The results indicate that an extra year of education yields a return of 10.4 per cent for entrepreneurs and 10.8 per cent for wage earners. Rates of return to schooling are lower for self-employed than

for employees, but the difference is small. Van der Sluis et al. (2004) summarized that the average return to a marginal year of education for entrepreneurs was 6.1 per cent across 94 previous studies. Our estimates of returns are higher than in most of the existing literature, but it should be noted that they are based on gross earnings. If we look at the results according to level of education, we can see that they show a higher return to highly educated entrepreneurs than for highly educated wage earners. For entrepreneurs, the return for a highest academic degree is 63 per cent higher than for those who do not have higher education, while for wage earners it is 57 per cent.

Table 4. Earning equations: results for entrepreneurs (E) and wage workers (W) from OLS estimations (the entire country)

Variable	E	W	E	W
<i>Education</i>				
Education years	.104*** (.006)	.108*** (.001)	-	-
Higher education	-	-	.630*** (.040)	.574*** (.005)
<i>Working experience</i>				
Wageexp1	.317*** (.028)	.361*** (.006)	.319*** (.028)	.357*** (.006)
Wageexp2	.585*** (.046)	.653*** (.006)	.590*** (.047)	.648*** (.007)
Entreexp1	.370*** (.029)	.109*** (.011)	.368*** (.029)	.100*** (.011)
Entreexp2	.681*** (.036)	.274*** (.024)	.665*** (.036)	.254*** (.024)
<i>Field of education</i>				
Edutrade	-.041 (.036)	-.027*** (.005)	.219*** (.031)	.204*** (.005)
Edu techn	-.095*** (.027)	.009* (.004)	.098*** (.024)	.173*** (.004)
Eduhesoc	.090* (.045)	.058*** (.006)	.272*** (.041)	.232*** (.006)
Edu servi	-.268*** (.036)	-.050*** (.006)	-.123*** (.035)	.066*** (.006)
<i>Other variables</i>				
Age	.005 (.008)	.049*** (.001)	.008 (.009)	.042*** (.001)
Age2	-.004 (.009)	-.039*** (.001)	-.008 (.009)	-.044*** (.001)
Woman	-.392***	-.354***	-.386***	-.346***

	(.023)	(.004)	(.023)	(.004)
Swedish	.040	.041***	.046	.049***
	(.038)	(.007)	(.038)	(.007)
Otherlan	-.177*	-.041**	-.208**	-.098***
	(.073)	(.013)	(.074)	(.013)
Public sector	-	-.137***	-	-.102***
		(.004)		(.004)
Constant	8.205***	7.587***	9.187***	8.651***
	(.183)	(.023)	(.173)	(.022)

Before commenting on the results for the other variables, it is necessary to increase the quality of the estimates. The first step is to use the Heckman method to correct for possible selection bias related to employment choice. Selectivity is controlled by estimating a Heckman selection model and including the inverse Mill's ratio ('Heckman's lambda') in the second-step regressions. Several variables are used to correct for the non-random nature of choice of employment. They include level of education and its field, gender, age, language, education of parents and regional factors. The results of the probit estimations for self-employment decision, in which education is measured in years are presented in Appendix 2, but these are not commented on here as they are not of direct interest.

The lambda turns out to be significantly negative for wage earners, which indicates the importance of controlling for self-selection bias. For entrepreneurs the lambda is positive, but significant only in the latter specification. A negative sign indicates negative covariance between the error terms in the employment choice and wage functions, whereas a positive sign indicates the opposite. Thus, a negative lambda suggests that unobserved factors tend to decrease the likelihood of selection to entrepreneurship, while they increase earned income. Due to the selection process, wage earners might have better earnings as self-employed. Comparison of the OLS and Heckman results shows them to be rather similar, with one important exception. The adjustment for selectivity has a substantial impact on the coefficient of the return to higher education for entrepreneurs, which is now much closer to the return for wage earners, although remaining higher. Thus, compared to the OLS estimation, this estimation shows a smaller risk premium for highly educated entrepreneurs. Otherwise, the effects on the estimated coefficients are minor. In the specification in which education is defined in years, the rate of return is approximately 10 per cent to a marginal year of education for both groups, but still remaining slightly bigger for wage earners.

The next step is to control for the possible endogeneity of choice of education together with self-selection. An instrumental variable method is combined with the Heckman method in which education of parents is used as identifying instruments for the choice of education. First, the selection model is estimated, from which the inverse Mills ratios are, then, included in the IV regression. In line with previous studies (e.g. van der Sluis et al., 2004; 2006), the use of the IV method produces higher returns to education in both specifications compared to the OLS and plain Heckman estimations. The estimated annual marginal effect of schooling is clearly higher for entrepreneurs than for wage earners: for entrepreneurs it is now 16.9% and for wage earners 13.4%. With respect to the results for wage earners, the estimated effects are higher than in many previous studies conducted in Finland (e.g. Asplund, 1993; Asplund 2000), but similar to the results obtained by Uusitalo (1999). The results of the latter specifications suggest that higher education raises the income of entrepreneurs by as much as 120%. For wage earners, the comparable rate of return is somewhat smaller (99.6 %), but still substantially high.

Thus, these results show strikingly higher returns to education than previous results. Which is more credible?¹ One approach is to test the validity and over-identification of the instruments in the IV estimation. These are tested with the Hansen-Sargan test. The test indicates whether the instruments are uncorrelated with the error term and whether the excluded instruments are correctly excluded from the estimated equation. For entrepreneurs the over-identification test supports the null hypothesis, but it does not do so for wage earners. This suggests that the instruments are not valid for wage earners. The reason for this is not clear, as education of parents is widely suggested and used as an instrument of schooling choice. Another common instrument is education of spouse (see e.g. Iversen et al., 2006), but as all individuals do not have a spouse this is not a

¹ Uusitalo (1999) suggests three distinct explanations for the difference between the OLS and IV estimates. First, the OLS estimates may be downward biased because of measurement error in schooling, while the IV estimates are consistent in the presence of measurement errors. Second, the instruments may be wrongly specified, which could bias the instrumental variable estimates upwards. Third, individuals may optimize their behaviour, which leads to a correlation between schooling and the earnings equation error and furthermore to a bias in the OLS estimates. Griliches (1977) considered that this correlation is negative in practice; educational attainment is cheaper for individuals who have less income to lose in the course of education. Thus, their optimal amount of schooling is larger and the OLS estimates tend to be downward biased.

workable solution.² For these reasons, the results given by the combined Heckman - IV estimation remain somewhat ambiguous, and more weight in the interpretation of the results should be given to the plain Heckman results.

² Instead, information on parental education for all individuals is available.

Table 5. Earning equations: results for entrepreneurs (E) and wage workers (W)
from the Heckman and combined Heckman - IV estimations

Variable	Heckman				Heckman - IV			
	(1)		(2)		(1)		(2)	
	E	W	E	W	E	W	E	W
<i>Education</i>								
Education	.103***	.107***	-	-	.169***	.132***	-	-
years	(.007)	(.000)			(.022)	(.003)		
Higher	-	-	.583***	.570***	-	-	1.200***	.996***
education			(.043)	(.006)			(.163)	(.025)
<i>Working experience</i>								
Wageexp1	.316***	.354***	.310***	.350***	.314***	.360***	.314***	.351***
	(.028)	(.006)	(.028)	(.006)	(.028)	(.006)	(.028)	(.006)
Wageexp2	.586***	.642***	.575***	.636***	.567***	.656***	.563***	.655***
	(.047)	(.006)	(.047)	(.007)	(.048)	(.006)	(.048)	(.007)
Entreexp1	.375***	.1114***	.376***	.106***	.375***	.112***	.378***	.116***
	(.029)	(.011)	(.029)	(.011)	(.029)	(.011)	(.029)	(.011)
Entreexp2	.683***	.287***	.666***	.268***	.686***	.287***	.668***	.284***
	(.031)	(.023)	(.036)	(.025)	(.037)	(.024)	(.037)	(.025)
<i>Field of education</i>								
Edutrade	-.035	-.039***	.201***	.186***	-.223**	-.081***	.172***	.180***
	(.037)	(.006)	(.031)	(.006)	(.071)	(.008)	(.033)	(.005)
Edutechn	-.093***	-.015**	.081**	.143***	-.195***	-.024***	.108***	.171***
	(.027)	(.005)	(.025)	(.005)	(.042)	(.005)	(.025)	(.005)
Eduhesoc	.087	.0858***	.024***	.263***	-.096	.031***	.134**	.269***
	(.045)	(.006)	(.042)	(.007)	(.074)	(.007)	(.052)	(.007)
Eduservi	-.268***	-.026***	-.125***	.092***	-.378***	-.067***	-.128***	.113***
	(.038)	(.006)	(.036)	(.007)	(.051)	(.008)	(.036)	(.006)
<i>Other variables</i>								
Age	.006	.049***	.018*	.054***	-.014	.034***	-.004	.038***
	(.009)	(.001)	(.009)	(.001)	(.011)	(.002)	(.011)	(.002)
Age2	-.005	-.046***	-.018	-.053***	.014	-.033***	.002	-.037***
	(.010)	(.002)	(.010)	(.002)	(.012)	(.002)	(.011)	(.002)
Woman	-.405***	-.387***	-.440***	-.383***	-.326***	-.344***	-.385***	-.345***
	(.031)	(.004)	(.030)	(.004)	(.040)	(.008)	(.023)	(.009)
Swedish	.043	.050***	.064	.058***	.018	.033***	.029	.034***
	(.038)	(.007)	(.038)	(.008)	(.039)	(.007)	(.039)	(.007)
Otherlan	-.179*	-.034*	-.183*	-.089**	-.182*	-.030*	-.214**	-.074***
	(.074)	(.014)	(.073)	(.015)	(.073)	(.013)	(.075)	(.014)
Public sector	-	-.269***	-	-.253***	-	-.162***	-	-.163***
		(.007)		(.007)		(.004)		(.005)
Constant	8.127***	7.474***	8.493***	8.510***	8.339***	7.473***	9.621***	8.607***
	(.253)	(.025)	(.253)	(.026)	(.243)	(.057)	(.355)	(.083)
Lambda	.036	-.596***	.150**	-.680***				
	(.092)	(.024)	(.055)	(.024)				
Sargan statistics					1.118	22.207***	3.551	86.619***
Number of observations					11 359	135 869	11 359	135 869

Notes: E = entrepreneurs, W= wage earners. Standard errors in parentheses. *p-value \leq .05, **p-value \leq .01 and ***p-value \leq .001.

To comment briefly on the results for the other variables, first to be noted is the importance of the other main determinant of the Mincerian earning equation, namely experience. All the experience variables are highly statistically significant, as expected. Prior experience in wage work is a more important determinant of employees' earnings, whereas experience in entrepreneurship matters more for entrepreneurs. Interestingly, experience in entrepreneurship does not have such a big effect on the earnings of wage workers as experience in wage work has on the earnings of entrepreneurs. The results suggest a non-linear relationship between earnings and age for wage earners, while age does not have a significant effect for entrepreneurs. Field of education has a widely significant effect, especially in the latter specifications. Gender has a distinct effect: females earn less than males. The language dummies show, first, that individuals who speak a language other than Finland's two official languages (Finnish or Swedish) as their mother tongue, i.e., immigrants, have lower earnings than natives, and second, that earnings are statistically higher for Swedish-speaking employees but not for Swedish-speaking entrepreneurs. Employees in the public sector have lower earnings than employees in the private sector.

Taken as a whole, most of our control variables proved to be significant and behave as expected. The results obtained using different methods and specifications are mostly consistent. As to the estimated return to education measured in years, this proved to be similar for entrepreneurs and wage earners in the plain OLS estimation and, especially in the Heckman estimation in which choice of employment is controlled for. On the contrary, in the comparison between higher education and other education, the results suggested a somewhat higher return for entrepreneurs than for wage earners. But no clear sign of a risk premium was found for educated entrepreneurs.

4.2 Regional results

Let us now turn to the regional estimations (Table 6). Results by type of region give us an insight into regional differences in returns to education for entrepreneurs and wage earners. The specifications include the same control variables as the equations for the entire country, but for brevity these results are not shown in Table 6.

To begin with, the results seem to diverge to some extent between the two types of regions. In urban areas, the estimated return to education is fairly similar between wage earners and entrepreneurs. An exception is the Heckman - IV specification in which education is measured in years: these results would suggest a higher return for entrepreneurs than for wage earners. Unfortunately, the over-identification test is still not supported for wage earners. For rural areas, the results suggest more systematically higher returns to education for entrepreneurs than for wage earners. The difference in impact measured for years is minor, but clearer difference is found for highly educated entrepreneurs: for this group the estimated return is clearly higher than for wage earners with the same level of education. The Heckman - IV estimation suggest 3.0 percentage units higher returns for entrepreneurs.

If we compare the results for the two types of regions, the estimations suggest that returns to education are higher for entrepreneurs in rural than in urban areas. In particular, well-educated entrepreneurs receive high returns to education in the countryside. In the group of wage earners, returns to education prove to be more alike between urban and rural regions. However, an urban environment might be a more attractive choice for employees in terms of returns to earnings.

These results support Bennett et al. (1995) who found that highly educated persons have relatively higher returns to education in rural areas compared to the lower educated. In our study, this is especially true for entrepreneurs. The results are, however, the reversed of those obtained by Goetz and Rupasingha (2004) who observed that returns to higher degrees of schooling are lower in rural areas compared to urban ones. Particularly for entrepreneurs, the relationship between education and income is more likely to be non-linear in rural areas (cf. Iversen et al., 2006). It is likely that the firms owned by higher educated entrepreneurs are focused on wider markets, while low-skilled entrepreneurs confine themselves to the local market, which tends to raise the relative returns for the highly educated. Online services provide an example of the kind of business run by the highly educated.

Table 6. Regional earning equations: results for entrepreneurs (E) and wage workers (W)

<i>Region type</i>	<i>Variable</i>	OLS				Heckman				Heckman - IV			
		(1)		(2)		(1)		(2)		(1)		(2)	
		E	W	E	W	E	W	E	W	E	W	E	W
Urban areas	Education years	.099*** (.008)	.107*** (.001)	-	-	.109*** (.008)	.107*** (.001)	-	-	.144*** (.022)	.124*** (.003)	-	-
	Higher education	-	-	.554*** (.045)	.559*** (.006)	-	-	.575*** (.047)	.561*** (.006)	-	-	.967*** (.151)	.935*** (.021)
Rural areas	Education years	.098*** (.012)	.102*** (.002)	-	-	.112*** (.013)	.102*** (.002)	-	-	.120*** (.044)	.103*** (.006)	-	-
	Higher education	-	-	.770*** (.089)	.574*** (.013)	-	-	.802*** (.092)	.572*** (.014)	-	-	1.219** (.468)	.921*** (.059)
Sargan statistics										2.447	17.016***	4.524	60.015***
										1.479	5.038***	2.133	18.918***

Notes: E = entrepreneurs, W= wage earners. Standard errors in parentheses. *p-value $\leq .05$, **p-value $\leq .01$ and ***p-value $\leq .001$. The specifications contain the same control variables as the specifications for the entire country (see Table 4), but these results are not presented here.

5 Conclusions

On average in Finland entrepreneurs are less educated than wage earners. This paper analysed the returns to education for entrepreneurs and wage earners in two types of Finnish regions, viz. urban and rural areas. These regions differ in many respects including entrepreneurial activity. In general, rural areas have a higher self-employment rate compared to urban areas. But if we look only at the group of highly educated self-employed, regional differences in entrepreneurial activity become considerably smaller.

Plain data on median income imply that level of education and type of region do not have such a distinct effect on income for entrepreneurs as for wage earners. In the estimation of earnings equations, the returns to education turned out to be much alike for entrepreneurs and wage earners, when educational attainment is measured in years. In the dichotomous comparison between the highly educated and others, the estimated return is slightly higher for entrepreneurs than for employees, although no clear-cut risk premium was found for educated entrepreneurs. Urban areas dominated the results for the entire country, whereas rural areas showed divergent results. In rural areas, returns to education are higher for entrepreneurs. The return for an extra year of education was up to 1.7 percentage units higher for these individuals depending on the estimation method used. Returns are particularly high for highly-educated entrepreneurs in rural areas.

What are the implications of these results for choice of employment and choice of residence? Our results suggest that entrepreneurship is a more attractive employment choice for highly educated individuals in rural areas, whereas in paid-work the situation is rather the opposite. Thus, educated entrepreneurs are financially better off in rural areas, especially if the employment choice is already made. For wage earners, region of residence does not have such importance in terms of returns to education.

Regionally, these findings raise the question of the causes of regional variation in the rate of entrepreneurship and the relative strength of pull and push factors. Are individuals pulled or pushed into self-employment? Is it market pull and higher expected earnings which dominate or are individuals pulled into entrepreneurship

because nothing else is available? Our results suggest that well-educated individuals get a high return to education, especially in rural areas. Even so the option of entrepreneurship is not a highly popular one among the highly-educated in rural areas. Regional differences in the intensity of entrepreneurship do not stem from differences among the well-educated, but rather from differences among the less-educated. Consequently, it is the push effect which dominates, not the pull effect. Regional differences in the rate of self-employment are more likely due to fewer paid-employment opportunities in rural and other weak economically areas than to higher expected earnings.

References

- Alba-Ramirez A, Sansegundo MJ (1995) The returns to education in Spain. *Economics of Education Review* 14:155-166
- Arrow KJ (1973) Higher education as filter. *Journal of Public Economics* 2:193-216
- Ashenfelter O, Harmon C, Oosterbeck H (1999) A review of the schooling/earnings relationship with tests for publication bias. *Labour Economics* 6:453-470
- Asplund R (1993) *Essays on human capital and earnings in Finland*. The Research Institute of the Finnish Economy, Series A 18, Helsinki
- Asplund R (2000) *Private returns to education in Finland: back to basics*. The Research Institute of the Finnish Economy, Discussion Papers 720, Helsinki
- Becker G (1975) *Human Capital*. Columbia University Press, New York
- Bennett R, Glennester H, Nevison D (1995) Regional rates of return to education and training in Britain. *Regional Studies* 29:279-295
- Bernard AB, Redding S, Schott PK, Simpson H (2003) *Relative wage variation and industry location*. National Bureau of Economic Research (NBER), Working Paper 9998
- Bernhardt I (1994) Comparative advantage in self-employment and paid work. *The Canadian Journal of Economics* 27:273-289
- Blanchflower DG (2000) *Self-employment in OECD countries*. National Bureau of Economic Research (NBER), Working Paper 7486
- de Blasio G, Di Addario S (2005) Do workers benefit from industrial agglomeration. *Journal of Regional Science* 45:797-827
- Brown S, Sessions JG (1998) Education, employment status and earnings: a comparative test of the strong screening hypothesis. *Scottish Journal of Political Economy* 45:586-591
- Brown S, Sessions JG (1999) Education, employment status and earnings: a test of the strong screening hypothesis in Italy. *Economics of Education Review* 18:397-404
- Brüderl J, Preisendorfer P, Ziegler R (1992) Survival chances of newly founded business organizations: *American Sociological Review* 57:227-242
- Card D (1999) The causal effect of education on earnings. *Handbook of Labor Economics* 3:1801-1863, North-Holland, Amsterdam
- Carrasco R, Ejrná M (2003) *Self-employment in Denmark and Spain: institution, economic conditions and gender differences*. Centre for Applied Microeconometrics, Institute of Economics, University of Copenhagen
- Dalmazzo A, de Blasio G (2007) Social returns to education in Italian local labor markets. *The Annals of Regional Science* 41:51-69
- Dumond M, Barry H, Macperson D (1999) Wage differentials across labour markets: does cost of living matter? *Economic Inquiry* 37: 577-589
- Duranton G, Monastiriotis V (2002) Mind the gaps: the evolution of regional earnings inequalities in the U.K., 1982-1997. *Journal of Regional Science* 42:219-256
- Evans DS, Leighton LS (1990) Small business formation by unemployed and employed workers. *Small Business Economics* 2: 319-330
- Firkin P (2003) Entrepreneurial capital. In: De Bruin A, Dupuis A. (eds) *New Perspectives in a Global Age*, Ashgate
- Fujita M, Krugman P, Venables AJ (1999) *The spatial economy*, The MIT Press, Cambridge
- García-Mainar I, Montuenga-Gómez VM (2004) Education returns of wage earners and self-employed workers: Portugal vs. Spain. *Economics of Education Review* 24:161-170
- Georgellis Y, Wall HJ (2000) What makes a region entrepreneurial? Evidence from Britain. *The Annals of Regional Science* 34:385-403

- Goetz SJ, Rupasingha A (2004) The returns to education in rural areas. *The Review of Regional Studies* 34:245-259
- Griliches Z (1977) Estimating the returns to schooling: some econometric problems. *Econometrica* 45:1-22
- Hamilton BH (2000) Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy* 108:604–631
- Heckman J (1979) Sample selection bias as a specification error. *Econometrica* 47:153–161
- Iversen J, Malchow-Møller N, Sørensen A (2006) Returns to schooling in self-employment. Centre for Economic and Business Research (CEBR), Discussion paper 19
- Johansson E (2000) Self-employment and the predicted earnings differential – evidence from Finland. *Finnish Economic Papers* 13:45–55
- Kangasharju A, Pekkala S (2002) The role of education in self-employment success in Finland. *Growth and Change* 33:216-237
- Labridianis L (2006) Fostering entrepreneurship as a means to overcome barriers to development of rural peripheral areas in Europe. *European Planning Studies* 14:3-8
- Lehmer F, Möller J (2009) Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort for Germany. *The Annals of Regional Science* (forthcoming)
- Marshall A (1890) *The principles of economics*. Macmillan, London.
- Mincer J (1974) *Schooling, experience and earnings*. Columbia University Press; New York
- Niittykangas H, Tervo H (2005) Spatial variations in intergenerational transmissions of self-employment. *Regional Studies* 39:319–332
- Parker SC, van Praag M (2004) *Schooling, capital constraints and entrepreneurial performance*. Tinbergen Institute Discussion Papers 04-106/3, Tinbergen Institute, Amsterdam
- Parker SC (1999) The inequality of employment and self-employment incomes: a decomposition analysis for the U.K. *Review of Income and Wealth* 45:263-274
- Parker SC (2004) *The economics of self-employment and entrepreneurship*. Cambridge University Press, Cambridge
- Power TM, Barrett RN (2001) *Post-cowboy economics: pay and prosperity in the new American West*. Island Press, Washington DC
- Psacharopoulos G (1979) On the weak versus the strong version of screening hypothesis. *Economic Letters* 4:181-185
- Rees H, Shah A (1986) An empirical analysis of self-employment in the UK. *Journal of Applied Econometrics* 1:95-108
- Robinson PB, Sexton EA (1994) The effect of education and experience on self-employment success. *Journal of Business Venturing* 9:141-156
- Rosen S (1981) The economics of superstars. *American Economic Review* 71:845-858
- Spence M (1973) Job market signalling: *Journal of Labor Economics* 87: 355-374
- Statistics Finland (2001) *Population census 2000 handbook*. Helsinki
- Storey D (1994) *Understanding the small business*. Routledge, London
- Tervo H (2008) Self-employment transitions and alternation in Finnish rural and urban labour markets. *Papers in Regional Science* 87:55- 76
- Tervo H, Haapanen M (2009) The nature of self-employment: how does gender matter? *International Journal of Entrepreneurship and Small Business* (forthcoming)
- Uusitalo R (1999) *Essays in economics of education*. Academic Dissertation, February, University of Helsinki, Faculty of Social Science, Helsinki
- Uusitalo R (2001) Homo entrepreneurus? *Applied Economics* 33:1631-1638

van der Sluis JM, van Praag M, van Witteloostuijn A (2004) Entrepreneurship, selection and performance: a meta-analysis of the role of education. Tinbergen Institute Discussion Paper 03-046/3, Tinbergen Institute, Amsterdam

Appendix 1. Descriptions of variables and their means/share

Variable	Definition	Entrepreneurs	Wage earners
Logincome	ln(income subject to state taxation - unemployment benefit - daily allowance and maternity allowance - home care allowance)	9.8	10.0
Urban	1 if location of workplace is in urban area, 0 otherwise	56.4	74.7
Explanatory variables			
<i>Education</i>			
Education in years	Number of years of education	11.2	11.7
Higher education	1 if higher-degree level tertiary education or doctorate or equivalent level tertiary, 0 otherwise	6.5	10.3
<i>Working experience</i>			
Wageexp1	1 if wage earning experience 5–10 years, 0 otherwise	29.1	24.3
Wageexp2	1 if wage earning experience more than 10 years, 0 otherwise	7.1	54.7
Entreexp1	1 if entrepreneurship experience 5–10 years, 0 otherwise	29.2	2.3
Entreexp2	1 if entrepreneurship experience. more than 10 years, 0 otherwise	40.4	0.4
<i>Field of education</i>			
Edutrade	1 if field of education business or social sciences, 0 otherwise	13.1	16.1
Edu techn	1 if field of education technology or natural sciences, 0 otherwise	28.9	26.7
Eduhesoc	1 if field of education health or welfare, 0 otherwise	6.9	10.9
Edu servi	1 if field of education services, 0 otherwise	10.4	9.3
<i>Personal characteristics</i>			
Age	Age in years	45.2	40.2
Age2	Age-squared divided by 100	21.3	17.5
Woman	1 if female, 0 if male	33.1	50.5
Swedish	1 if native language is Swedish, 0 otherwise	6.7	5.5
Otherlan	1 if native language other than Finnish or Swedish, 0 otherwise	1.8	1.5
Native	1 if residing in sub-region of birth, 0 otherwise	64.1	60.1
<i>Family background</i>			
Fatinedu	1 if father has secondary education, 0 otherwise	11.9	16.5
Fathiedu	1 if father has tertiary education, 0 otherwise	9.9	14.7
Motinedu	1 if mother has secondary education, 0 otherwise	15.0	20.4
Mothiedu	1 if mother has tertiary education, 0 otherwise	6.7	10.9
Entrfat	1 if father was self-employed in 1970, 1980 or 1990, 0 otherwise	16.3	9.2
Entrmot	1 if mother was self-employed in 1970, 1980 or 1990, 0 otherwise	13.6	7.5
<i>Regional characteristics</i>			
Gdp	Annual index of gdp in the region	90.6	95.7
Size of firms	Average size of firms in the region (employees)	4.5	4.8
Employment rate	Unemployment rate in the region	12.6	12.0
Number of observations		11 528	137 333

Appendix 2. Results on the probit estimations of the Heckman selection equation (self-employment decision; education in years)

Variables	Entire country	Urban areas	Rural areas
Education years	-.061*** (.003)	-.049*** (.004)	-.083*** (.006)
Edutrade	.073*** (.018)	.060** (.022)	.171*** (.034)
Edu techn	-.024 (.014)	-.050** (.018)	.042 (.024)
Eduhesoc	.032 (.021)	.077** (.026)	-.040 (.038)
Eduservi	.218*** (.018)	.232*** (.023)	.223*** (.031)
Age	.082*** (.004)	.079*** (.005)	.091*** (.006)
Age2	-.068*** (.004)	-.064*** (.006)	-.0791*** (.007)
Woman	-.395*** (.012)	-.352*** (.014)	-.481*** (.020)
Swedish	.063*** (.022)	.040 (.029)	.040 (.034)
Otherlan	.262*** (.040)	.312*** (.044)	.195* (.099)
Native	.035** (.011)	.024* (.013)	.000 (.019)
Fatinedu	.032* (.016)	.031 (.020)	.063* (.028)
Fathiedu	.062*** (.019)	.083*** (.022)	.083* (.039)
Motinedu	.025 (.015)	.039* (.018)	.019 (.026)
Mothiedu	.065** (.022)	.071** (.026)	.098* (.045)
Entrfat	.339*** (.019)	.302*** (.023)	.393*** (.032)
Entrmot	.223*** (.020)	.228*** (.025)	.199*** (.035)
Gdp	-.001** (.001)	-.001** (.001)	-.000 (.000)
Size of firms	-.070*** (.006)	.055*** (.010)	-.035*** (.008)
Employment rate	-.001 (.002)	.010*** (.002)	-.010*** (.002)
<i>Constant</i>	-2.496*** (.092)	-3.412*** (.123)	-2.215*** (.152)