Education and entrepreneurial success*

Ivar Kolstad**

and

Arne Wiig,***

November 2010

Abstract

Education is commonly believed to be important for the success of entrepreneurial activity. To estimate the returns to education in terms of entrepreneurial profits, however, one must address the challenge that both education and entrepreneurship are endogenous. Using data from Malawi for 1900 firms, this paper estimates returns to education using distance to school as an instrument for education, and land availability as an instrument for entrepreneurship. The results suggest that the effect of education on profits is sizeable for at least some groups of entrepreneurs.

Keywords: Entrepreneurship; returns to schooling; endogeneity; Malawi

JEL Codes: L26, J24, C30

---

* The authors thank Erik Ø. Sørensen, Magnus Hatlebakk, Bertil Tungodden, Eyolf Jul-Larsen and Øivind Anti Nilsen for valuable comments and advice. We are grateful to the National Statistical Office (NSO) of Malawi for providing the data. However, further processing and application of the data was the responsibility of the authors and the views expressed are those of the authors and not of the NSO. The usual disclaimer applies.

** Corresponding author. Chr. Michelsen Institute, P.O.Box 6033 Postterminalen, N-5892 Bergen, Norway. Phone: +47 47 93 81 22. Fax: +47 55 31 03 13. E-mail: ivar.kolstad@cmi.no.

*** Chr. Michelsen Institute, P.O.Box 6033 Postterminalen, N-5892 Bergen, Norway. Phone: +47 47 93 81 23. Fax: +47 55 31 03 13. E-mail: arne.wiig@cmi.no.

1 Introduction

Do more educated people make better entrepreneurs? And if so, by how much does an added year of education increase the profits of an average entrepreneur? Accurately estimating the returns to education is of obvious importance to public policy, in deciding how much public funds to channel into education versus other sectors such as health or infrastructure. A large literature has emerged in recent decades on the impact of education on pay in wage employment, and on entrepreneurial profits. Harmon et al. (2003) find that an added year of education increases wage income by on average 6.5 per cent, based on a meta analysis of micro level studies of wage earners. Similar meta studies of entrepreneurs suggest that an added year of education raises entrepreneurial profits by on average 5.5 per cent in developing countries, and 6.1 per cent in developed economies (van der Sluis et al., 2005; 2008).

Questions remain, however, about how accurate estimates of entrepreneurial returns to education really are. In identifying causal effects of education, one faces the challenge that neither educational nor entrepreneurial status captured by standard surveys reflect anything close to a randomized experiment. Education and entrepreneurial success likely depend on unobserved variables such as ability, the omission of which leads to biased estimates of returns. It is also a well known problem that we only observe profits for those who have chosen to be entrepreneurs, representing a sub-sample of all potential entrepreneurs, which may result in selection bias. The literature on wage returns to education has addressed these challenges through the use of instruments for education and employment (reviews are found in Card (2001), Harmon et al. (2003), and Belzil (2007)). The literature on entrepreneurial returns to education does not, however, exhibit a corresponding emphasis on identifying causal effects. The large majority of studies use ordinary least squares estimation whose selection on observables assumptions are unlikely to hold. The few studies that address either endogeneity of education or selection into entrepreneurship, focus on developed economies or impose exclusion restrictions that seem questionable (van der Sluis et al., 2005; 2007).

This paper attempts to identify the causal effect of education on entrepreneurial profits in a developing country context, using data from the Malawi Second Integrated Household Survey (IHS-2) 2004-2005. Through a three-stage estimation procedure, we address both the problem of self-selection into entrepreneurship, and the endogeneity of schooling. The application of this procedure to the question of entrepreneurial success is to our knowledge novel, allowing us to simultaneously correct for the two types of bias to which estimates of entrepreneurial returns to schooling have been susceptible. This can be viewed as one step towards greater methodological convergence and comparability with the literature on education and wages. Similar to the literature on wages, we find that estimates of entrepreneurial returns to education increase substantially when taking the endogeneity of schooling into account.

The paper is structured as follows. Since a good contextual understanding is needed to find appropriate instruments, section 2 combines a discussion of context and methodology. Information presented on the economy and education system of Malawi is used to motivate our choice of instruments. The methodological approach which integrates a selection model with instrument variable regression, is explained in some detail. Section 3 presents the data used and descriptive statistics. Section 4 presents our main results, followed by a discussion of local average treatment effects and robustness. Section 5 concludes with a look at implications for policy and further research.

2 Background and methodology

Malawi is a least developed country of 15.3 million inhabitants, landlocked between Mozambique, Zambia and Tanzania. Agriculture constitutes 36 per cent of GDP and farming is the most common occupation. Almost 60% of Malawi’s exports stem from tobacco (Republic of Malawi/World Bank 2006). This study focuses, however, on non-agricultural entrepreneurship. A number of people have activities in the informal sector, mainly in petty trade, fisheries and simple service industries, and there are also some larger enterprises mainly in the Southern town of Blantyre. Nevertheless, the private sector remains small in Malawi, and its expansion is an aim of domestic industrial policy (IMF, 2007; Record, 2007). Education is suggested as one possible means to making the private sector more profitable and productive (Republic of Malawi/World Bank, 2006). While the introduction of free primary schools in 1994 likely raised attendance, almost 30 per cent of the official school age children do not start primary school and the average level of schooling in Malawi remains low. In the 8-4-4 education system of the country, only 25% have completed eight years of primary education, 17 % of the relevant age cohort are enrolled at the secondary level and less than 1% are enrolled in tertiary education (Mkandawire and Mulera, 2010).

The Mincer (1974) equation provides the classic setup for estimating the returns to education. In the entrepreneurship literature, most studies use some variant on this, where ordinary least squares (OLS) is used to estimate equation (1). [1] The log of profits of firm *i* is regressed on the education of its owner, using his or her age as a proxy for experience (which is assumed to have a positive but decreasing marginal effect), and controlling for a vector of other firm- and owner-specific variables X_i.

$$\ln(profits_i) = \alpha + \beta_1 age_i + \beta_2 (age_i)^2 + \beta_3 (education_i) + \gamma X_i + \varepsilon_i \qquad (1)$$

The main problem in estimating equation (1) is that there may be selection on unobservables into both education and entrepreneurship. If education is correlated with some unobserved element of the profit equation, OLS estimates are not consistent: unobserved ability may for instance impact positively on both education and profits, leading to an upward bias in OLS estimates of the returns to education. Furthermore, in terms of entrepreneurship, there is only data on profits for people who have chosen to be entrepreneurs, which need not be a representative sample of all potential entrepreneurs. If becoming an entrepreneur is affected by some unobserved variable correlated with unobserved elements of the profit equation, OLS estimates are again not unbiased. In principle, the bias from this selection problem can go either way. In sum, OLS estimates do not capture the causal effect of education, and we cannot surmise a priori which way the results are biased.

Endogeneity of education can be dealt with through instrument variable estimation, by finding a variable correlated with education but not with profits. A number of instruments for education have been suggested in the literature on wages, including family background variables and different types of policy characteristics and reform, and many of these may apply equally well to the question of entrepreneurial returns to education. The problem of selection into entrepreneurship is standardly addressed through the Heckman (1979) selection model. Identification in this case requires a variable correlated with becoming an entrepreneur but not with profits, essentially an instrument. Different types of family background variables

[1] The only other study of entrepreneurial success we are aware of from Malawi, conducted by Chirwa (2008), uses this type of approach. His results show a positive effect on profits of education, measured by dummies for completion of primary, secondary and tertiary school. In the following, we use the more standard measure of education as years of schooling.

have been suggested as instruments for entrepreneurship, such as the occupational status of parents or religious affiliation (van der Sluis et al, 2007).[2]

Perhaps the most convincing studies of returns to education use some feature of school policy as an instrument for education. The seminal study of Angrist and Krueger (1991), for instance, used quarter of birth as an instrument based on legal restrictions on school dropout age. While the introduction of free primary education in Malawi in 1994 provides one potential policy experiment to exploit, the data we use is collected only ten years later, which means that few people affected by this reform will have matured into adult entrepreneurs. Other types of policy experiments also appear to be unavailable. We focus instead on another type of cost likely to affect parental investment in schooling, the time spent travelling to and from school. Parents in households located at a greater distance from a school face greater opportunity costs in sending their kids to school, which is likely to affect their education negatively. We therefore use distance to school (measured in minutes) as our instrument for education, which is similar to the approach taken by Card (1995) in studying higher education in the US. There is considerable variation in how distantly households in Malawi are located from a school, and particularly in less densely populated rural areas travel time is likely to become a binding constraint on investment in education.

Subsistence farming is the most common form of activity among households in Malawi. Our instrument for entrepreneurship builds on the observation that there are limited alternative options besides entrepreneurship for people who cannot make a living as farmers in Malawi. While a number of people also do *ganyu* work, i.e. work as day labourers, more formal employment opportunities are limited. Access to public sector jobs is for the few and well-connected, and there is little private industrial activity on any substantial scale. Migration represents one alternative strategy to farm work, but migration opportunities have become more restricted, in particular to other countries in the region such as South Africa. Individuals from households that have little access to land per household member, are hence more likely to move into entrepreneurial activities. We hence use access to land per household member as our selection variable. Our instrument might be weak if land constrained households could simply acquire more land, but little land changes hands in Malawi due to ambiguities in land titling (Jul-Larsen and Mvula, 2009) and there is also limited new land available particularly

[2] While our data contains information on religious affiliation, religion is not strongly linked to entrepreneurship in our case.

in the more densely populated areas in the South of the country. Since our data suggests that there is likely a u-shaped relationship between access to land and entrepreneurship, meaning that the probability of entrepreneurship is higher for individuals from households with little land and with a lot of land (possibly due to investment of surplus from agricultural activities into business), we also add access to land squared in the selection equation. However, we exclude the very largest land owners from our sample. These are typically owners of large estates, foreigners or politically well connected locals with investment opportunities abroad, and therefore not representative of the general population.[3] While casual interviews we conducted with entrepreneurs in Malawi suggest that parental and elder sibling occupation may also be important predictors of entrepreneurship, there are too few observations for these variables in our data set to use them in estimations.

We would argue that our instruments for education and entrepreneurship are valid in the Malawi context, i.e. they have no direct effect on entrepreneurial profits. Firstly, the possibility that distance to school or access to land are correlated with unobserved geographical profit premiums is addressed through the inclusion of urban/rural and district dummies. Secondly, there is a strong link between land ownership and identity in Malawi, and limited trade in land due to ambiguities in titling. This makes it unlikely that families with a stronger emphasis on education, and consequently more able or highly motivated kids, choose to relocate closer to a school. Parents often send their children to boarding schools instead of relocating the entire household. The problem of mobility is thus more applicable to developed countries such as the US where Card (1995) originally employed the distance instrument, than to Malawi. For similar reasons, it is unlikely that people with greater unobserved entrepreneurial ability choose to live on smaller land plots. This is backed up by the fact that we see little complete specialization in terms of occupation in Malawi, and the average entrepreneur spends considerable time on farming activities (a point to which we return in sections 3 and 4.2).

Addressing endogeneity of education and entrepreneurship separately is technically relatively straightforward. However, addressing both problems at the same time requires a more complicated set-up. Here, we apply the approach outlined by Wooldridge (2002, section 17.4.2). This is a three stage estimation procedure, where the first stage is a probit regression

[3] See for instance Orr (2000) on the ownership of Malawian estates.

of entrepreneurship using access to land and its square as instruments (equation 2 below). The predicted Mills ratio from the probit regression is then used to correct for selection bias in a subsequent instrument variable regression, where we use distance to school as an instrument for education (equations 3 and 4). In addition, all three equations contain individual specific control variables X_{1i} (including age and age squared), and equations 3 and 4 contain firm specific controls X_{2i}.

$$\begin{aligned} \Pr obit(entrepreneur_j) = \alpha_1 + \beta_{11} distance_j + \beta_{12} land_j + \beta_{13}(land_j)^2 \\ + \gamma_{11} X_{1j} + v_{1j} \end{aligned} \tag{2}$$

$$\begin{aligned} Education_i = \alpha_2 + \beta_{21} distance_i + \beta_{22} land_i + \beta_{23}(land_i)^2 + \beta_{24} Mills_i \\ + \gamma_{21} X_{1i} + \gamma_{22} X_{2i} + v_{2i} \end{aligned} \tag{3}$$

$$\begin{aligned} \ln(profits_i) = \alpha_3 + \beta_{31} education(predicted)_i + \beta_{32} Mills_i + \gamma_{31} X_{1i} \\ + \gamma_{32} X_{2i} + \varepsilon_i \end{aligned} \tag{4}$$

Our exclusion restriction is hence that neither distance to school nor access to land feature in the profit equation. All the instruments, however, feature in both equations 2 and 3. The reason for including distance in the probit equation is to avoid bias in the estimates, maintaining $v_{1j} \sim N(0,1)$. Not omitting relevant variables is crucial in non-linear models. Given that distance is included in the first stage, the Mills ratio becomes a one-dimensional reduction of access to land and distance. For identification, equation 3 needs to contain information from one more dimension than equation 4. By including both distance and access to land (in addition to the Mills ratio) we ensure that equation 3 has information from two dimensions, thus ensuring that there is different information in the predicted Mills ratio and the predicted education values. We hence correct for the endogeneity of both entrepreneurship and education in the final profit equation. An added complication in estimating the system of equations is that the Mills ratio is a generated regressor, implying that standard errors are not accurate. Given the survey structure of our data, we follow the standard approach of reporting jackknifed standard errors in order to correct for this.

3 Data

The data used in this paper is taken from the Malawi Second Integrated Household Survey (IHS-2) 2004-2005.[4] The survey covers 11280 households and 52707 individuals. The survey includes a module on entrepreneurship comprising 3913 enterprises. Some individuals own more than one firm, and some firms have more than one owner. In order to merge the enterprise module with other modules we have excluded firms with more than one owner and randomly selected one firm where an individual owns several. This reduces the number of enterprises to 3556. Excluding entrepreneurs under 18 years and large estate owners cuts the sample to 3287 firms. Due to missing data for our main variables this number is further reduced to 1900 enterprises, which constitute our main sample of entrepreneurs. The substantial reduction in observations due to missing data raises the concern that the resulting sample may not be representative; we address this question in a separate section on robustness (section 4.2).

All the variables used for the main estimations are summarized in Table 1. As our dependent variable, we use the log of the monthly profits reported by the owner.[5] Education is measured as years of education, constructed from responses to a survey question of highest class attended. We follow the Mincerian tradition of including age and its square as controls, in addition to a range of other firm- and individual specific controls.[6] Distance, our instrument for education, is the minimum time of travel to school in the household, measured in minutes. Land, our instrument for entrepreneurship, is measured in acres per household member.

---

[4] See www.worldbank.org/html/prdph/lsms/country/malawi04/docs/IHS2%20Basic%20Information.pdf for further documentation.

[5] While one may question the accuracy of reported profits, this appears to be the best available indicator of entrepreneurial success (cf. de Mel et al, 2009). We have checked the consistency of this variable with reported revenues less costs, and the correlation is high (0.81).

[6] We have chosen not to include industry dummies in our estimations, as these are likely to be endogenously determined and influenced by education. A number of other possible control variables suggested by previous studies proved highly insignificant in preliminary estimations and have not been included in our main estimations. These include the number of household members working in business (another measure of firm size), ethnic minority status of the owner and the marital status of the owner.

Table 1. Main variables

Variable	Explanation
Dependent variable	
Profits	Reported profits last month (in Malawi Kwacha), logged
Independent variables	
Education	Years of education
Firm specific controls	
Firm age	Years the firm has been in existence
Registered	Dummy = 1 if enterprise registered with government
Firm size	No. of employees from outside household
Individual specific controls	
Urban	Dummy = 1 if located in urban area
Age	Reported age
Age-squared	
Male	Dummy = 1 if male
Chronic illness	Dummy = 1 if reported to suffer from chronic illness
North	Dummy = 1 if located in the Northern region
Centre	Dummy = 1 if located in the Centre region
Instruments	
Distance	Minimum distance to school of household (minutes)
Land	Available land (in acres) divided by no. of household members
Land-squared	

Table 2 below reports summary statistics for the main variables included in our estimations. The average profits in annual terms are about 270 USD at current exchange rates (1USD =150KZ), and the average firm employs 0.22 non-household members. The median firm has profits of about 80 USD, so the firms in our sample are generally small. There is, however, a great deal of variation. The largest firm in terms of profits earns more than 33,000 USD annually, and the largest firm in terms of employment employs 70 people. The average individual in our sample has between four and five years of education.

Firms in the Northern region correspond to just less than 10 per cent of the sample, firms in the Central region almost 40 per cent, and the remaining 50 per cent are in the Southern region, which roughly reflects relative population sizes in these regions. While not reported in Table 2, 60 per cent of the enterprises in our sample are in the service industries, while manufacturing and primary industries comprise 34 and 6 per cent, respectively. Retail trade is the main entrepreneurial activity, comprising 84 per cent of service sector activities. In manufacturing, handicrafts, beer brewing, distilling, and baking are the main activities, representing about 70 per cent of the sector. Primary industries mainly consist of semi-industrial fishing enterprises, and fishing, forestry and logging activities add up to a 94 per cent share of the sector.

Table 2. Summary statistics

Variable	Obs.	Mean	Std. err. of mean	Min	Max
Profits	1900	3406,16	447,41	10	420000,0
Education	1900	4,69	0,12	0	17,0
Firm age	1900	5,88	0,20	0	54,0
Registered	1900	0,06	0,01	0	1,0
Firm size	1900	0,22	0,08	0	70,0
Urban	1900	0,15	0,01	0	1,0
Age	1900	39,02	0,33	18	98,0
Male	1900	0,55	0,01	0	1,0
Chronic illness	1900	0,16	0,01	0	1,0
North	1900	0,09	0,01	0	1,0
Centre	1900	0,38	0,01	0	1,0
Distance	1900	22,11	0,63	0	120,0
Land	1900	0,35	0,01	0	1,2

Note: Profits in are Malawi Kwacha earned over the past month. Education is years of schooling. Firm age and age are in years, firm size the number of employed not from own household. Registered, urban, male, chronic illness, North and Centre are dummy variables. Distance is minimum household distance to school in minutes, land is in acres per household member.

Diversification is a common livelihood strategy in Malawi, and though 57 per cent of our sample report using more than half their time on entrepreneurship, almost 35 per cent use most of their time on agriculture. Correspondingly, the average entrepreneur spends about 20 hours a week running the enterprise, 11 hours on farming, and 3 hours on other activities. In a separate survey question where individuals are asked to name their main activity, only 26 per cent report being self-employed, while more than 40 per cent report being farmers. There may therefore be some discrepancies between the individuals that are included in the enterprise module of the IHS-2 and those that can be properly characterized as entrepreneurs, a point we return to in section 4.2.

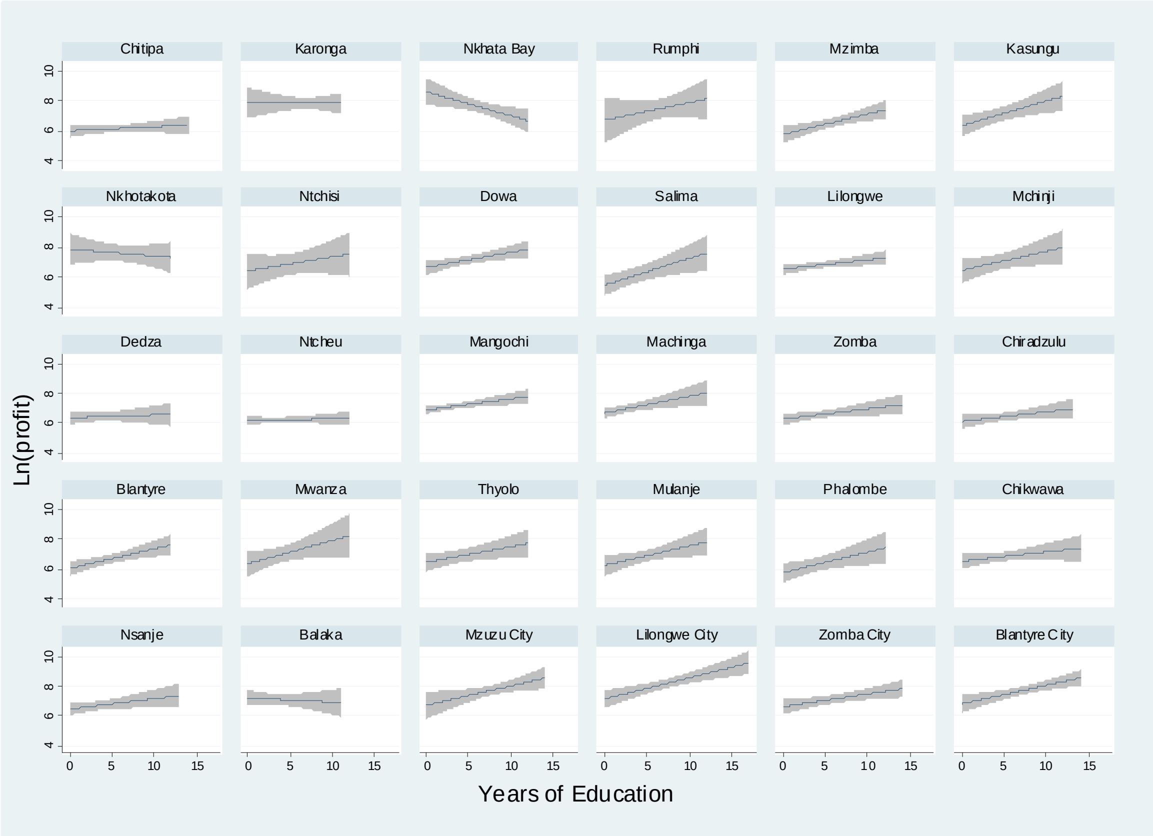
Since we are running a probit model of entrepreneurship, the sample of entrepreneurs is contained within a larger sample including individuals who are not entrepreneurs, comprising a total of 14829 observations. A comparison of entrepreneurs and non-entrepreneurs is instructive, as significant differences exist between them. Consistent with our selection argument, entrepreneurs on average have significantly less land than non-entrepreneurs ($p<0.001$) but the difference is not significant at higher levels of land ownership. In addition, entrepreneurs are on average significantly older ($p<0.001$), they are more likely to be male

($p<0.001$), live in the south ($p<0.001$) and suffer from chronic illness ($p<0.003$). There are no significant differences in education or urban proportions between the two groups.

As an initial assessment of whether the reduction in the sample of entrepreneurs due to missing data leads to a sample that is not representative, we have compared the 1900 entrepreneurs in the main sample with the 1656 entrepreneurs excluded. There are no significant differences in mean profits or education between the two samples. The entrepreneurs in our sample have significantly less land ($p<0.054$) than the excluded entrepreneurs, but this is due to the presence of large estate owners among the excluded entrepreneurs. Entrepreneurs in our sample are older and more likely to be women. While there are no significant differences in the proportion of firms from each region, there appears to be significant overrepresentation of manufacturing firms, and underrepresentation of primary sector and service firms in our sample when compared to the excluded firms.

The recent literature on education and wages has focused on heterogeneity in the returns to education. This is something that should also be considered in estimating entrepreneurial returns to education. In Malawi, for instance, education levels vary considerably across regions, with the highest average education level in the North and the lowest in the Central region. If marginal returns to education are decreasing in education levels, as suggested by Card (1999), local average treatment effects may be greater in the Central region and lower in the North. Simple bivariate correlations suggest that there may be heterogeneity in entrepreneurial returns to education. As an illustration of this, Figure 1 presents fitted values for the relation between profits and education broken down by district, of which there are 30 in Malawi. As the figure shows, the slope of the fitted line varies a lot across districts, and while most districts exhibit an upward sloping line, there are three districts where the correlation between profits and education is negative. In the presence of heterogeneity in returns, the question of which local average treatment effects our instrument picks up becomes important. This is analyzed at some length in the subsequent section, albeit from a different angle, following the presentation of our main results.

Figure 1. Profits and education by district (fitted values).



4 Results

4.1 Main results

The results from the respective stages of the three stage estimation are presented in columns one to three in Table 3, while the fourth column presents OLS estimates for purposes of comparison. The results from the first stage of the three stage estimation (first column) show that access to land and its square work well as instruments for entrepreneurship. Both are highly significant, and indicate the expected u-formed relationship between access to land and the probability of being an entrepreneur. The turning point at which the marginal effect of access to land goes from being negative to being positive is at about 0.71 acres/person, which is reasonable in light of the descriptive statistics in Table 2. The signs of the other controls also appear reasonable, suggesting that the probability of entrepreneurship increases with age (until about age 42), with being male, and with living in the more industrialized Southern region. Chronic illness is positively related to entrepreneurship, possibly due to illness resulting in lower productivity in farming.

Table 3. Main regression results.

	First stage	Second stage	Third stage	OLS
Dependent variable	*Entrepreneurship*	*Education*	*Profit (logged)*	*Profit (logged)*
Education			0.214*	0.060***
			(0.12)	(0.01)
Firm age		-0.038***	0.030***	0.024***
		(0.01)	(0.01)	(0.00)
Registered		1.150***	0.693**	0.886***
		(0.40)	(0.30)	(0.16)
Firm size		0.109	0.046	0.063*
		(0.21)	(0.24)	(0.04)
Urban	-0.001	2.176***	0.390	0.831***
	(0.06)	(0.38)	(0.27)	(0.09)
Age	0.083***	0.790	0.027	0.025**
	(0.01)	(1.00)	(0.09)	(0.01)
Age-squared	-0.001***	-0.009	-0.000	-0.000***
	(0.00)	(0.01)	(0.00)	(0.00)
Male	0.221***	3.979	0.372***	0.661***
	(0.03)	(2.59)	(0.11)	(0.06)
Chronic illness	0.158***	1.664	-0.195	-0.156*
	(0.05)	(1.82)	(0.17)	(0.08)
North	-0.230***	-0.553	-0.190	0.072
	(0.06)	(2.71)	(0.46)	(0.10)
Centre	-0.179***	-1.751	-0.012	-0.017
	(0.04)	(2.08)	(0.22)	(0.08)
Distance	-0.003***	-0.057*		
	(0.00)	(0.03)		
Land	-0.520***	-8.698		
	(0.18)	(6.17)		
Land-squared	0.365**	6.958		
	(0.16)	(4.41)		
Mills ratio		13.293	-0.228	
		(14.48)	(1.15)	
Constant	-2.665***	-30.195	5.267	5.640***
	(0.13)	(42.03)	(4.32)	(0.24)
r2		0.245	0.123	0.268
N	14829	1900	1900	1900

*Jackknifed standard errors in parentheses, *** indicates significance at the 1% level, ** at 5%, * at 10%.*

It turns out that the predicted Mills ratio is not significant in the education equation (column two) or profit equation (column three). There hence does not appear to be a selection bias in estimating the entrepreneurial returns to education. In other words, just running an IV estimation instrumenting for education would give very similar results.[7]

Moving to the results of the second stage of the three stage estimation (column two), distance has the expected negative relation with education, indicating that entrepreneurs living further from a school are less educated. This variable is only significant at the 10% level, and an F-test of whether distance enters the education equation yields an F statistic of 2.84, well below the conventionally required level of 10 (cf. Staiger and Stock, 1997). However, the low test statistic reflects increased inefficiency due to the inclusion of the insignificant land variables and Mills ratio. When dropping the first stage, and simply running an IV estimation, the

[7] Results from the corresponding IV regression can be found in Table A1 in the Appendix. We report the full three stage estimation results here, since we get different results on selection bias when addressing the question of missing data in Section 4.2.

distance variable has an F statistic of 23.15, which is well above conventional levels. The instrument for education is therefore stronger than its significance level in Table 3 would suggest. For the other control variables, the proprietors of older firms appear to have less education (possibly reflecting increasing education levels among younger cohorts), and owners of registered firms in urban areas are more educated.

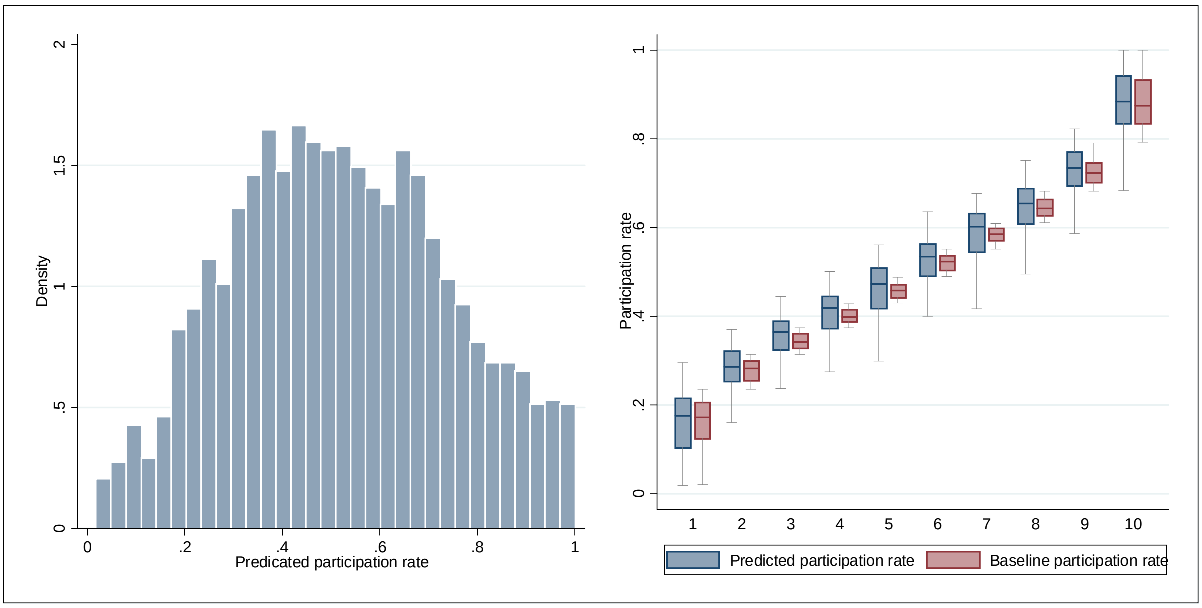
The three stage estimation suggests that the returns to education in terms of entrepreneurial profits is at about 21 per cent (column three). This point estimate is well above the OLS estimate of about 6 per cent return (column four). One possible explanation for the magnitude of the estimated effect is that there may be heterogeneity in returns to education across different groups. Rather than average treatment effects, our results may therefore reflect local average treatment effects for the groups whose education are affected by our instrument for education. Formal analysis of which groups are moved by the distance instrument is complex in this case. Since our education variable takes on multiple values, there is the possibility of variable treatment intensity. 2SLS results are then a weighted average of unit causal effects of schooling (Angrist and Imbens, 1995). The weights can be calculated to tell us how the groups moved by our instrument are distributed over the range of education values. This is, however, complicated by the fact that our instrument for education is (in principle) continuous.

We rely instead on the graphical representation techniques used in Moffitt (2008). In an analysis of returns to higher education in the UK, Moffitt uses probit regressions to generate predicted participation rates in higher education. By comparing participation rates with and without the inclusion of instruments in the probit equation, it is possible to discern where the instruments create action in terms of educational outcomes. Since our education variable is not dichotomous, we have to modify Moffitt's approach. We generate dichotomous variables for having at least one year of education, at least two years, and so on. These are then used as dependent variables in a series of probit regressions, including and excluding distance as an explanatory variable. The resulting distributions of participation probabilities provide a picture of where in the range of education values the instrument has an effect on education, and for what participation probabilities.

It turns out that distance does little if anything to alter the distribution of probabilities of education in secondary and tertiary education. Where the distance variable does have an effect is in primary education. Which part of the distribution of participation probabilities distance affects is fairly similar across all the years of primary education, but the effects appear more marked around year 4 and 5, i.e. around mean education levels. We have therefore included representations of participation probabilities for 5 or more years of education in Figure 2. The right hand panel of the figure shows the range of participation probabilities for each decile of these probabilities. The red boxes convey the ranges when distance is held at its mean (baseline participation rates), while the blue boxes represent the ranges when distance is allowed to vary (predicted or actual participation rates). For instance, for the 10 per cent of individuals with lowest participation rates (decile 1), probabilities of taking 5 or more years of education range from just above zero to a little above 0.2 when distance is held at its mean (cf. tick marks red box), and from just above zero to approximately 0.3 when distance is allowed to vary (cf. tick marks blue box). As the figure suggests, the distance instrument has more of an effect at participation rates between 0.3 and 0.7, but very little at high or low participation rates. In other words, distance affects primary education most for those with medium probabilities of acquiring such education. As observed by Moffitt (2008), this also means that our instrument is strong for those with medium participation rates, but weak for high and low rates.

The left hand panel of the figure includes a histogram of predicted participation rates. For year 5, the majority of entrepreneurs have participation rates around the levels where the instrument does the most work. In addition to the impact of distance on education participation being greatest around year 4 and 5, these are also the years for which the greatest number of individuals are likely to be moved by the instrument, By contrast, the distributions of participation rates at lower or higher years of primary education are more skewed to the right and left, respectively, thus putting less weight on the medium participation rates where the instrument has most of an effect. This means that those with medium predicted participation rates at or around grades 4 and 5 are overrepresented among those whose education are affected by our particular instrument.

Figure 2. Predicted and baseline participation in education at and above grade 5



In the right panel, the upper and lower points of the rectangles are the 75th and 25th percentile points of the distribution, respectively. The horizontal lines inside the rectangles are the medians. Upper and lower tick marks are upper and lower ranges.

In simpler terms, Figure 2 indicates where the distance instrument has most of an effect on education (right hand panel) for the most individuals (left hand panel). Distance has a relatively strong effect for those with medium probabilities of getting at least 4 or 5 years of education, and this group also comprises a large part of the population. In other words, this is where the instrument creates the most action in terms of education outcomes, and the returns our estimations pick up reflect the returns of this group. This provides the basis for a possible interpretation of our estimate. Moffitt (2008) shows that the effect of education decreases with participation rates, i.e. returns to education go down as larger parts of a population, and hence more lower-return individuals, are drawn into education. In our case, the distance instrument disproportionately affects individuals with medium participation rates. Our estimate of returns to education may thus reflect a high return among this group, potentially above what would be the case for groups with higher participation rates. Further characterization of compliers is, however, difficult in our case. In the first stage of the IV regression it seems that the effect of distance on education may be relatively stronger for women, and in the Central region of Malawi, but the differences are not significant. Since education levels are on average lower for women than men, and lower in the Central region than elsewhere in Malawi, our estimate could reflect a high return among those with medium participation rates in these two sub-populations.

4.2 Robustness of results

As noted earlier, there is a substantial drop in the number of observations due to missing data for some of the variables. This is particularly the case for school distance, our instrument for education. This raises the question of whether our results are biased as the entrepreneurs in our main sample may not be representative. In a way, there are two selection problems here, selection into entrepreneurship, and selection of entrepreneurs into our main sample. To assess this issue, we reran the main estimation substituting mean distance at the primary sampling unit (psu) level for missing values of the school distance variable. While somewhat inaccurate, individuals are likely to live at a distance from school similar to that of their neighbours.

The results of the three-stage estimation replacing psu distance means for missing values are presented in Table 4. Including distance means increases the number of observations from 1900 to 2962. The sample increase does not appear to affect the strength of our instruments; access to land and its square are still significantly related to being an entrepreneur (column one), and distance to school is significantly related to education (column two). However, the results from the profit equation change markedly (column three). Firstly, the mills ratio is now significant, indicating that there is a selection bias in the estimation of entrepreneurial returns to education. Secondly, the estimated returns to education are now marginally negative and insignificant, suggesting that the returns are essentially zero. Finally, the estimated effect of a number of the covariates changes when the sample is increased; entrepreneurs with registered firms in urban areas appear to do better as compared to previous estimations, and chronic illness has a significant and substantial negative relation to profits.

Table 4. Regression results substituting psu means for missing values of distance

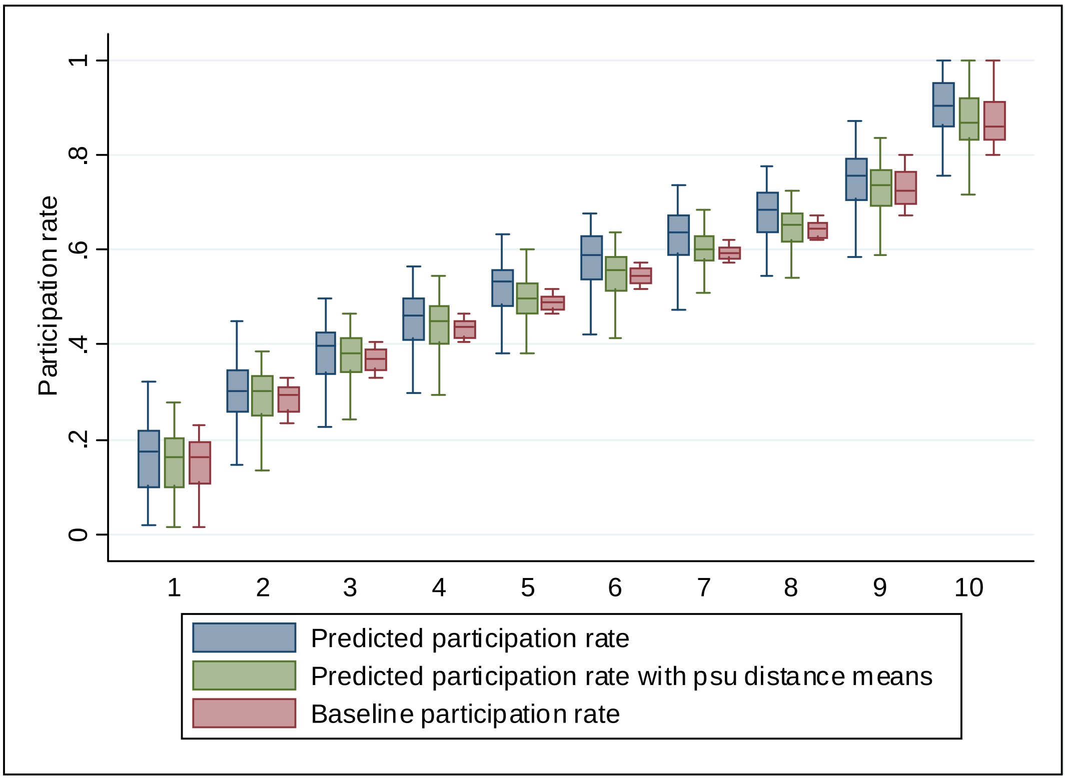
	First stage	Second stage	Third stage
Dependent variable	*Entrepreneurship*	*Education*	*Profit(logged)*
Education			-0.041
			(0.12)
Firm age		-0.034***	0.019***
		(0.01)	(0.01)
Registered		1.349***	1.073***
		(0.38)	(0.28)
Firm size		0.138	0.092
		(0.27)	(0.32)
Urban	0.010	2.467***	0.982***
	(0.06)	(0.31)	(0.28)
Age	0.065***	1.213**	-0.072
	(0.00)	(0.59)	(0.05)
Age-squared	-0.001***	-0.014**	0.001
	(0.00)	(0.01)	(0.00)
Male	0.330***	7.878***	0.315**
	(0.03)	(2.92)	(0.13)
Chronic illness	0.165***	3.073**	-0.404***
	(0.04)	(1.44)	(0.15)
North	-0.240***	-2.850	0.689*
	(0.06)	(2.13)	(0.39)
Centre	-0.164***	-3.169**	0.325***
	(0.04)	(1.45)	(0.13)
Distance	-0.004***	-0.099***	
	(0.00)	(0.03)	
Land	-0.556***	-12.964***	
	(0.15)	(4.95)	
Land-squared	0.442***	10.448***	
	(0.13)	(3.96)	
Mills ratio		24.599**	-1.941***
		(10.90)	(0.75)
Constant	-2.259***	-55.260**	11.139***
	(0.09)	(27.98)	(2.77)
r2		0.247	0.232
N	22211	2962	2962

*Jackknifed standard errors in parentheses, *** indicates significance at the 1% level, ** at 5%, * at 10%.*

There are at least two possible explanations for the changes in the returns to education estimate. One explanation is that our main sample of 1900 entrepreneurs is not representative, and a bias arises from not including the entrepreneurs where there is missing information. In other words, some unobserved characteristic that affects whether an entrepreneur is able to report his/her distance to school, is correlated with an unobserved variable in the profit equation. If returns to education are homogeneous, this casts doubt on the robustness of our main result. However, another explanation is that the two estimates of returns to education pick up fundamentally different things. As illustrated above, if returns are heterogeneous, our main estimate picks up the local average treatment effect for a specific subset of entrepreneurs. When we introduce mean distance values for those with missing data, whose returns we pick up becomes much less specific.

We illustrate this in Figure 3, which is essentially an elaboration on the right hand side panel of Figure 2. As before, the red boxes represent the range of predicted participation in education (at or above grade 5) when distance is held at its mean for everyone. The blue boxes show participation rates using actual values for distance. To this we have added the green boxes which are the predicted participation rates when missing values for distance are replaced by psu means. As the figure reveals, the green boxes are more compressed and centred around the red boxes than are the blue ones. In other words, introducing psu means for missing values of distance leads to less dispersion in predicted education values, than using only actual values for distance. Changes in education induced by the instrument are thus less attributable to specific parts of the population when using a distance instrument which includes psu means for some observations. The estimated returns to education in Table 4 may then be more of an equally weighted average of treatment effects across different groups, compared to the returns in Table 3 which represent local average treatment effects for more specific groups. In short, with heterogeneous returns, we pick up returns for different people when replacing missing values for distance with psu means.

Figure 3. Predicted participation in education (at or above grade 5) with and without using psu means for missing values of distance



Returning to our main sample of 1900 enterprises, the results in section 4.1 are robust to a range of specification changes. Adding district dummies instead of regional dummies raises the estimate only marginally. Including industry dummies reduces the estimate by a couple of percentage points, but this is mostly due to a reduction in sample due to missing data on industry affiliation. The inclusion of other explanatory variables that have been suggested in the entrepreneurship literature (ethnic minority and marital status) does not change results. Extending the sample to include 16- and 17-year-olds adds only marginally to the number of observations, and does not affect results.

As noted in section 3, Malawian households are often quite diversified in terms of the productive activities they are involved in. Though the majority of entrepreneurs in our sample spend most of their time on their business, they also typically do other work such as farming or wage work. The above results are for the full sample of entrepreneurs, not restricted to those that have entrepreneurship as their main activity. This raises the question of whether the estimated entrepreneurial returns to education are misrepresented by including a number of individuals that should not be characterized as entrepreneurs. As it turns out, however, this is not much of a problem. In Table 5, we present abridged results from regressions where the sample is restricted to individuals who spend more than 50% of their time on entrepreneurial activities (first column), individuals who spend more than 50% and at least 7 hours per week on such activities (second column), and individuals who name self-employment as their main activity (third column).[8] The returns to education drop only marginally in the first two cases, as seen in columns one and two. In the third column, the education coefficient becomes low and insignificant. This, however, is because the distance instrument becomes weak when the sample is restricted to only those 493 individuals reporting self-employment as their main activity.

Table 5. IV-regressions, entrepreneurship main activity. Dependent variable: ln(profits).

	More than 50% of time in business	More than 50% and 7 hours/week in business	Main activity self-employment
Education	0.200**	0.204**	0.024
	(0.09)	(0.09)	(0.14)
r2	0.121	0.105	0.285
N	1027	1001	493

*Controls included but not reported. Standard errors in parentheses. *** indicates significance at the 1% level, ** at 5%, * at 10%.*

The estimated entrepreneurial returns to education are above typical estimates of wage returns in developing countries. In order to see whether our estimates have some reasonable relation to wage returns in Malawi, we therefore also ran similar regressions using wages as the dependent variable. Wages in this case include cash pay and allowances/gratuities for wage/salaried work (*ganyu* excepted). We do not have an instrument for selection into employment, and so cannot rule out there being a selection bias. Table 6 reports IV regression results where we instrument for education using distance to school, and employ a specification similar to that used earlier. Results for the first stage education equation in the first column shows distance to be highly significant also for the sample of wage earners, and has an F-value of 34.8 which is well above conventionally required levels, suggesting a strong instrument. The estimated wage returns to education in column two are almost exactly the same as the entrepreneurial returns estimated earlier. There thus seems to be a correspondence between estimated entrepreneurial and wage returns to education, which is perhaps not surprising as the instrument likely picks up the same groups in both cases.

[8] The results presented are from IV regressions, due to the fact that our three stage estimator does not converge in the first two cases.

Table 6. IV-regression. Dependent variable: ln(wages)

	First stage	Second stage
Dependent variable	*Education*	*Wages (logged)*
Education		0.216***
		(0.04)
Urban	3.463***	0.037
	(0.38)	(0.16)
Age	0.194***	0.065***
	(0.05)	(0.01)
Age-squared	-0.003***	-0.001***
	(0.00)	(0.00)
Male	0.830***	0.076
	(0.23)	(0.06)
Chronic illness	-0.589**	0.018
	(0.29)	(0.08)
North	1.414***	-0.308***
	(0.33)	(0.10)
Centre	0.142	0.132*
	(0.33)	(0.07)
Distance	-0.045***	
	(0.01)	
Constant	3.798***	5.014***
	(1.08)	(0.24)
r2	0.218	0.292
N	2260	2260

*Standard errors in parentheses. *** indicates significance at the 1% level, ** at 5%, * at 10%.*

5 Concluding remarks

In this paper, we have used a three stage approach which addresses the potential problems related to self-selection into both entrepreneurship and education. This brings the literature on entrepreneurial returns to education methodologically closer to the corresponding literature on wages and education. Our estimate of entrepreneurial returns to education suggests that it is sizeable at least for some entrepreneurs. The final part of this statement should be stressed for policy purposes; if there is heterogeneity in returns to education, our results do not necessarily reflect an average effect of education on profits across all entrepreneurs. Rather, they indicate a local average treatment effect of education, i.e. the impact of education on those groups whose education is affected by school distance. One should therefore be careful in using these findings to assess the impact of policies to increase education generally among entrepreneurs in Malawi. They are nevertheless useful in assessing policies that more specifically affect the groups whose return we have identified. For instance, the results provide some indication of the impact of interventions that aim to reduce distance to primary schools in relevant areas, such as building new schools or improving transport facilities in more remote parts of the country. These interventions may affect the schooling of groups similar to those whose education is affected by the distance to school instrument, i.e. those with medium participation rates in the middle years of primary school.

These observations nevertheless point to the need for further analysis to map other parts of the return function. Such analysis would, however, require additional instruments for education, which are not readily available in most typical household data sets. Policy experiments provide an important source of variation in this sense, but this has proved hard to exploit in the case of Malawi. While this paper has focused on the returns to formal education, other types of education more specifically aimed at entrepreneurs may be more amenable to randomized experiments. For instance, Bjorvatn and Tungodden (2010) use this approach to estimate the impact of business training on microcredit clients in Tanzania. While some progress is being made, more research on causal effects of human capital on entrepreneurial success is nevertheless needed to inform policy debates in this area.

References

Angrist, J. D. and Imbens, G. W. (1995), “Two-stage least squares estimation of average causal effects in models with variable treatment intensity”, *Journal of the American Statistical Association*, 90, 430-42

Angrist, J. D. and Krueger, A. B. (1991), “Does compulsory schooling attendance affect schooling and earnings?” *Quarterly Journal of Economics*, 106, 976-1014

Belzil, C. (2007). “The return to schooling in structural dynamic models: a survey”. *European Economic Review*, 51, 1059–1105

Bjorvatn, K. and Tungodden, B. (2010), ”Teaching entrepreneurship in Tanzania: Evaluating participation and performance”, *Journal of the European Economic Association*, 8, 2-3, 561-570

Card, D. (1995), "Using geographic variation in college proximity to estimate the return to schooling," in Christofides, L. N., Grant, E. K., and Swidinsky, R. (eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, Toronto: University of Toronto Press, 201-222

Card, D. (1999), “The causal effect of education on earnings”, in Ashenfelter, O. and Card, D. (eds.), *Handbook of Labour Economics*, vol. 3A, Amsterdam: North-Holland

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

Chirwa, E. W. (2008), “Effects of gender on the performance of micro and small enterprises in Malawi”, *Development Southern Africa*, 25, 3 347-362

de Mel, S., McKenzie, D. J. and Woodruff, C. (2009), “Measuring microenterprise profits: Must we ask how the sausage is made?” *Journal of Development Economics*, 88, 1, 19-31

Harmon, C., Oosterbeek, H., and Walker, I. (2003), “The returns to education: Microeconomics”, *Journal of Economic Surveys*, 17, 2, 115-155

Heckman, J. J. (1979), “Sample selection bias as a specification error”, *Econometrica*, 47, 1, 153-61

IMF (2007), *Malawi: Poverty Reduction Strategy Paper—Growth and Development Strategy*, IMF Country Report No. 07/55, Washington D.C.: International Monetary Fund

Jul-Larsen, E. and Mvula, P. (2009), ” Security for many or surplus for the few? Customary tenure and social differentiation in Southern Malawi”, *Journal of Southern African Studies*, 35, 1, 175-190

Mkandawire, F. R., and Mulera, D., (2010), “Malawi” in Peterson, P., Baker, E. and McGaw, B. (eds.), *International Encyclopedia of Education*, Amsterdam:Elsevier, 666-672

Moffitt, R. (2008), *Estimating marginal treatment effects in heterogeneous populations*, mimeo, revised January 2008, Johns Hopkins University

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press

Orr, A. (2000), “’Green gold’? Burley tobacco, smallholder agriculture, and poverty alleviation in Malawi”, *World Development*, 28, 2, 347-363

Record, R. (2007), “From policy to practice: Changing government attitudes towards the private sector in Malawi”, *Journal of International Development*, 19, 805-816

Republic of Malawi/World Bank (2006). *Malawi Poverty and Vulnerability Assessment: Investing in Our Future*. http://www.aec.msu.edu/fs2/mgt/caadp/malawi_pva_draft_052606_final_draft.pdf, accessed 29 November 2010

Staiger, D. and Stock, J. H. (1997), “Instrumental variable regressions with weak instruments”, *Econometrica*, 65, 3, 557-586

van der Sluis, J., van Praag, M. and Vijverberg, W. (2005), ”Entrepreneurship selection and performance: A meta-analysis of the impact of education in developing economies”, *World Bank Economic Review*, 19, 2, 225-261

van der Sluis, J., van Praag, M. and van Witteloostuijn, A. (2007). *Why are the returns to education higher for entrepreneurs than for employees*. IZA Discussion Paper 3058, Bonn: Institute for Labor

van der Sluis, J., and M. van Praag, (2008). "Education and entrepreneurship selection and performance: A review of the empirical literature”, *Journal of Economic Surveys* 22(5): 795-841

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

Appendix

Table A 1. IV regression

	First stage	Second stage
Dependent variable	*Education*	*Profit(logged)*
Education		0.207***
		(0.07)
Firm age	-0.039***	0.030***
	(0.01)	(0.01)
Registered	1.168***	0.705***
	(0.35)	(0.18)
Firm size	0.113***	0.047
	(0.03)	(0.03)
Urban	2.517***	0.428**
	(0.34)	(0.20)
Age	-0.125***	0.042***
	(0.03)	(0.01)
Age-squared	0.001	-0.000***
	(0.00)	(0.00)
Male	1.626***	0.424***
	(0.18)	(0.13)
Chronic illness	0.026	-0.165*
	(0.24)	(0.09)
North	1.837***	-0.218
	(0.26)	(0.17)
Centre	0.110	-0.043
	(0.23)	(0.08)
Distance	-0.027***	
	(0.01)	
Constant	7.895***	4.618***
	(0.84)	(0.53)
r2	0.239	0.136
N	1900	1900

*Standard errors in parentheses, *** indicates significance at the 1% level, ** at 5%, * at 10%.*