

Education and entrepreneurial success

Ivar Kolstad · Arne Wiig

Accepted: 26 November 2014 / Published online: 6 December 2014
© Springer Science+Business Media New York 2014

Abstract This paper estimates the effect of education on the success of entrepreneurial activity, using survey data from Malawi. An instrument variable approach is used to address the endogeneity of education. We find a significant and substantial effect of an added year of primary education on entrepreneurial profitability. This is consistent with theoretical arguments that primary schooling provides a generalised form of competence that underpins the variety of skills an entrepreneur needs to succeed in business. Results are robust to non-random selection into entrepreneurship.

Keywords Entrepreneurship · Returns to schooling · Endogeneity · Malawi

JEL Classifications L26 · J24 · C30

1 Introduction

Do more educated people make better entrepreneurs? A number of recent studies have used randomized

controlled trials to identify the impact of business training programmes for entrepreneurs, some of which find positive effects on profits and/or sales (see McKenzie and Woodruff 2014 for a review). Similar methods are more difficult to apply to the question of the effect of formal education on entrepreneurial success. Nevertheless, it is likely that formal education, and in particular basic skills of literacy and numeracy, is a prerequisite for absorbing and putting to use the information learnt in a business training programme, or more generally to make use of business strategies above some minimal level of complexity. In his jack-of-all-trades theory of entrepreneurship, Lazear (2004) stresses that an entrepreneur must have a variety of skills to success in business, which similarly requires basic literacy and numeracy. This implies that the effect of primary education (which enhances people's generalised human capital) on entrepreneurial success is potentially large. This article uses survey data from Malawi, a country where a large percentage of the population has not completed primary school, to estimate the rate of return of primary education in terms of entrepreneurial profits.

The article makes two main contributions. Firstly, it addresses the methodological challenge that both education and entrepreneurship are endogenous variables. Education and entrepreneurial success likely depend on unobserved variables such as ability, the omission of which may lead to biased estimates of returns. It is also a well-known problem that we only observe profits for those who have chosen to be

I. Kolstad (✉) · A. Wiig
Chr. Michelsen Institute, Postterminalen, P.O. Box 6033,
5892 Bergen, Norway
e-mail: ivar.kolstad@cmi.no

A. Wiig
e-mail: arne.wiig@cmi.no

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; 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.¹ The few studies that address either endogeneity of education or selection into entrepreneurship, impose exclusion restrictions that seem questionable (van der Sluis et al. 2005, 2007). We address the endogeneity of education through an instrumental variable approach, using distance to school as an instrument. Moreover, through a three-stage estimation procedure where land availability is also used as an instrument for being an entrepreneur, we show that our results are robust to controlling for both the endogeneity of education, and self-selection into entrepreneurship.

Secondly, our methodological approach permits us to identify the effect of *primary* education on entrepreneurial success. By modifying the approach of Moffitt (2008) to the case of a non-dichotomous endogenous variable, we show that our distance to school instrument has the largest effect on the schooling choices of individuals with medium probability of completing grades 4 and 5 of primary school. This means that if there are heterogeneous returns to education, our instrumental variable estimation results are picking up the local average treatment effect of education for these types of individuals, we are in other words measuring the returns to treating them with an added year of primary education. Consistent with our expectation that primary schooling provides a generalised form of competence that underpins the variety of skills an entrepreneur needs to succeed in business, we find a significant and substantial effect of an added year of primary education on entrepreneurial profitability. We see this as a contribution to a

literature on entrepreneurial returns to education which has so far focused on average treatment effects across all years of education, or estimated the effect of primary, secondary and tertiary completion without addressing the challenge that completion of each level is an endogenous choice.

The paper is structured as follows. Section 2 presents the theoretical backdrop of our analyses and the main hypothesis, and it explains our empirical approach. Our main instrumental variable estimation, and the three-stage approach used as a robustness check to also take into account the endogeneity of entrepreneurship, are presented in some detail, and the choice of instruments motivated through contextual information on the economy and education system of Malawi. Section 3 presents the data used and descriptive statistics. Section 4 presents our main instrumental variable estimation results, followed by the analysis of local average treatment effects, and a robustness section showing the results to be robust to endogenous selection into entrepreneurship, and other challenges. Section 5 concludes with a look at implications for policy and further research.

2 Conceptual framework and empirical specification

The effect of human capital on labour market outcomes has been analysed in a number of studies; for many of these the Mincer (1974) model provides the theoretical basis. In the Mincer model, education has a positive impact on worker productivity. In the original Mincer model, individuals are assumed identical *ex ante*, and the model expresses the idea that since schooling requires earnings to be forgone individuals require a compensating differential in wages in occupations that require a longer schooling period (Mincer 1958). In a second model, Mincer (1974) builds on an accounting identity model and focuses on the relationship between observed earnings, potential earnings, and human capital investment (both in terms of on-the-job investment and formal education). The rate of return to formal schooling is assumed constant for all years of schooling, and the rate of return to post-school investment is also constant. By assuming a linearly declining rate of post-school investment and equating the present value of earnings across years of schooling, a log linear

¹ In two meta-studies van der Sluis et al. (2005, 2008) suggest that an added year of education raises entrepreneurial profits by on average 5.5 % in developing countries, and 6.1 % in developed economies. See also the review by Dickson et al. (2008).

relationship between earnings and schooling is derived. The coefficient in a Mincer regression hence yields an estimate of the internal rate of return to education.

While the Mincer model was developed to study employment, several of its basic components are also relevant for entrepreneurship. One can view entrepreneurial activities as a subset of the range of occupations available to individuals in the Mincer framework, with a compensating differential in profits required for forms of business that demand more years of education (Kolstad et al. 2014). The basic assumption that education has a positive impact on worker productivity likely also extends to productivity in entrepreneurship, but here there may be some important differences as well. As argued in Lazear (2004), an entrepreneur must be able to handle a variety of different tasks to stay in business, which requires a wide set of skills; entrepreneurs must in a sense be ‘jacks-of-all-trades’. This implies that ‘human capital investment patterns should differ between those who end up being entrepreneurs and those who end up working for others’ (ibid:208), with entrepreneurs having more generalised and employees more specialised education. As specialisation tends to increase through the course of formal education, this indicates that the ratio of returns to primary versus secondary or tertiary education should be higher for entrepreneurs than for employees.² Since primary education provides basic skills of literacy and numeracy required to acquire and use a number of other skills necessary to be a successful entrepreneur, returns to primary education are likely to be positive and economically substantial for entrepreneurs. The returns are also likely to be more substantial for groups that have had little exposure to formal primary education, since there is more of an untapped potential for efficient entrepreneurship in such groups.

The main hypothesis tested in this article is hence that returns to primary education for entrepreneurs are significantly positive and economically substantial, and in particular for groups that have had little access to primary education. To test this hypothesis requires data from a context in which there is substantial

variation in the extent to which people have primary education. Data from a developed country would hence not be appropriate, and we choose instead to use data from a developing country with relatively low average education levels. Malawi is a least developed country of 15.3 million inhabitants where almost 30 % of the official school age children do not start primary school. In the 8-4-4 education system of the country, only 25 % have completed 8 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). In terms of entrepreneurship, the private sector in Malawi is small. While there are some larger enterprises mainly in the Southern town of Blantyre, most entrepreneurs are subsistence entrepreneurs in the informal sector (mainly in petty trade, fishery and simple service industries), and farming remains the most important occupation.³ The country hence provides a highly relevant site in which to test the effects of primary education on entrepreneurial outcomes.

In existing empirical analyses of the entrepreneurial returns to education, most studies use some variant of the Mincer (1974) equation as their specification. Ordinary least squares (OLS) is then used to estimate Eq. (1): The log of profits of enterprise i is regressed on the years of 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 enterprise- and owner-specific variables X_i .

$$\ln(\text{profits}_i) = \alpha + \beta_1 \text{age}_i + \beta_2 (\text{age}_i)^2 + \beta_3 (\text{education}_i) + \gamma X_i + \varepsilon_i \quad (1)$$

The main problem in estimating Eq. (1) is that there may be selection on unobservables into (a) education and (b) entrepreneurship. In terms of the first challenge, if education is correlated with some unobserved element of the profit equation, OLS estimates are not consistent. Unobserved innate ability may for instance

² Consistent with these arguments, Barouni and Broecke (2014) estimate wage returns of education for 12 African countries, and find a substantially higher effect of secondary and tertiary education than for primary education.

³ Agriculture constitutes 36 % of GDP in Malawi, and almost 60 % of Malawi’s exports stem from tobacco (Republic of Malawi/World Bank 2006). The expansion of the private sector is an aim of domestic industrial policy (IMF 2007; Record 2007), and education has been suggested as one possible means to making the private sector more profitable and productive (Republic of Malawi/World Bank 2006).

impact positively on both education and profits, leading to an upward bias in OLS estimates of the returns to education. In principle, however, OLS estimates could be biased either way depending on the unobserved variable and its relation to education and profits. Dickson (2013), following Lang (1993), suggests that unobserved discount rates may relate negatively to education and positively to wages (or in our case profits) as more impatient individuals have more determination and ambition to make money, or interact with unobserved ability in more complex ways that generate a potential downward bias in OLS estimates.

We address the endogeneity of education by using distance to school (measured in minutes) as an instrumental variable. 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. 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. We hence expect (and find) a negative relation between distance to school and years of education. Our empirical approach is similar to that taken by Card (1995) in studying the wage effect of higher education in the USA, and Söderbom et al. (2006) in analysing wage returns to education in Kenya and Tanzania.

Our main results are hence based on a standard instrument variable regression reflected in Eqs. (2) and (3). The education of entrepreneur i is first regressed on an instrument Z_i and a set of covariates X_i (including age and age squared). Predicted education values from this estimation are then used in the profit equation, in addition to the covariates X_i .

$$\text{Education}_i = \alpha_1 + \beta_1 \text{distance}_i + \gamma_1 X_i + v_i \quad (2)$$

$$\ln(\text{profits}_i) = \alpha_2 + \beta_2 \text{education}(\text{predicted})_i + \gamma_2 X_i + \varepsilon_i \quad (3)$$

We would argue that our instrument for education is valid in the Malawi context, i.e., it has no effect on entrepreneurial profits through other variables than education. 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 school instead of relocating the entire household. The problem of mobility is thus more applicable to developed countries such as the USA where Card (1995) originally employed a distance instrument, than to Malawi.

The second methodological challenge is that there is selection on unobservables into entrepreneurship. The basic challenge in this case is that 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, just as in the case of having an endogenous regressor. 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.

To test the robustness of our main instrument variable results, we perform additional estimations where we also address selection. We use access to land per household member as our instrument for being an entrepreneur, for the following reason: Subsistence farming is the most common form of activity among households in Malawi. 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. 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 in the more densely populated areas in the South of the country.

Since our data suggest 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.⁴ As for the validity of our selection variable, it is unlikely that people with greater unobserved entrepreneurial ability choose to live on smaller land plots, as there is limited mobility and trade in land. This is backed up by the fact that we see little complete specialisation in terms of occupation in Malawi, and the average entrepreneur spends considerable time on farming activities (a point to which we return in Sects. 3 and 4.2).⁵

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. In our robustness section, we apply the approach outlined by Wooldridge (2002, Sect. 17.4.2). This is a three-stage estimation procedure, where the first stage is a probit regression of entrepreneurship using access to land and its square as instruments (Eq. 4 below). The predicted Mills ratio from the probit regression is then used to correct for selection bias in a subsequent instrumental variable regression, where we use

distance to school as an instrument for education (Eqs. 5 and 6).⁶ In addition, all three equations contain individual-specific control variables X_{1i} (including age and age squared), and Eqs. 3 and 4 contain enterprise specific controls X_{2i} .

$$\text{Entrepreneur}_j = (\alpha_1 + \beta_{11}\text{distance}_j + \beta_{12}\text{land}_j + \beta_{13}(\text{land}_j)^2 + \gamma_{11}X_{1j} + \nu_{1j} > 0) \quad (4)$$

$$\text{Education}_i = \alpha_2 + \beta_{21}\text{distance}_i + \beta_{22}\text{land}_i + \beta_{23}(\text{land}_i)^2 + \beta_{24}\text{Mills}_i + \gamma_{21}X_{1i} + \gamma_{22}X_{2i} + \nu_{2i} \quad (5)$$

$$\ln(\text{profits}_i) = \alpha_3 + \beta_{31}\text{education}(\text{predicted})_i + \beta_{32}\text{Mills}_i + \gamma_{31}X_{1i} + \gamma_{32}X_{2i} + \varepsilon_i \quad (6)$$

Our exclusion restriction is that neither distance to school nor access to land feature in the profit equation. All the instruments, however, feature in both Eqs. 4 and 5. The reason for including distance in the probit equation is to avoid bias in the estimates, maintaining $\nu_{1j} \sim N(0, 1)$. Not omitting relevant variables is crucial in nonlinear 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, Eq. 5 needs to contain information from one more dimension than Eq. 6. By including both distance and access to land (in addition to the Mills ratio), we ensure that Eq. 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.

⁴ See for instance Orr (2000) on the ownership of Malawian estates.

⁵ While land may in principle be used as collateral for loans, facilitating access to capital, additional regressions showed no significant relationship between access to land and business related loans.

⁶ While one may argue that the education decision is usually made before the entrepreneurship decision, reversing the order of the first two stages produces very similar estimates of returns to education.

3 Data

The data used in this paper are taken from the Malawi Second Integrated Household Survey (IHS-2) 2004–2005.⁷ The survey covers 11,280 households and 52,707 individuals. The survey includes a module on entrepreneurship comprising 3,913 non-agricultural enterprises. Some individuals own more than one enterprise, and some enterprises have more than one owner. In order to merge the enterprise module with other modules, we have excluded enterprises with more than one owner and randomly selected one enterprise where an individual owns several. This reduces the number of enterprises to 3,556. Excluding entrepreneurs under 18 years and large estate owners cuts the sample to 3,287 enterprises. Due to missing data for our main variables (in particular the distance to school variable), this number is further reduced to 1,900 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 the robustness section (Sect. 4.2).

All the variables used for the main estimations are summarised in Table 1. As our dependent variable, we use the log of the monthly profits reported by the owner.⁸ 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 enterprise- and individual-specific controls.⁹ Distance, our instrument

⁷ See www.worldbank.org/html/prdph/lsm/s/country/malawi04/docs/IHS2%20Basic%20Information.pdf for further documentation.

⁸ While one may question the accuracy of reported profits, this appears to be the best available indicator of entrepreneurial success (cf. Mel et al 2009). We have checked the consistency of this variable with reported revenues less costs, and the correlation is high (0.81). Using the log of profits means we lose 95 non-positive observations.

⁹ Rather than age, some previous studies use age minus years of schooling –6 as a measure of labour market experience, doing so does not change results. We have chosen not to include industry dummies in our estimations, as these are likely to be endogenously determined and influenced by education (cf. e.g. Wang 2013). 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

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	

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.

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 (1 USD = 150 KZ), and the average enterprise employs 0.22 non-household members. The median enterprise has profits of about 80 USD, so the enterprises in our sample are generally small. There is, however, a great deal of variation. The largest enterprise in terms of profits earns more than

Footnote 9 continued

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 2 Summary statistics

Variable	Observations	Mean	SE	Min	Max
Profits	1,900	3,406.16	447.41	10	420,000.0
Education	1,900	4.69	0.12	0	17.0
Firm age	1,900	5.88	0.20	0	54.0
Registered	1,900	0.06	0.01	0	1.0
Firm size	1,900	0.22	0.08	0	70.0
Urban	1,900	0.15	0.01	0	1.0
Age	1,900	39.02	0.33	18	98.0
Male	1,900	0.55	0.01	0	1.0
Chronic illness	1,900	0.16	0.01	0	1.0
North	1,900	0.09	0.01	0	1.0
Centre	1,900	0.38	0.01	0	1.0
Distance	1,900	22.11	0.63	0	120.0
Land	1,900	0.35	0.01	0	1.2

Profits are in Malawi Kwacha earned over the past month. Education is years of schooling. Firm age and age are in years, firm size is 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

33,000 USD annually, and the largest enterprise in terms of employment employs 70 people. The average individual in our sample has between 4 and 5 years of education.

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

Diversification is a common livelihood strategy in Malawi, and though 57 % of our sample report using more than half their time on entrepreneurship, almost 35 % use most of their time on agriculture. Correspondingly, the average entrepreneur spends about 20 h a week running the enterprise, 11 h on farming, and 3 h on other activities. In a separate survey question where individuals are asked to name their

main activity, only 26 % report being self-employed, while more than 40 % 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 characterised as entrepreneurs, a point we return to in Sect. 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 14,829 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.¹⁰ 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

¹⁰ In the full IHS2 data set, almost 50 % of urban inhabitants live in a household that has access to land (compared to more than 95 % in rural areas). The proportions are about the same in our sample. There is hence no reason to believe that the access to land instrument is not relevant in urban areas. Running the estimation of the selection equation for only urban dwellers, access to land and its square displays much the same relation to entrepreneurship status as in the full sample.

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 1,900 entrepreneurs in the main sample with the 1,656 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 enterprises from each region, there appears to be significant overrepresentation of manufacturing enterprises, and underrepresentation of primary sector and service enterprises in our sample when compared to the excluded enterprises.

4 Results

4.1 Main results

The results from our main instrumental variable regression, using distance to school as an instrument for education, are presented in Table 3. The first two columns show the results from the first and second stage of the instrumental variable regression, respectively, and third column reports the corresponding OLS results for comparison. As can be seen from the first column, distance to school is negatively related to the number of years of education, as expected. The instrument is strong, the F value for the distance variable in the first stage is above 23, as reported in the last row of the table, and well above the conventionally required level of 10 (cf. Staiger and Stock 1997). Our main result is at the top of column two and shows that an added year of education increases profits by 20.7 %. This estimate is highly significant and significantly larger ($p < 0.031$) than the OLS estimate of 6 % reported in the last column.

The survey structure of our data means that a number of the standard instrumental variable tests cannot be performed. However, the estimates do not differ that much if we run our estimations without the

Table 3 Main results, instrumental variable regression

Dependent variable	First stage Education	Second stage Profit (logged)	OLS Profit (logged)
Education		0.207*** (0.068)	0.060*** (0.009)
Firm age	−0.039*** (0.012)	0.030*** (0.006)	0.024*** (0.005)
Registered	1.168*** (0.350)	0.705*** (0.185)	0.886*** (0.161)
Firm size	0.113*** (0.032)	0.047 (0.032)	0.063* (0.035)
Urban	2.517*** (0.345)	0.428** (0.204)	0.831*** (0.092)
Age	−0.125*** (0.033)	0.042*** (0.013)	0.025** (0.010)
Age squared	0.001 (0.000)	−0.000*** (0.000)	−0.000*** (0.000)
Male	1.626*** (0.177)	0.424*** (0.126)	0.661*** (0.065)
Chronic illness	0.026 (0.237)	−0.165* (0.087)	−0.156* (0.083)
North	1.837*** (0.257)	−0.218 (0.169)	0.072 (0.104)
Centre	0.110 (0.230)	−0.043 (0.084)	−0.017 (0.082)
Distance	−0.027*** (0.006)		
Constant	7.895*** (0.835)	4.618*** (0.529)	5.640*** (0.241)
r ²	0.239	0.136	0.268
N	1,900	1,900	1,900
F instrument	23.149		

Standard errors in parentheses

*** Significance at the 1 % level, ** at 5 %, * at 10 %

survey structure. The instrumental variable estimate of the effect of education then becomes 21.5 %. A Durbin-Wu-Hausman test rejects the hypothesis that education is exogenous ($p < 0.016$). The first-stage F -statistic for the distance instrument is 27.74, well above the 16.38 critical value of the Stock and Yogo (2005) weak identification test with 10 % maximal IV size (meaning the bias of the IV estimate should not be >10 % of the bias of the OLS estimate). Unsurprisingly, weak instrument robust inference tests therefore also confirm our main results; based on the Anderson–

Rubin test, the coefficient of education is both significantly different from zero ($p < 0.002$) and from the OLS estimate ($p < 0.017$). Since we have only one instrument, we cannot use overidentification tests to assess instrument validity.

The results for the covariates conform largely to expectations. Older, registered firms have higher profits, but there is no significant relation between firm size and profitability. Urban, older, male and healthier entrepreneurs run more profitable businesses, with a diminishing marginal effect for age. Owners of registered and larger firms have more education, as do entrepreneurs who are male, reside in urban areas, and in the Northern region. Older entrepreneurs and owners of older firms have less education, which likely reflects education levels in the population going up over time.

Our main hypothesis is that returns to primary education for entrepreneurs are significantly positive and economically substantial for groups that have had little access to primary education. We now proceed to show that returns to primary education for these groups is what our instrument variable estimate is picking up, using a modified version of the approach developed by Moffitt (2008). The methodological background for this is the observation that if there are heterogeneous returns to education in a population, an instrumental variable estimate picks up not the average treatment effect of education across the population, but a local treatment effect for the subgroups of the population whose education are affected by the instrument. In other words, we need to look into for whom the distance instrument has the greatest effect on education to identify the groups for which our estimate is relevant.

Formal analysis of which groups are moved by the distance instrument is complex in our 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 1 year of education, at least 2 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. Intuitively, switching the instrument on and off, and looking at how this affect participation probabilities in education, tells us where the instrument has the greatest effect on education.

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 Fig. 1. The right hand panel of the figure shows the range of participation probabilities for each decile of these probabilities. The dark grey boxes convey the ranges when distance is held at its mean (baseline participation rates), while the light grey boxes represent the ranges when distance is allowed to vary (predicted or actual participation rates). For instance, for the 10 % 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 dark grey box), and from just above zero to approximately 0.3 when distance is allowed to vary (cf. tick marks light grey 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

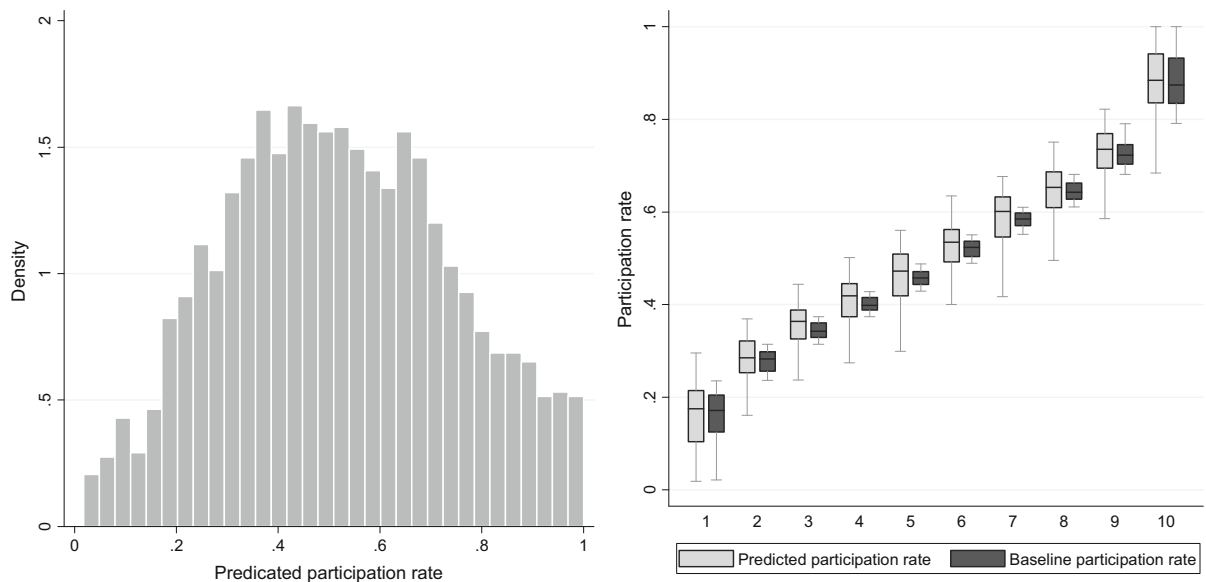


Fig. 1 Predicted and baseline participation in education at and above grade 5. *Note* In the right panel, the upper and lower points of the rectangles are the 75th and 25th % points of the

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 is 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 is affected by our particular instrument.

In simpler terms, Fig. 1 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

distribution, respectively. The horizontal lines inside the rectangles are the medians. Upper and lower tick marks are upper and lower ranges

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 support for our hypothesis of significant and high returns to primary education for groups with limited access to such education. Our instrumental variable estimate of a return of more than 20 % to an added year of education identifies the effect of primary education, specifically for groups with medium participation probabilities in this form of education. With heterogeneous effects of education, this also means that we cannot from our estimates say anything about the effects on entrepreneurial profits of secondary or tertiary education. Our results also do not speak to the effect of primary education on groups with a high (or very low) probability of getting primary education. Nevertheless, our sizeable effect of primary education for groups with medium participation rates is consistent with the results of Moffit (2008), who shows that the effect of education decreases with participation rates, i.e., returns to education go down as larger parts of a group, and hence more lower-return individuals, are drawn into education. It is likely that the differences in our estimates from those of other studies using similar

methods reflects differences in how the instruments in these other studies affect education, but since few of these studies discuss heterogeneous effects and local average treatment effects, it is difficult to be more specific on this point.

4.2 Robustness

As noted, there is a possibility of a selection bias as the choice of becoming an entrepreneur is also endogenous. Our results are, however, robust to this type of endogeneity. Table 4 shows the results from the three-stage estimation which combines a Heckman selection approach with an instrumental variable regression. The results from the first stage of the three-stage estimation (first column) show that access to land and its square provide strong 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. While the selection instruments are strong in the first stage, the generated Mills ratio is not significant in the subsequent to stages. This suggests that there is no selection bias of this type in our results, and consistent with this, the estimated effect of education on profits is in the same order as in the instrumental variable regression presented earlier, at 21.4 %.

Adding land area and its square as instruments in the second stage does, however, create a challenge of instrument weakness. The F value of the excluded instruments in the second stage is 3.55, which is below critical values for reasonable bias ratios on the Stock and Yogo (2005) test. As previously noted, the survey structure of our data does not permit standard instrument variable tests to be performed, so we perform these tests with no survey structure imposed. Since in the three-stage estimation we have an overidentified model, a conditional likelihood ratio test has better properties for weak instrument inference than an Anderson–Rubin test. Results from the conditional likelihood ratio test show that the estimated effect of education is both significantly different from zero ($p < 0.033$) and from the corresponding OLS estimate ($p < 0.094$). In addition, the limited information likelihood estimator, which with an overidentified

Table 4 Results from three-stage estimation

Dependent variable	First stage	Second stage	Third stage
	Entrepreneurship	Education	Profit (logged)
Education			0.214* (0.123)
Firm age		−0.038*** (0.012)	0.030*** (0.007)
Registered		1.150*** (0.399)	0.693** (0.305)
Firm size		0.109 (0.212)	0.046 (0.244)
Urban	−0.001 (0.062)	2.176*** (0.376)	0.390 (0.265)
Age	0.083*** (0.006)	0.790 (0.997)	0.027 (0.093)
Age squared	−0.001*** (0.000)	−0.009 (0.011)	−0.000 (0.001)
Male	0.221*** (0.034)	3.979 (2.594)	0.372*** (0.110)
Chronic illness	0.158*** (0.048)	1.664 (1.820)	−0.195 (0.175)
North	−0.230*** (0.059)	−0.553 (2.710)	−0.190 (0.459)
Centre	−0.179*** (0.044)	−1.751 (2.079)	−0.012 (0.222)
Distance	−0.003*** (0.001)	−0.057* (0.034)	
Land	−0.520*** (0.184)	−8.698 (6.166)	
Land squared	0.365** (0.164)	6.958 (4.414)	
Mills ratio		13.293 (14.484)	−0.228 (1.153)
Constant	−2.665*** (0.132)	−30.195 (42.033)	5.267 (4.317)
r2		0.245	0.123
N	14,829	1,900	1,900
F instruments		3.55	

Jackknifed standard errors in parentheses

*** Significance at the 1 % level, ** at 5 %, * at 10 %

model has less bias than two-stage least squares (2SLS) in finite samples, suggests an effect of education that is higher than the 2SLS estimate, at

24.7 %. With an overidentified model and no survey structure, we can also perform overidentification tests. The Sargan overidentification test in our case does not reject the validity of our instruments ($p > 0.298$). In other words, our main results from the instrumental variable regression presented in Sect. 4.1 are robust to endogenous selection into entrepreneurship.

The results presented above 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. The same applies if we control for the time school-going household members spend working on the household business, or the number of household members working on the business, addressing any concern that individuals living close to a school may have more time to spend working on the business. 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 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. To assess this issue, we ran further estimations where we substituted mean distance at the primary sampling unit 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. Results from instrumental variable estimation on the resulting sample of 2,962 observations proved almost identical to results for the original sample of 1,900 entrepreneurs, with an estimated impact of an added year of education at 20.6 %. Our results therefore do not seem to be affected by limitations in the sample incurred through missing observations for our instrument for education.

In Sect. 3, we saw that 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 characterised as entrepreneurs. As it turns out, however, this is not much of a problem. In Table 4, 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 h per week on such activities (second column), and individuals who name self-employment as their main activity (third column). 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.

5 Concluding remarks

The results of this paper show that the private returns to entrepreneurs from certain forms of education can be substantial. Specifically, primary education has a large effect on the profits of entrepreneurs, for groups of entrepreneurs whose participation in this form of education is limited. This is consistent with theoretical arguments that generalised knowledge of the kind provided through primary education is important for entrepreneurs to acquire the diverse skills required to run a business successfully. While the same theoretical arguments imply that returns to secondary or tertiary education may be lower for entrepreneurs, our results do not really shed light on this, this would require a different set of instruments for education which affects participation at other levels of schooling. Similarly, returns to primary education are likely to be lower for groups whose participation in primary school is high, but our instrument does not pick up effects for these other groups. We note that lower estimates for the effect of education from other studies may reflect instruments working at other levels of education, or for groups with higher levels of participation in primary

Table 5 IV-regressions, entrepreneurship main activity

	More than 50 % of time in business	More than 50 % and 7 h/week in business	Main activity self- employment
Education	0.200** (0.09)	0.204** (0.09)	0.024 (0.14)
r ²	0.121	0.105	0.285
N	1,027	1,001	493

Dependent variable: ln(profits)

Controls included but not reported. Standard errors in parentheses

*** Significance at the 1 % level, ** at 5 %, * at 10 %

school, the latter being a particularly relevant possibility for studies of developed economies. Mapping returns across a wider range of education levels and at various participation rates is an important area for further studies (Table 5).

Methodologically, this paper addresses both the challenge that education is endogenous, and the problem of selection into entrepreneurship. Moreover, we build on and modify previous methods to more precisely determine what type of effect an instrument variable estimate is identifying. The application of these methods to the question of entrepreneurial activity is to our knowledge also novel. This can be viewed as one step towards greater methodological convergence and comparability with the literature on education and employment.

Acknowledgments The authors thank Erik Ø. Sørensen, Magnus Hatlebakk, Bertil Tungodden, Eyolf Jul-Larsen, Øivind Anti Nilsen, and two anonymous reviewers 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.

References

- Angrist, J. D., & 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, 430–442.
- Barouni, M., & Broecke, S. (2014). The returns to education in Africa: Some new estimates. *Journal of Development Studies*. doi:10.1080/00220388.2014.936394.

- Belzil, C. (2007). The return to schooling in structural dynamic models: A survey. *European Economic Review*, 51, 1059–1105.
- Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In L. N. Christofides, E. K. Grant, & R. Swidinsky (Eds.), *Aspects of labour market behaviour: Essays in honour of John Vanderkamp* (pp. 201–222). Toronto: University of Toronto Press.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5), 1127–1160.
- de Mel, S., McKenzie, D. J., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, 88(1), 19–31.
- Dickson, M. (2013). The causal effect of education on wages revisited. *Oxford Bulletin of Economics and Statistics*, 75(4), 477–498.
- Dickson, P. H., Solomon, G. T., & Weaver, K. M. (2008). Entrepreneurial selection and success: does education matter? *Journal of Small Businesses and Enterprise Development*, 15(2), 239–258.
- Harmon, C., Oosterbeek, H., & 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), 153–161.
- IMF. (2007). Malawi: Poverty reduction strategy paper—Growth and development strategy, IMF Country Report No. 07/55, Washington DC: International Monetary Fund.
- Jul-Larsen, E., & 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.
- Kolstad, I., Wiig, A., & Moazzem, K. G. (2014). Returns to education among entrepreneurs in Bangladesh. *Journal of Asian Economics*, 34, 54–65.
- Lang, K. (1993). *Ability bias, discount rate bias and the return to education*, mimeo. Boston: Boston University.
- Lazear, Ed. (2004). Balanced skills and entrepreneurship. *American Economic Review*, 94(2), 208–211.
- McKenzie, D., & Woodruff, C. (2014). What are we learning from business training and entrepreneurship evaluations around the developing world? *World Bank Research Observer*, 29(1), 48–82.
- Mincer, J. (1958). Investment in human-capital and personal income-distribution. *Journal of Political Economy*, 66(4), 281–302.
- Mincer, J. (1974). *Schooling, experience, and earnings*. New York: Columbia University Press.
- Mkandawire, F. R., & Mulera, D. (2010). “Malawi” in Peterson. In P. E. Baker & B. McGaw (Eds.), *International encyclopedia of education* (pp. 666–672). Amsterdam: Elsevier.
- Moffitt, R. (2008). *Estimating marginal treatment effects in heterogeneous populations*, mimeo, revised January 2008, Johns Hopkins University, Baltimore.
- 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 November 29 2010.
- Söderbom, M., Teal, F., Wambugu, A., & Kahyarara, G. (2006). The dynamics of returns to education in Kenyan and Tanzanian manufacturing. *Oxford Bulletin of Economics and Statistics*, 68(3), 261–288.
- Staiger, D., & Stock, J. H. (1997). Instrumental variable regressions with weak instruments. *Econometrica*, 65(3), 557–586.
- Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge: Cambridge University Press.
- van der Sluis, J., & van Praag, M. (2008). Education and entrepreneurship selection and performance: A review of the empirical literature. *Journal of Economic Surveys*, 22(5), 795–841.
- van der Sluis, J., van Praag, M., & van Witteloostuijn, A. (2007). *Why are the returns to education higher for entrepreneurs than for employees?* IZA Discussion Papers 3058, Bonn: Institute for Labour.
- van der Sluis, J., van Praag, M., & 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.
- Wang, L. (2013). How does education affect the earnings distribution in urban China? *Oxford Bulletin of Economics and Statistics*, 75(3), 435–454.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.