TI 2003-046/3
Tinbergen Institute Discussion Paper

Entrepreneurship Selection and Performance

Version November 2004

Justin van der Sluis[1]
C. Mirjam van Praag[1]
Wim Vijverberg[2]

[1] *Department of Economics, Faculty of Economics and Econometrics, University of Amsterdam, and Tinbergen Institute,*
[2] *University of Texas at Dallas, and IZA.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Entrepreneurship Selection and Performance:
A Meta-analysis of the Impact of Education in Less Developed Countries.

JUSTIN VAN DER SLUIS
University of Amsterdam and Tinbergen Institute
MIRJAM VAN PRAAG
University of Amsterdam and Tinbergen Institute
WIM VIJVERBERG*
University of Texas at Dallas and IZA
This draft: 1 November 2004

Abstract

This paper provides an analytical review of empirical studies into the impact of schooling on entrepreneurship selection and performance in less developed countries. We analyze the variation of this impact across various characteristics of studies. We find that a marginal year of schooling raises enterprise income by an average of 5.5 percent, which is closely similar to the average return in industrial countries. The return varies by gender, rural/urban residence, and the share of agriculture in the economy. Furthermore, more educated workers typically end up in wage employment but also prefer non-farm entrepreneurship to farming. The education effect that separates workers from self-employment into wage employment is stronger for women, possibly stronger in urban areas, and also stronger in the least developed economies where agriculture is more dominant and literacy rates are lower.

Keywords: Meta-analysis, schooling, education, entrepreneurship, self-employment, performance, occupational choice, developing countries
JEL-codes: J23, J24, J31, M13, O15.

*Corresponding author. Address: School of Social Sciences, University of Texas at Dallas, P.O. Box 830688, Richardson, TX 75083-0688, U.S.A. E-mail: vijver@utdallas.edu; telephone: +1-972-883-2042. We gratefully acknowledge the financial support by The World Bank as well as comments from the editor and referees. The findings, interpretations, and conclusions expressed in this paper are those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to the members of its Board of Executive Directors or the countries they represent.

1. Introduction

Human capital theory states that one of the main drivers of the investment in schooling is the notion that schooling produces skills that raise worker productivity and income. Supposedly, therefore, education is beneficial for economic growth. Thus, in the development literature, great effort has been spent on the quantification of the rate of return to education. Psacharopoulos (1994) has brought together the evidence from 140 studies from all over the world in a way that allows both international comparisons and trend analyses. However, almost without exception, returns to schooling refer to the returns *employees* generate from their years at school (Bennell, 1996). In contrast, the literature that pertains to the measurement of the rate of return to schooling in *entrepreneurship* or, its most common empirical equivalent, self-employment is actually still poorly defined, both for developed countries (Van der Sluis, Van Praag, and Vijverberg, 2003) and for developing countries, to which we confine ourselves in this study.[1]

Our objective is to assess whether and to what extent there is an effect of schooling on entrepreneurship entry and performance in developing countries.[2] We bring together more than 80 studies that measure these effects. A careful reading of the studies reveals that a simple summary is problematic because variable definitions, empirical models and data sources differ so much. Thus, using factors that characterize each study, we investigate the education effects by means of meta-analytical techniques. A meta-analytical approach yields a quantitative assessment of the literature that generates benefits that differ from the more traditional literature survey: it forces precise comparisons of the utilized research practices and methodologies, and it yields a quantitative explanation of the variation among the many research outcomes.[3] Thus, we gain a deeper understanding of the gaps and opportunities in this fairly poorly developed area of entrepreneurship research. We compare our results to both the better-developed studies pertaining to employees' returns to schooling and to the just as poorly developed studies pertaining to entrepreneurs in developed countries.

This topic clearly is of great relevance and interest. Researchers and practitioners alike are fully aware of the (potential) contributions of entrepreneurs to the economy. Entrepreneurs generate a substantial part of the national income and employment in most countries. Small enterprises form a large flexible buffer

[1] In a parallel paper, Van Der Sluis, Van Praag, and Vijverberg (2003) examine the same set of relationships that exist in industrial countries. Since the general level of schooling is lower and a substantial portion of the labor force works on the farm, an analysis of the developing country studies is not easily fused with one of industrial country studies.

[2] For more general surveys on entrepreneurship, see, e.g., Mead and Liedholm (1998), King and McGrath (1999), and Kiggundu (2002). For a good survey on the separate literature of the impact of education on farm production, see Jamison and Lau (1982), and Lockheed, Jamison, and Lau (1987). This literature is also summarized in a meta analysis by Phillips (1994). In this paper, we do not cover the role of education in agricultural activities.

[3] Meta analysis offers an overview of the literature that complements the more standard literature survey, which highlights particular high-quality pieces of research, and thus shines a different light on the same matter. Due to space constraints, we are not able to describe individual studies in detail, which in itself will be a useful contribution.

between salaried employment and incorporated business. Moreover, entrepreneurship may generate benefits for society through the development and maintenance of human and social capital that inherently occurs when entrepreneurial activity takes place.

More specifically, in developing countries, the size and economic importance of the entrepreneurial sector have been underestimated for a long time. In line with Lewis (1954) and Ranis and Fei (1964), studies of economic development emphasized agriculture and industry. The work by Harris and Todaro (1969) illustrated that migrants from agriculture to industry may face a period of unemployment or, if unemployment is unaffordable, may be forced to provide for themselves through a low-productivity household enterprise. The latter idea took more shape when an International Labour Office (1972) report defined the so-called informal sector, which has proved to be one of the more influential concepts in development economics, for right or for wrong reasons (e.g., House 1984; Peattie 1987; Mead and Morrison 1996). As defined, the informal sector covered all economic activity that was hidden away from official oversight and tended to be not very productive. Soon, the perception ruled that the informal sector consisted mainly of small enterprises, undesirable for anyone striving to make a decent living and unable to make any significant contribution to national economic growth. This idea was a bit of a caricature. It may be true that many small enterprises appear unproductive, but they do make useful contributions to household income and some do blossom into large operations. Only when large-scale household surveys were administered did it become clear that there was indeed much hidden entrepreneurial activity.[4]

Realizing how extensive household entrepreneurship really was, the question arose what determines the income from household enterprises: for instance, what does schooling contribute? Moreover, if employment in the formal sector is so much more desirable, why do people want to start a household enterprise? There are good (theoretical) reasons to presume that education plays a role as a potential determinant of entrepreneurship selection and entrepreneurial success (Section 2), in both the formal and informal sector of a developing economy. If education leads to a higher quality of entrepreneurial performance and to a greater number of entrepreneurs, this justifies appropriate investments in education. Therefore, research is required into whether and how this is achieved. As a starting point, we wish to assess whether and to what extent schooling currently affects entrepreneurship selection and performance. As a by-product we will be able to evaluate the state-of-the-art of research of this kind.

The paper proceeds as follows. Section 2 briefly summarizes economic theory about the relationship between entrepreneurship entry, performance and educational attainment. Section 3 describes the data gathering process and some of the characteristics of the resulting database. Section 4 describes the current

[4] Studies of entrepreneurship relied on surveys sampled from lists of registered (and therefore larger-scale) enterprises and therefore presented a biased view of entrepreneurship: only successful entrepreneurs would grow and eventually register. For an example of Côte d'Ivoire, see Vijverberg (1992); for an example across the countries of Botswana, Kenya, Malawi, Swaziland and Zimbabwe, see Mead (1994), and Mead and Liedholm (1998).

practice of research into the relationship between schooling and entrepreneurship entry and performance. Section 5 describes the construction of subsamples for the meta-analysis used to explain cross-study differences in the relationship between schooling and entrepreneurship (performance and entry). Section 6 discusses the results from the meta-analysis pertaining to performance and compares them to the relevant findings concerning employees and entrepreneurs in developed countries. Section 7 is similar to section 6 and pertains to the relationship between schooling and entrepreneurship selection. Section 8 concludes.

2. Economic Theory

Theory offers several determinants of entrepreneurship selection and performance. Among them are risk attitude, access to capital, various types of labor market experience, economic conditions, family background, psychological traits, income diversification, access to credit, and, last but not least, education. This section briefly reviews the theoretical arguments pertaining to the relationship between schooling and entrepreneurship that are prevailing in economics.

Education as a determinant of entrepreneurship selection

Le (1999) argues that there are several channels through which the level of education might influence the propensity to become self-employed. Calvo and Wellisz (1980), inspired by Lucas' general equilibrium model (1978), explain the impact of educational attainment on the probability of selection into an entrepreneurial position through managerial ability. Education would enhance managerial ability, which in turn increases the probability of entrepreneurship.

The other channel of influence as indicated by Le causes an opposite, negative effect of education on entrepreneurship selection. It points to the possibility that higher levels of education might generate better outside options (i.e., more lucrative paid wage employment under better working conditions) and thus decrease the likelihood of entrepreneurship as the preferred choice. It is yet unclear what the predicted effect of these offsetting forces on the effect of schooling on entrepreneurship selection might be.

Education as a determinant of entrepreneurship performance

According to the Mincerian specification of the determinants of individual earnings, the main factors affecting earnings are schooling and experience. This specification and the implied positive returns to schooling have found empirical support in the wage sector. Although there is little reason to doubt its applicability in other occupational sectors such as entrepreneurship, so far, little systematic work has been done on this subject as yet.

Schooling is acknowledged not only for its productive effect on the quality or quantity of labor supplied, as assumed by Mincer, it also has value as a signal of productive ability in labor markets without complete information (Spence, 1973; Riley, 2002).

Integrated models of choice and performance

Another type of models, known as structural models, simultaneously explains the occupational choice and performance of labor market participants. In these models, the actual division between entrepreneurs and wage labor turns on the distribution of individual characteristics among the utility maximizing population. In Lucas (1978) and Van Praag and Cramer (2001), this characteristic is individual entrepreneurial ability as determined by, for instance, education. In such models, education generates higher levels of (expected) entrepreneurial ability that, in turn, causes higher levels of entrepreneurial performance (in terms of profit and firm size). This higher level of expected performance and thereby (psychological) income, increases the expected utility attached to entrepreneurship and thereby favors this occupational choice. Similarly, Vijverberg (1986, 1993) models occupational choice as a time allocation problem where people choose from different income-generating activities; see also Roy (1951). Across activities, education has a different effect on productivity. Moreover, people with different education backgrounds may have varying preferences for those activities. Thus, education impacts sorting outcomes but the net direction of the impact is an empirical matter.

Education may impact sorting outcomes in several other ways. First, it interacts with the seasonality of on-farm work. During the slack season, many seek off-farm employment, but the scarcity of jobs, partially due to the lack of tolerance of non-farm business ventures to seasonal fluctuations, forces farm workers to enter some sort of non-farm self-employment activity (Haggblade, Hazell and Brown, 1989; Lanjouw and Lanjouw, 2001). Manual jobs that lend themselves to short-term self-employment require less education: more educated workers establish themselves in more permanent activities. Second, households seek to diversify their income. Thus, they may operate a non-farm enterprise to offset uncertainty in farming outcomes, and the education of household members may decide who gets to do which work.[5] Third, as credit markets function poorly, non-farm enterprises raise the cash needed for the operation of and especially investment in the farm (Reardon et al. 2000; Lanjouw and Lanjouw, 2001). As well, education is associated with greater household wealth: the start-up of entrepreneurial activity is often financed with family assets rather than with loans, and loans themselves are easier to get if the household has some wealth to offer as security (e.g., Paulson and Townsend, 2001). In all this, education plays a role, as it helps people to perceive economic opportunities (Schultz, 1980).

Hence, there are many economic reasons why education impacts entrepreneurship choice and entrepreneurial performance. As Le (1999:386) notes, educational attainment is "one of the major theoretical determinants." We now turn to the database that contains information on the empirical studies pertinent to these issues.

[5] See Reardon et al. (2000), De Janvry and Sadoulet (2000), Lanjouw and Lanjouw (2001), Haggblade, Hazell, and Reardon (2002).

3. Construction of the Database

Meta-analysis is a quantitative tool that is applied to synthesize previous research findings that share common aspects that can be addressed in a statistical sense. The set of meta-analytical techniques have largely been developed and applied outside the academic field of economics in the medical and natural sciences. Rare examples of the application of these techniques in economics are Phillips (1994), Card and Krueger (1995), Ashenfelter et al. (1999), Nijkamp and Poot (2002), and Van der Sluis et al. (2003).

In building a database that can be analyzed by means of a meta-analysis, the first issue of concern is coverage, i.e., the extent to which the retrieved documents are representative of the literature (Nijkamp and Poot, 2002). Our aim is a complete coverage of empirical studies that estimate a quantified relationship between entrepreneurship (entry and/or performance) and education. However, as the relevant literature is widely scattered, several restrictions are imposed for studies to be included in the database: The studies must (*i*) be written in English; (*ii*) be written for an academic audience; (*iii*) pertain to less developed countries or countries in transition; and (*iv*) be published after 1980.[6]

We consider journal articles, book chapters, books, and working papers. This is a wide net to cast. Our first avenue of search is the Internet, with Web of Science as the primary source for published journal articles. Besides journal articles, we also include unpublished papers in the database.[7] The motivation for this is twofold: it is the only way of including the most recent research output, and it enlarges our sample. We consider working papers up to June 2003, the date at which we completed the construction of the database, as far as they are available on the Internet.[8] Our primary (virtual) search engines for working papers are the SSRN (Social Science Research Network), WOPEC (WOrking Papers in EConomics), and working papers series of well-known research institutes such as NBER, the World Bank, WIDER, IZA, and the Institute of the Study of African Economies at Oxford University. Our second avenue of search for both published and unpublished papers is a scan through the references of each sampled paper. Furthermore, the Web of Science allows one to do a search of citations. Thus, after an article is published, it is straightforward to find all other articles (in the journals covered) that refer to the studies that were already captured in our sample.

The second part of the references section shows the resulting selection of studies. We gathered 84 studies that each included at least one valid observation on either the quantified relationship between schooling and entry/selection (a transition to entrepreneurship) or the quantified relationship between

[6] The 1980 cutoff is imposed for practical reasons of access but is virtually innocuous since this literature got going only in the mid 1980s.

[7] Older working papers are particularly difficult to find. Consequently, the database contains only one working paper from before 1995.

[8] In order to prevent double counts and to preserve independence of observations, we check whether working papers have more recently appeared as publications (sometimes with a different title or authorship).

schooling and performance (earnings, duration, etc.). Altogether, the 84 studies yielded 203 observations. Among these, 129 (64 percent) examine performance, 19 (9 percent) investigate entry into entrepreneurship, and 55 (27 percent) specify the dependent variable as "being self-employed." The latter is a stock (rather than flow) variable that is a hybrid of entry (everyone who *is* self-employed has entered this occupational status) and performance (it generates an overrepresentation of survivors). We therefore keep "stock studies" as a separate category.

4. Description of Studies in the Database

With the database in hand, we now turn to a description of the studies contained in it. We consider facets like the definition of the primary variables of interest, such as entrepreneurial outcomes and education, as well as the type of data that were used, the analytical techniques that researchers employed, and so forth. A subsequent section will focus more explicitly on the evidence regarding the relationship between entrepreneurial outcomes and education.

Measurement of entrepreneurship, enterprise performance, and education

As it turns out, researchers of entrepreneurship do not adhere to uniform definitions of the key variables of interest in our study, i.e., entrepreneurship, enterprise performance, and education. This represents one of the challenges of performing a meta-analysis.

Empirical definitions of entrepreneurship are fairly well comparable to each other, though their meanings are much more prosaic than those in most (untestable) theories that refer to the innovative free mind of the resourceful spiritual entrepreneur. Indeed, most researchers empirically define entrepreneurs as self-employed. The next question is how one should model the entrepreneurship choice. Should one contrast the (characteristics of the) entrepreneur against everyone, against employees, or against (self-employed) farmers? Is it a binary choice, or should one view the entrepreneur as someone who chooses from many alternatives? Table 1, panel A, illustrates the wide variety of choices that researchers have made in this respect. Most studies look at non-farm entrepreneurship as a 0/1 option, whereas another large group of studies models it as a multinomial choice. But this is only part of the complexity. A binomial (0/1) model logically contrasts self-employment against a single alternative, but as seen in the table, the 38 binomial studies utilize six different alternatives. Furthermore, the multinomial models spread out over seven alternatives while omitting the two most frequently used options of the binomial models.[9] Apart from this is the discrepancy in the choice of unit of analysis: some studies look at the choice of individuals, other of households. Needless to say, there is little homogeneity among the studies.

[9] The column total of 62 reflects the number of log-odds combinations derived from the 26 studies using a multinomial model.

Similarly, the literature has not yet converged upon standardized definitions of performance and educational achievement. Panel B of Table 1 shows the various empirical definitions of performance that have been used: 70 of 129 (54 percent) of performance observations focus on self-employment earnings, defined in various ways. Another 16 percent look at inputs (typically employment) as a measure of size or growth, 15 percent examine duration or survival, which can be easily translated into each other.[10]

Panel C of Table 1 illustrates the great diversity in the use of education variables. The largest number of studies enters education as a straightforward linear variable reflecting years of schooling. Some studies embellish the relationship by means of squared years or by spline functions that allow a different slope at the different schooling levels. Many other studies use dummy variables indicating levels of schooling, rather than years needed to attain the various levels. Among these, there are a few studies that distinguish lower from upper elementary and lower from upper secondary schooling.[11] Moreover, a small number of studies distinguish training and apprenticeships as less formal ways of skill investment. Overall, it is important to keep in mind that schooling systems in developing countries are highly heterogeneous (Kurian, 1988). In some countries, cognitive skills by grade level are at a par with industrial countries; in other countries, basic reading, writing, and arithmetic skills are still suspect at the end of elementary school (e.g., Lee and Barro, 2001:485). The lack of uniformity in the measurement of schooling may generate additional problems for a quantified meta-analysis of the relationship between schooling and entrepreneurship; the variation in the quality of education will lead to a spread in measured effects.[12]

The variety of definitions of entrepreneurship, entrepreneur performance and education points to the need to be very careful in the design of the conceptual framework that synthesizes the available evidence in this field of research.

An additional complexity arises from the use of various estimation strategies. We distinguish structural studies from reduced form studies into the same relationship. Several authors have acknowledged that the selection into self-employment is an endogenous choice, dependent on the expected performance of the enterprise or on the utility from income; failing to account for such selectivity effects may well bias the

[10] Ten studies first derive a technical efficiency index from a production frontier analysis and then examine this index as an indicator of entrepreneurial success. Eight studies found other performance measure to analyze, such as self-employment income as a share of total household income, a private benefit cost ratio, the growth rate of profits, a business diversification index, and so forth.

[11] The motivation for this is threefold. One might speculate that there is a threshold effect of education, such that benefits are gained only when cognitive skills reach a minimum level that is achieved after, say, three years of schooling as it appears to be the case in agriculture (e.g., Phillips, 1994). Furthermore, given the low rates of schooling especially among older generations in many parts of the world, it makes sense to distinguish among these levels of schooling. In addition, unlike in some industrial countries, secondary schooling is often broken up into two levels (Kurian, 1988).

[12] Not reported in the table are three studies that count the number of people in the household in each of several education categories, one study that refers to the English-type O-level and A-level of education, and five studies that did not include any education variable but utilized information on training and/or apprenticeships.

estimated return to education. Those studies that we label as "structural" attempt to incorporate at least some kind of a deliberate occupational choice of labor force participants.[13] Nineteen percent (=24/129) of the performance observations are structural. Amongst the stock and entry studies this number is negligible. However, none of the studies attempts to address the endogeneity of the schooling decision of individuals. This is striking, because doing so is becoming common practice in studies that measure the effect of education for wage employees (see Ashenfelter et al., 1999).

-Insert Table 1 about here-

The studies and their data samples

Table 2 characterizes several aspects of the origin of the studies in the sample. Panel A starts off by considering the field of publication where the studies appeared. Of the 203 observations in the database, 161 have been published. As the field of entrepreneurship research is well known for its multidisciplinary character, and since researchers in the various disciplines may use different control variables, which may well have an impact on the estimated impact of education, we cover (and distinguish) five academic fields: general economics journals (13.7%), labor economics and education journals (9.3%), development economics journals (57.8%), small business and entrepreneurship journals (12.4%), and management and sociology journals (6.8%). Comparing the row proportions with the overall proportions, it appears that studies of structural performance are over-represented in the category of general economics and development economics journals, whereas stock studies are over-represented in labor and development economics journals. Reduced form performance models appear more frequently in the small business and entrepreneurship journals and in working papers. Studies of entrepreneurship entry are primarily found in the management and sociology category.

Panel B highlights some trends in the literature. In particular, analyses of entrepreneurship have become more popular recently, although this appears to have ebbed slightly in the last few years. Perhaps more important, however, is the evident trend in the nature of the research in this field: structural form studies are a more recent phenomenon.

[13] There are actually many choices that could fall under this heading. For example, it is often assumed that the individual is working anyway and that the only choice to be modeled is whether to be self-employed as opposed to working for a wage. However, this choice model could be augmented with the choice whether to work, a choice between working in the public sector or the private sector, a choice to work for a large corporate organization as opposed to a smaller business that offers a similar environment as one's own enterprise, a choice to find employment in an urban area as opposed to a more rural setting, a choice to quit schooling in favor of employment, etc. Obviously, there is no study that includes all of these features. The point is that structural studies attempt to remove the bias caused by ignoring one or perhaps several of these choices but that one could easily think of other omitted selectivity factors that still may bias the estimated returns to schooling. Therefore, while any comparison between reduced form and structural model estimates has obvious limitations, the comparison of various structural studies is still not entirely straightforward either.

Panel C shows the geographical distribution of studies into entrepreneurship entry and performance. We aggregate the countries studied into six groups, namely the Sub-Saharan Africa, North Africa and the Middle East, South Asia, East and Southeast Asia, Latin America, and East-European and Asian countries in transition. Sub-Saharan Africa dominates, contributing 36.5 percent (72/203) observations. This is followed by Latin America and the Caribbean with 31.5 percent (65/203). Studies that use data from countries in transition are all more recent.

Figure 1 depicts the sample size of each of the four different types of studies in our database, sorting these studies by sample size from smallest to largest. A number of studies did not report sample sizes; these are represented by the horizontal offset of the four graphs; e.g., 16 of the 105 studies on reduced form performance did not report the size of their sample. As is shown, stock and entry studies tend to use larger samples than performance studies, for the simple reason that they include the non-self-employed in order to study the entrepreneurship choice. Of the reduced form performance studies, 73 of the 89 that reported the sample size contained less than 1000 observations, going as low as only 30. The trend is towards using larger datasets: in general, the correlation between the size of the sample and the year the sample was gathered was positive (0.15).

-Insert Table 2 about here-

-Insert Figure 1 about here-

In doing research on entrepreneurship choice and performance, there are good reasons why one might want to study men and women separately: they face different constraints and act on different opportunities. At the same time, a household enterprise is often a joint activity of different household members, in which men and women work together for the benefit of not so much the enterprise but more explicitly the household. Among the studies included in the database, the stock and entry studies separate the sample by gender more often but they are also more likely to treat the entrepreneurship choice as a household outcome (Panel D of Table 2). More than half of the performance studies utilize samples that mix men and women, frequently because the unit of observation is not necessarily the entrepreneur but rather the enterprise, which may employ both men and women.[14]

5. Implementing meta-analysis

Constructing subsets for meta-analysis

Before turning to a description and explanation of the (average) effects of education on entrepreneurship performance and selection, we set out the procedures we will employ for the analysis of variation in the effects. We shall use regression techniques to "explain" the effect of schooling, referred to as *b*, by the

[14] The appropriate education measure then pertains to the entrepreneur, the leader of the enterprise.

various characteristics Z that we have gathered for each study; we estimate a model $b = Zb + e$. It is important to select subsets of studies for this regression analysis that are homogenous in the definitions used for education and for enterprise performance or entry/stock: the less variation there is across studies in the measurement of these variables, the more meaningful will be the result of the meta-analysis. However, insisting on strict homogeneity reduces the subset of eligible studies. In particular, the requirement of homogeneity conflicts with the fractured definitions of stock/entry, performance, and education. For example, studies that specify enterprise income in linear form end up in a different subset than those that use a logarithmic form; studies using years of schooling are separated from those specifying a set of dummy variables and so on and so forth. Many of the resulting subsets are therefore too small to permit estimation of a model $b = Zb + e$.

There is, however, a way to pool small subsets in meaningful ways. Namely, t-statistics of *b* reflect the sign and significance of the estimated relationship, where it does not matter so much whether the dependent variable is measured in linear or logarithmic form. Better yet, one may pool across all forms of performance measures, as long as the parameter estimates and t-statistics are recorded in such a way that the hypothesized effect of education points in the same direction. To this end we proceed in two steps: we first recode all the effects of schooling for performance measures for which "the more, the better" does not hold, i.e., exit from self-employment and the hazard out of self-employment. We then define a recoded variable t^* that takes on the value 0 for observations that have established a significantly negative effect, 1 for all studies that find an insignificantly negative effect, 2 for studies that show an insignificantly positive effect, and 3 for significantly positive effects. This ordered variable is then regressed on characteristics of the studies by means of an ordered probit model: $t^* = Zb + e$. The advantage of this approach is that it allows us to merge different entrepreneurial performance indicators into a single analysis.

As a result, we shall employ three subsamples of performance studies: (*i*) studies measuring performance as log(income); here we can use a quantitative approach, i.e., estimate the model $b = Zb + e$; (*ii*) studies measuring performance in terms of any sort of income; and (*iii*) pooling all performance studies, measured in any manner. The variation in the effect of education on performance in the latter two subsamples can be estimated qualitatively only, by means of the ordered probit model $t^* = Zb + e$.

In building suitable subsets, we need to address one statistical requirement that is to be met in any regression analysis: independence of observations. Thus, at issue is whether studies represent independent measurements of the impact of education. A detailed examination of the studies reveals the potential for a violation of the independence assumption. In many cases, multiple observations of the same type (for

instance, performance) come from one publication that uses a single dataset. For example, many publications typically report several estimated models in an effort to demonstrate the robustness of the reported results. To preserve independence, we therefore select only one of the estimated models, in particular the one that best shows the estimated relationship. But if one publication presents separate estimates for, e.g., men and women, we include both in our database, since independence is not in danger.[15] If it happens that several studies use the same data source and (roughly) the same subsample to examine the same entrepreneurial outcome, we must retain only one study to preserve independence. In our case, this requirement forced us to drop one observation (Lanjouw, 1999).[16,17]

Determinants in meta-analytical models

Since our subset sizes are not particularly large, to use an understatement, we must be careful to formulate parsimonious meta-analytical models. There is always a temptation to create rich specifications and many "what if" scenarios, but subset sizes of at most 70 studies do not permit rich models, and the effects we do estimate should be cautiously interpreted.

Actually, theory provides limited guidance in generating hypotheses about the determinants of the returns to schooling for entrepreneurs, or about the determinants of the relationship between entry and schooling. If anything, theory would only loosely guide us to ask the following questions. Given that schooling supposedly is more beneficial in more vibrant economies, is the return to schooling or the effect of schooling on entry, for example, higher in the East Asia than in Africa? Is it higher in urban areas than in rural areas? Is the return to schooling increasing over time? Moreover, theory also predicts that part of the returns to education derive from optimal choices concerning enterprise inputs and sector of economic activity. Thus, studies that control for input use and/or business sector should be reporting smaller estimated returns. This may also bear on the effect of using a structural model, although the impact of education on selection into entrepreneurship is not clear a priori, as argued in Section 2. Thus, the elimination of self-selection bias by estimating a structural model instead of a reduced form model may well affect the relationship between schooling and performance (or between schooling and entry) but the direction is not predicted by theory.

[15] We do acknowledge that two estimates drawn for different subsamples from one study (or from two studies by the same author) inherit a scientific approach from a single source and might therefore still be correlated, statistically speaking, from the perspective of a meta analysis. Our use of the term "independence of observations" pertains to the statistical independence of the samples that generate the estimated education effects.

[16] In a few other cases there is some overlap in the samples, such as one study that uses several rounds of a survey of which another study uses only one year. The gain in independence among observations was judged to be less than the loss of an observation, so no studies were dropped in such circumstances.

[17] For different reasons, four studies are eliminated at this stage as well. Vijverberg (1995) and Henderson (1983) reported the effect of education interacted with some other variable; Honig (1996, 1998) reported t-statistics that were so high as to be implausible in the light of the recorded R-squared statistic.

In other aspects, the meta-analysis is more exploratory. Are returns higher for men than for women? Is there any distinction in the effect of education across schooling levels? Moreover, does the performance measure selected affect the estimated return to schooling? Do estimates vary by sample size, or by the scientific weight of the journal (measured by its so-called impact factor), or by the field of the journal in which the study is published? Is there something like a publication or reporting bias in the sense that there is an overrepresentation of significant results? With this in mind, we now turn to the description of the explanatory variables of the meta-analytical models.

SAMPLE CHARACTERISTICS OF OBSERVATIONS The first group of control variables is meant to control for three sample characteristics: the region of the world that the sample is taken from, the percentage of females in the sample, and the percentage of sample individuals living in urban areas.[18]

USE OF SECTOR AND INPUT CONTROL VARIABLES As hypothesized above, the estimated impact of education may depend on whether the study includes controls for sector of business (or when the sample was comprised of enterprises in a single industry) and whether the regression model incorporates enterprise inputs.

MACROECONOMIC CONDITIONS A variable is included that indicates the (earliest) year from which the observations in the sample have been drawn. This variable captures temporal effects associated with technology, level of development, and such. Other variables that we examine comprise the sectoral composition of the economy (agriculture, industry, services); income per capita as a more direct measure of development; gross investment, as human capital could be either a substitute or a complement of physical capital; and rates of illiteracy, since human capital may be more precious in an environment where people are less skilled. These variables are drawn from World Bank (2003). We would have liked to incorporate in our analysis information about the competitive business climate as reported by the World Competiveness Scoreboard outcome (IMD Yearbook 2002). However, this database included only 19 of our 50 countries, too few to permit a meaningful analysis of this macroeconomic dimension.

CHARACTERISTICS OF JOURNALS Forty-two of the 240 observations are drawn from working papers or book chapters. Where feasible, we include a dummy variable indicating whether the observation was found in a published source rather than a working paper or book. In preliminary work, we found no effect of the branch of journal in which a study is published. This factor is therefore omitted. We also consider

[18] Sometimes, when the sample consisted of a mix of rural and urban residents and when the study did not supply the relevant statistics, we inserted the general ratio applying to the working population in the relevant country as taken from World Bank statistics. Moreover, when the percentage of females was not reported, we inserted a value of 0.50 when the unit of observation was the household (i.e., when the research question was whether the household operates an enterprise), or we inserted a value of 0.2 when there was a sense that the majority of the respondents would be male, or we use World Bank statistics on the female labor force participation rates.

the impact factor of the journal in the equations because if the impact factor stands for journal quality, then the study that is included in the better quality journal should itself be of a better quality. If all better quality studies report higher (or lower) schooling effects, then we have reason to believe that the "true" effect is indeed higher (or lower) than what a simple average effect suggests. For journals without an impact factor and for working papers and books, the impact factor is set to 0.

PUBLICATION/REPORTING BIAS Like Ashenfelter et al. (1999), we are concerned that the observed universe of published results has emerged solely because they were statistically significant. It might be the case that studies remain unpublished because they failed to provide a statistically significant rejection of the null-hypothesis of no effect. It is also quite possible that authors systematically drop insignificant determinants from their models, and thus that a statistically insignificant education effect leads them to omit all education variables: if so, such studies would not appear in the database either.

In order to test for this so called publication or reporting bias of the schooling measures, we include the standard error of the parameter estimate for education in the observed studies. Without publication bias, standard errors shouldn't have any significant relationship with the coefficients. In the presence of publication bias however, we would expect a positive relationship between the standard error and the coefficient *b* of the schooling measure.[19]

In the ordered probit models of t^*, the standard error cannot be used as a control variable: the dependent variables in these subsets are based on the t-values, which are partly determined by these standard errors themselves. Therefore, in these models, we include the square root of the sample size *N*, since the standard error of the parameter estimate declines with the sample size at a rate of $N^{0.5}$. Consider the effect of $N^{0.5}$: (*i*) if the true effect of the education variable is positive, its t-statistic is more likely significantly positive in a larger sample, and therefore the parameter on $N^{0.5}$ in the ordered probit model should be positive (and along the same lines, if the true effect is negative, the impact of $N^{0.5}$ will be negative); (*ii*) if the true effect of the education variable is 0, its t-statistic should hover around 0, no matter what the sample size is, so that the parameter on $N^{0.5}$ should be near 0; (*iii*) if publication bias exists, studies with small data samples are also reporting statistically significant education effects, so that, regardless of whether the true education effect is positive or negative, the parameter on $N^{0.5}$ should be near 0. The latter situation is therefore indicative of publication bias. Yet, caution applies: if the true effect of education varies such that it is positive in some places (samples) and negative in others, the parameter on $N^{0.5}$ could be near zero even in the absence of publication bias.

PERFORMANCE MEASURES USED In meta-analytical regression models that are estimated on subsets that

[19] Hedges (1992) offers another approach to the study of publication bias. We explore his method briefly to confirm our own findings below.

pool all performance measures, we control for the types of performance measures used to assess whether the impact of schooling differs between these measures. After some preliminary attempts, we include a dummy for earnings-related performance measures in these equations.

ESTIMATION METHODS AND TYPES OF DATA Finally, we include a dummy variable for whether a structural model has been used. In the few studies where panel data were used, the panel aspect appears to be largely ignored, so this distinction is not further explored.

6. The effect of schooling on entrepreneurial outcomes: performance

Having taken stock of the characteristics of the many studies, we are now in a position to examine the relationship between entrepreneurial outcomes and education, which is the centerpiece of our analysis. This section takes aim at the link between education and performance; in the next, we turn to entrepreneurial choice.

Effect of schooling

The first step is to illustrate whether the various estimated coefficients indeed imply that there exists a relationship between schooling and entrepreneurship performance. Accordingly, Table 3 records the sign and statistical significance of the impact of education on performance in all its forms. The preponderance of evidence is that the effect is positive, though it is not always easily teased out of the data. Perhaps the most successful specification is one that uses years entered linearly. Thirty-three of 40 observations are positive, with 19 of them statistically significantly so. Attempts to uncover nonlinearity are not particularly successful, unless the evidence is interpreted as indicating that upper secondary schooling yields greater returns than primary or lower secondary education. But this is not convincing as yet. Entering schooling in a quadratic form has only yielded insignificant parameter estimates (15 studies).[20]

Studies that use dummy variables generate the same kind of evidence. It should be noted that the database records the dummy variable effects in comparison with the base category of no schooling, even if a study actually utilized another group as the base category.[21] Upper elementary schooling shows clearer evidence of positive returns compared to lower elementary schooling, supporting the threshold notion that was found in agriculture (Phillips, 1994). However, the impact of lower secondary schooling is not more positive than that of upper elementary schooling, and indeed for post-secondary and college education the impact is not always different from the no-schooling category. Yet, six of seven estimates

[20] Note that with years entered in quadratic form, the overall effect of education is not recorded in Table 3, nor is the effect of years per se, which is not interpretable without reference to the squared term.

[21] This is not a fully innocuous choice, even if most studies use no schooling as the comparison group. People who have no education may be relatively heterogeneous in that some of them may not have had access to schooling in their youth and therefore have more native ability on average than those who choose to forego schooling.

for upper secondary schooling are significantly positive, which is consistent with the summary of the years effect above. In all, therefore, the relationship appears to be positive but not strongly evident.

A small number of studies incorporate information about training and apprenticeships in the regression model. This is done both by counting years or by dichotomous indicator variables. Clearly, both training and apprenticeships are informal, heterogeneous means of accumulating human capital and are difficult to capture precisely in a questionnaire. Across studies, mild evidence of a positive relationship emerges.

-Insert Table 3 about here-

It turns out that the most common model utilizes the log of earnings, enterprise income or profits as the dependent variable.[22] On top of this, many of these studies use years of education as the measure of education: in such a model, *b* represents the proportional increase in income resulting from a marginal year of schooling. We have pooled the estimates of all of these studies. In cases where a splined (piecewise linear) relationship was estimated, we treat the slope of each segment of the education-income function as a separate observation: in effect, such studies yield several estimates of *b* pertaining to different levels of education. Since it is impossible for us to characterize the correlation among these estimates from the information provided in the studies, we assume that the slopes of the segments are uncorrelated. Moreover, quadratic and interacted relationships are evaluated at means (or, if means were not provided, at six years of schooling and 35 years of age).[23] In all, then, we obtain a subsample of 49 observations. As shown in Table 4, the average value of *b* is 5.5 percent, indeed a very plausible rate of return per year of schooling. The standard deviation of 6.4 percent indicates a large spread around this mean across the 49 studies.

-Insert Table 4 –

Explaining the variation in the effect

To find out the cause of this spread, we estimate a number of regression models that are all variations of a base model that includes the proportion of females and the proportion of urbanites in the sample that the

[22] The use of the logarithm, together with an assumption that hours of work are predetermined, allows one to pool such studies regardless of the time dimension of the income concept. Estimation results from linear and log-linear models might also be made comparable by expressing the education effect in elasticity form or by standardizing the parameter estimates with the standard deviation of earnings, but that requires descriptive statistics on earnings and education values that are often not reported in the studies.

[23] t-statistics were adjusted similarly, though that must be done informally since not all of the necessary information is available. For example, in a quadratic specification, if the linear and quadratic terms are both positive and the linear term tends toward statistical significance, a simple linear model most often yields a significant positive parameter estimate. But if the linear term is positive and the squared term has a negative parameter such that a full inverted U-shape results, a simple linear model yields an insignificant coefficient. One could compute the exact t-statistic only if the covariance were reported, but that is never the case.

study uses. Owing to the small sample size, we are unable to include all variables at the same time; thus, we expand the base model with variables one (or one group) at the time.

The bottom panel of Table 4 shows our findings. According to the base model, returns are higher for females and in urban areas, by about four percentage points each. The gender effect is also observed in developed countries (Van der Sluis et al., 2003). Among the many additional factors we consider (groups A through M), the main finding is that b tends to be higher in studies that report a less precise estimate, which is consistent with the notion of a publication bias, but the effect of the standard error of b is significant at a 10 percent level only. We explore this in greater detail with the method designed by Hedges (1992). We find that the odds that a study with a statistically insignificant parameter estimate appears in the literature is only 0.65 (not reported in Table 4), but it is not statistically different from 1.0 at a p-value of 0.21.[24] Again, there is only weak evidence of publication bias here. Thus, it appears that the entrepreneurship literature is more tolerant of insignificant estimates than the parallel literature on the returns to education in wage earnings (Ashenfelter et al. 1998), perhaps because *a priori* expectations among researchers and journal editors are not as strong.

Other explanations of the variation in b are weak. Estimated returns may be lower when the regression model includes sector dummies, conforming to expectations. Over time, returns may be dropping. Returns in South-East Asia might be a bit lower than elsewhere. There is a hint that returns are lower in countries with a higher concentration of industry, and higher when more of GDP derives from the service sector. Returns may be lower in higher income countries. Human and physical capital appear to be substitutes, as the return diminishes in countries with greater rates of gross investment. However, none of these effects attain statistical significance.

In the next column of Table 4, we expand the subset to include all studies that examine enterprise income/earnings/profits, whether in linear or in logarithmic form, and we examine the sign and significance of the estimated relationship by means of an ordered probit model of t^*. Once again, when the proportion of females and urbanites in the sample is higher, it is more likely that the study finds a positive and significant education effect. The base model also contains the square root of the sample size: consistent with the hypothesis that education raises income and unsupportive of publication bias, the effect of $N^{0.5}$ is positive and significant.[25] Among the other determinants we examined, there is weak evidence that part of the effect of education plays through the allocative input choices. Moreover, estimates in a structural model tend to be more positive and significant.

[24] This pertains to the version of Hedges' model that merely estimates the average rate of return. This model suggests an average return of 4.4 percent and a variability across studies of 3.5 percentage points. In our "base model" the unexplained variability across studies already declines to 2.9 percent. Moreover, the estimate of the odds rises to 0.69 with a p-value of 0.30, reinforcing our conclusion that the evidence of publication bias is weak.

[25] The Hedges method cannot be applied to the ordered probit model of t^*.

The final column of Table 4 utilizes all studies of enterprise performance that specify education in years. Compared to the second column, the main added finding is that enterprise earnings models may well find a higher proportion of significant positive education effects than, e.g., studies of enterprise survival or technological efficiency. This phenomenon is also observed in developed countries (Van der Sluis et al., 2003).

Table 5 repeats this entire exercise while using the subset of studies that specify education through a series of dummy variables. The columns refer to primary, secondary, and post-secondary schooling and each combine the various sublevels that authors have used. Note, once again, that the base category is no schooling. Thus, on average, primary education yields a 19 ($=e^{0.174}$-1) percent gain, which is comparable with the 5 percent annually found in Table 4, since the average person in this category has less than 6 years of schooling (as the category combines lower and upper elementary schooling). Entrepreneurs with secondary schooling earn 34 ($=e^{0.294}$-1) percent more than unschooled individuals, which on an annual basis is a bit lower than the result in Table 4. Those with post-secondary schooling gain 140 percent.

Each of these averages is accompanied by a large standard deviation: the variation among studies is large. Which factors are associated with this variation? The subsets (15, 20, and 11 studies) are actually too small to examine this with meta-analytical regressions or, for that matter, to apply Hedges' model; thus, we should take the regression estimates with a grain of salt. Returns to primary schooling appear lower in urban areas and in economies with larger service sectors. Secondary schooling returns are higher in more recent samples, in societies with more illiterate people, a larger agricultural sector, and a smaller industrial sector. But the strongest determinant is the effect of the standard error, suggesting a strong publication bias (consider also the *in*significance of $N^{0.5}$): large estimates with large standard errors are tolerated, but smaller estimates had better be more precise.[26]

Adding all other income studies augments the subsets substantially but, as mentioned, necessitates an ordered probit approach, taking away two degrees of freedom. Pooling all performance studies that use dummy schooling variables raises the subsets to reasonable numbers: 27, 43 and 25 studies respectively. For primary schooling the negative urban effect weakens. Larger (or more precisely estimated) education effects occur in societies with lower literacy rates and more extensive agricultural activity. As to secondary schooling, the effect appears smaller for females and perhaps larger in agricultural societies. Other factors show little correlation with b. Returns to postsecondary schooling appear lower in regressions that include controls for inputs and perhaps sector and are higher when a structural model is estimated or in illiterate agricultural societies. As before, the impact is also stronger on earnings than on other performance measures.

[26] No regressions were attempted to explain the variation in postsecondary returns.

In sum, there is some evidence of higher returns for women and in urban areas, yet at the same time also in agricultural societies where literacy rates are lower. Inserting controls for inputs removes the allocative portion of the gain that education generates. Adding sector dummies removes another choice-related portion of the returns. Publication bias is evident in secondary schooling dummies: this would suggest that researchers utilize the series of dummy variables only if, in particular, secondary schooling shows up positive and significant. When designing a model, one would like to leave open the possibility of a nonlinear education effect, but while quadratic and splined functions usually are ineffective in detecting nonlinearity, dummy variables may indeed be able to indicate a nonlinear relationship.

-Insert Table 5 about here-

7. The effect of schooling on entrepreneurial outcomes: entrepreneurship

The second dimension of the impact of schooling deals with the choice of entrepreneurship. The logical focus of an analysis of this choice is the behavior of individuals as they make their career decisions: who starts their own business; who decides to go into wage employment? As Table 2 already indicated, such studies of *entry* are relatively scarce, to a large degree due to data limitations and in part due to researchers' inattention. The majority of studies examine being, rather than becoming, self-employed. Because the database is small anyway, we aggregate the stock and entry studies, though we will briefly check if estimated effects differ between stock and entry studies.

The literature offers many different empirical models. Estimation methods vary from simple binomial to elaborate structural models.[27] Studies also use many different comparison categories (Table 1), as the choices to people living in developing countries are broad. Altogether, this makes a comparison across studies tedious.[28] To make the parameter estimates comparable, we express them in terms of the marginal impact on the probability of non-farm self-employment.

[27] More advanced studies recognize that non-agricultural self-employment is one alternative among several and that therefore a multinomial choice model is preferable. The econometric model of choice is the multinomial logit model in which the parameters are identified relative to a base category. A multinomial probit model avoids the independence of irrelevant alternatives (IIA) assumption but is more difficult to estimate. A nested multinomial logit model does not always produce plausible nesting structures.

[28] However, different studies use different base categories, for example farming, wage employment, or non-employment. This renders the estimated schooling parameter of the self-employment selection equation incomparable across studies. Given the possible categories j (= 1, ..., J, in general), studies report the estimates and t-statistics (or standard errors) of b_j and b_k, but not of $b_j - b_k$ for every combination of j and k. Of course, one may, and we actually do, compute estimated values of $b_j - b_k$, but studies do not provide enough information to conclude anything definite about the significance level of this difference. Since in this survey we desire to understand the impact of education on the self-employment choice in detail, we make some reasonable assumptions such that we are able to evaluate the (significance levels of the) impact of education on the choice between every combination of economic activities. For example, if b_j is statistically significantly positive and b_k is statistically

Table 6 describes the evidence from studies that analyze the impact of education on the self-employment choice, sorted according to the base category. The table shows only those studies for which at least one type of education variable has been used more than five times.[29] As it turns out, there is a great deal of consistency between studies that specify education in the form of years of schooling and studies that employ categorical variables. Overall, relative to a heterogeneous set of other forms of employment (panel A), education lowers the likelihood of non-farm self-employment. The effect is frequently statistically significant and amounts to an average of 1.3 percentage point decline per year of schooling. The contrast with wage employment (panel C) is much sharper negative at a 6.8 percentage point decline per year. Moreover, a rise in the attained schooling level pulls people out of farming (panel D) at a rate of 8.1 percent per year of schooling. Panel B combines alternative forms of employment with non-employment as the base and shows that relative to this amalgam education perhaps weakly favors non-farm self-employment. In combination, panels A and B suggest that more educated individuals are less likely to be non-employed than to be engaged in a non-farm enterprise, but this tends to be contradicted by the more-often-negative estimates in panel E that summarizes the explicit comparisons between non-employment and non-farm self-employment. Finally, panel F explicitly compares entry against non-entry, but the ten observations come from one study on transitions into entrepreneurship after communism in Hungary. This evidence shows an ambiguous effect of education. Altogether, schooling is associated with a distinct sorting in the labor market of developing countries.

-Insert Table 6 about here-

The wage/self-employment comparison has been studied most often, and the subsets of studies are large enough to permit a deeper meta-analysis, which is reported in Table 7. As before, we explore various hypotheses from a base model that includes the proportion of the sample that is female and that has an urban residence. The base model results suggest that, as a person's level of education rises, a woman is more likely to choose wage employment over non-farm entrepreneurship than a man. Similarly, a more educated urban resident is more likely to select a wage job than a more educated rural resident, though the estimated difference is not as strong as between the genders.

---

insignificantly different from 0 (or is statistically significantly negative), we assume that $b_j - b_k$ is statistically significantly positive. Or if both b_j and b_k are statistically significantly positive, we assume that $b_j - b_k$ is statistically insignificantly different from 0 (but of course still positive if $b_j - b_k > 0$).

[29] The category of primary education combines lower and upper levels, as does the category of secondary education. Post-secondary education includes college.

Adding variables one at the time helps explain the variation between studies.[30] The year in which the data were generated proves irrelevant: the global level of technology or globalization has no discernible impact on local labor market sorting patterns. In regard to the macroeconomic variables, the meta-analysis of the effect of years of schooling yields a distinctly different association pattern than the education category analyses. The years model suggests that *b* is more negative in agricultural societies with higher illiteracy: as educated workers are scarcer, education opens up more opportunities in wage employment. The category models show no such association. The years model finds that *b* is less negative when the economy is more industrial; the category models show the opposite. The category models also show that *b* becomes less negative in service-oriented economies; the association in the years model is inconclusive. Per capita income matters in the years model (reducing the negative size of *b*) but is largely absent in the category models. The only macroeconomic variable on which all models agree is gross investment, reducing the negative value of *b*, suggesting that in a faster growing economy there are more entrepreneurial opportunities for more educated individuals. Overall, the macroeconomic variables show a more plausible effect in the years model; some of the estimated effects in the category models are exceedingly large, possibly indicating the kind of spurious effects that one sometimes finds in regressions run on small samples.[31]

In sum, the literature shows that more educated workers typically end up in wage employment, shunning non-farm entrepreneurship. This effect is stronger for women, possibly stronger in urban areas, and also stronger in less developed economies where agriculture is more dominant and literacy rates are lower. Relative to farming, however, more educated workers seek out non-farm entrepreneurship opportunities.

-Insert Table 7 about here-

8. Conclusion, evaluation and suggestions for further research

Conclusion

This paper's objective has been to give an overview of the literature pertaining to developing countries on the relationship between entrepreneurship selection/performance and schooling. We summarize and examine variation among the existing studies by means of meta-analysis. We compare the findings that pertain to developing countries to the existing evidence on the returns to schooling for employees.

[30] In regressions that are not reported in the table, we examine the field of the journal publication, which did not matter, and the difference between the six regions of the world. The latter showed statistically significant parameters, but with such a small sample there is little economic significance one should attach to such a result.

[31] Whether the author formulated a structural model or examined entry rather than stock makes no difference. Publication bias is not evident. We do not include the standard error of *b* in the education category equations, since this standard error cannot be derived from the studies in the cases that the authors have utilized a different base category for their multinomial logit model than wage employment.

Moreover, Van der Sluis et al. (2003) provide the same sort of overview for industrial countries, and we compare our findings with their results.

We find that an additional year of schooling raises enterprise income by an average of 5.5 percent. This is somewhat lower than the return to education in wage employment in developing countries, where averages range from 7.2 percent to more than 11 percent (Psacharopoulos, 1994).[32] Yet, it is closely similar to the average return (6.1 percent) to schooling in entrepreneurial pursuits in the United States. It is worth noting that this estimate of 6.1 percent for entrepreneurs is smaller than the usual comparable estimates for employees that are between seven and nine percent. Thus, the comparison between entrepreneurs and employees runs parallel in industrial and developing countries. The return tends to be higher for women, like in developed countries, and for urban residents, yet at the same time also in agricultural societies where literacy rates are lower. As is to be expected, the measured return is sensitive to model specification: for example, inserting controls for inputs removes the allocative portion of the gain that education generates.

In regards to entrepreneurship choice, our descriptive summary of the effect of education indicates that more educated workers typically end up in wage employment, shunning non-farm entrepreneurship. Relative to farming, however, more educated workers seek out non-farm entrepreneurship opportunities. For reasons of sample size limitations, a meta-analysis to explain the heterogeneity of results is feasible only for the wage employment/non-farm self-employment comparison. The education effect that separates workers out of self-employment into wage employment is stronger for women, possibly stronger in urban areas, and also stronger in less developed economies where agriculture is more dominant and literacy rates are lower. Many studies report that uneducated women are concentrated in low-income sectors of food commerce and textiles. Thus, it appears that education leads women towards more rewarding opportunities to be found not in higher-income entrepreneurial activities but in wage jobs. The differential in the labor market sorting process that education brings about in developing countries stands in contrast with the lack of relationship between an individual's schooling level and the probability of selection into entrepreneurship that was reported in industrial countries. Economic theory points out that there are opposing effects that education might have on entrepreneurship entry, which may play out differently in developing countries: for example, the opportunity set in developing countries is larger, as are the income differentials between the sectors.

Evaluation of the "State of the Art" and suggestions for further research

[32] Psacharopoulos (1994:1330) also reported a 10.8 percent average rate of return to education in self-employment, which, just as we find, is less than the 12.2 percent average return in "dependent employment." This value of 10.8 percent is a summary of estimates in a small early entrepreneurship literature and is, given the large standard error in Table 6, not necessarily out of line with our average of 5 percent.

We conclude that the state of the art of research into the effect of education on entrepreneurship is somewhat disappointing. Though much research effort has been directed towards the issue, we have found many lacunae, issues that either have not been addressed at all, or have not been addressed in a satisfactory manner. We benchmark against the common practice in the returns to schooling in employment literature.

A first drawback is the lack of homogeneity in the definitions of schooling, performance, and entrepreneurship. Only about 35 percent of the studies use a simple "years of schooling" as the measure of educational attainment. Most researchers utilize a haphazardly designed set of dummies for specific levels of schooling in their performance and entry equations: the comparison group varies so much across studies that it is difficult to make generalizations. The same holds true for performance where various (by themselves useful) definitions of performance are in fashion, and for entrepreneurship selection where comparison groups are quite random. To build a body of knowledge, more attention must be paid to both the systematic operationalization of the entrepreneurship concepts and the reporting of the results (including, it must be said, a proper description of the data that are utilized). Novelty for the sake of product differentiation does not have as much value in the scientific realm as it does in the market place.

A second issue that should receive more attention in the literature is the role that ability and other often unobserved factors might play in determining entrepreneurial selection and performance. It is quite plausible that the "effect" of schooling that is typically estimated is not completely causal and therefore biased: ability and other factors might increase performance and also lead to more schooling, thus potentially leading to a spurious positive effect of schooling on performance. A deeper theoretical concern is that schooling itself is endogenous to one's performance in the labor market: although future earnings are not the only reason to pursue an education, the prospect of earning higher incomes entices many students to stay in school longer (e.g., Glewwe, 1999). In the established returns-to-schooling literature that focuses on wage employment, this issue is well recognized, and whenever the data permit, researchers attempt to correct for the ability bias and the endogeneity of schooling by including measures of innate ability into the specifications, by using instrumental variables of one type or another, or by running controlled experiments as is the case in "twins" studies. In the entrepreneurship counterpart of this literature, none of the studies made reference to the endogeneity of schooling,[33] and there are virtually no studies that incorporate any kind of ability measure.[34]

Of the 129 studies on performance that we surveyed, 19 percent corrected for selection biases and 81 percent did not. Omission of such a correction should be acknowledged in the type of recommendations

[33] This kind of analysis is further complicated by the fact that, because of cultural differences or financial constraints or the availability of schools, people with zero years of schooling may be a heterogeneous group.

[34] Exceptions are Escher et al. (2002), where cognitive ability mattered but educational attainment was not controlled for, and Vijverberg (1999), where the effect of various ability measures was statistically insignificant.

these studies put forth. Actually, one might question the standard model that many researchers use to correct of selectivity. The multinomial logit model assumes that each person makes one and only one labor market choice. However, it has become quite clear that many of the labor market participants, including the non-farm self-employed, are active in more than one sector at the time or, even more so, during the course of a year (Vijverberg, 1992; Lanjouw, 2001; Vijverberg and Haughton, 2002). The MNL model is not suited to analyze such behavioral patterns (nor, by extension, are the logit and probit model), and the selectivity correction may therefore not be appropriate either.[35,36]

For employees, the distinction between (the effects of) general education and specific education is quite well known. In the entrepreneurship counterpart of the literature, there is much that remains to be explored regarding the type of curriculum, the effect of training, and the benefits of apprenticeships. It may well be true that the type of curriculum is more important than the level of schooling pursued, and it is conceivable that both curriculum and level of schooling impact entrepreneurial outcomes through productivity (human capital theory) and sorting (screening or signaling theory; see Wolpin, 1977, and Riley, 2002). Related, while we found a limited number of studies that included prior entrepreneurship experience of the person him-/herself or of the father of the respondent as a determinant of the likelihood that the person pursues entrepreneurial interests, the precise process that leads to this choice is not clear and likely constitutes a fruitful area of future research.

Finally, entrepreneurship might be described as the process of bringing inputs, technologies, and output markets together. An important part of this process is acquiring financial capital during start-up and expansion. One might view obtaining credit as one of the many dimensions of entrepreneurial performance, and indeed there is little doubt that schooling is related to the likelihood of getting loans (e.g., Raturi and Swamy, 1999; Bigsten et al., 2000; McKernan, 2002, Parker and Van Praag, 2003). Yet, while there is therefore a close link with this branch of the finance literature, a review of the precise role of education is left as a future task. This paper addresses the overall effect of education on performance and selection into entrepreneurship rather than the component effects.

To summarize, there are many challenges left in the study of the relationship between entrepreneurship and education, not only quantitatively but also qualitatively. We hope that the present paper will stimulate efforts towards a deeper, more robust understanding of the role of education in determining the decision to become an entrepreneur and in determining the returns to education among entrepreneurs. If this meta-analysis demonstrates anything, it is that in developing countries the selection of entrepreneurs seems to

[35] One attempt to develop a solution to this issue is found in Vijverberg (1986).

[36] A related issue that has never been discussed is the potential bias resulting from non-representative participation in samples. It might well be the case that more successful entrepreneurs do not take the time to fill out questionnaires, or, conversely, that poorly performing entrepreneurs are unwilling to reveal their bad state of affairs in a questionnaire.

be rising at low levels of formal education and falling at higher levels, and that performance has a positive relationship with education pursued. If schooling is so positively related "even" to entrepreneurship performance, this makes a stronger case for investment in human capital through schooling (perhaps including lifelong learning) at all levels.

References

Note that studies that contribute to the database are listed separately below. Studies that are referred to in the text but not below may be found in the database reference list.

Ashenfelter, Orly, Colm Harmon and Hessel Oosterbeek, 1999. "A review of the schooling / earnings relationship with tests for publication bias." *Labour Economics* 6, 453-470.

Bennell, Paul, 1996. "Rates of return to education: Does the conventional pattern prevail in sub-Saharan Africa?" *World Development* 24:1, 183-199.

Bigsten, Arne, Paul Collier, Stefan Dercon, Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Mans Soderborn, Abena Oduro, Remco Oostendorp, Cathy Patillo, Francis Teal, and Albert Zeufack. 2000. "Credit Constraints in Manufacturing Enterprises in Africa." Discussion Paper 2000.24, Institute for the Study of African Economies, Oxford University.

Calvo, G. and S. Wellisz. 1980. "Technology, entrepreneurs and firm size", *Quarterly Journal of Economics* 95, 663-678.

Card, D. and A.B. Krueger. 1995. "Time-series Minimum-Wage Studies - A Meta-analyis." *American Economic Review* 85:2, 238-243.

Escher, Susanne, Rafal Grabarkiewicz, Michael Frese, Gwenda van Steekelenburg, Maartje Lauw, and Christian Friedrich. 2002. "The Moderator Effect of Cognitive Ability on the Relationship between Planning Strategies and Business Success of Small Scale Business Owners in South Africa: A Longitudinal Study." *Journal of Developmental Entrepreneurship*, 7:3, 305-318.

Glewwe, Paul. 1999. "A method for estimating the determinants of schooling outcomes." In P. Glewwe, ed. *The Economics of School Quality in Developing Countries: An Empirical Study of Ghana*. New York: St. Martin's Press.

Haggblade, Steven, Peter Hazell, and James Brown. 1989. "Farm-nonfarm linkages in rural Sub-Saharan Africa." *World Development*, 17:8, 1173-1201.

Haggblade, Steven, Peter Hazell and Thomas Reardon. 2002. *Strategies for stimulating poverty-alleviating growth in the rural nonfarm economy in developing countries*. Washington, D.C. The World Bank. Rural Development Department; International Food Policy Research Institute (IFPRI) Discussion paper, http://www.ifpri.org/divs/eptd/dp/eptdp92.htm

Harris, J.R., and M. Todaro. 1969. "Migration, unemployment and development: a two-sector model." *American Economic Review*, 60:1, 126-142.

Hedges, L.V. 1992. "Modeling publication effects in meta-analysis." *Statistical Science*, 7:2, 246-255.

House, William J. 1984. "Nairobi's informal sector -- dynamic entrepreneurs or surplus labor?" *Economic Development and Cultural Change*, 32:2, 277-302.

International Labour Office. 1972. *Employment, Incomes, and Equality: A Strategy for Productive Employment in Kenya.* Geneva: ILO.

I.M.D. World Competitiveness Yearbook 2002. www.imd.ch

Jamison, Dean T., and Lawrence J. Lau. 1982. *Farmer Education and Farm Efficiency*. Baltimore: Johns Hopkins University Press

Kiggundu, Moses N. 2002. "Entrepreneurs and entrepreneurship in Africa: What is known and what needs to be done." *Journal of Developmental Entrepreneurship*, 7:3, 239-258.

King, K., and S. McGrath. 1999, *Enterprise in Africa: Between Poverty and Growth.* London: Intermediate Technology.

Kurian, George T. 1988. *The World Education Encyclopedia*. Oxford: Facts on File.

Lanjouw, Jean O., and Peter Lanjouw. 2001. "The rural non-farm sector: issues and evidence from developing countries." *Agricultural Economics*, 26:1, 1-23.
Le, A.T. 1999. "Empirical studies of self-employment." *Journal of Economic Surveys* 13:4, 381-416.
Lee, Jong-Hwa, and Robert J. Barro. 2001. "Schooling quality in a cross-section of countries." *Economica*, 68:272, 465-488.
Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." *The Manchester School*. 139-191.
Lockheed, Marleen E., Dean T. Jamison, and Lawrence J. Lau. 1987. "Farmer Education and Farm Efficiency: A Survey." *Economic Development and Cultural Change*, 29, 37-76.
Lucas, R.E. 1978. "On the size distribution of business firms." *Bell Journal of Economics* 9:2, 508-523.
McKernan, Signe-Mary. 2002. "The Impact of Microcredit Programs on Self-Employment Profits: Do Noncredit Program Aspects Matter." *Review of Economics and Statistics*, 84:1, 93-115.
Mead, Donald C. 1994. "The contribution of small enterprises to employment growth in Southern and Eastern Africa." World Development, 22:12, 1881-1894.
Mead, Donald C., and Carl Liedholm. 1998. "The dynamics of micro and small enterprises in developing countries." *World Development*, 26:1, 61-74.
Mead, Donald C., and Catherine Morrison. 1996. "The informal sector elephant." *World Development*, 24:10, 1611-1619.
Nijkamp, Peter and Jacques Poot. 2002. "Meta-analysis of the impact of fiscal policies on long-run growth", Tinbergen Institute Discussion Paper 28/3
Parker, Simon C., and C. Mirjam Van Praag. 2003. "Explaining entrepreneurial performance: The effects of education and financial capital constraints." University of Amsterdam.
Peattie, Lisa. 1987. "An idea in good currency and how it grew: the informal sector." *World Development*, 15:7, 850-860.
Phillips, J.M. 1994. "Farmer education and farm efficiency." *Economic Development and Change*, 43:1, 149-165.
Psacharopoulos, George, 1994. "Returns to investment in education: a global update." *World Development* 22, 1325-1343.
Ranis, Gustav, and John C.H. Fei. 1964. *Development of the labor surplus economy: theory and policy*. Homewood, Ill.: R.D. Irwin.
Raturi, Mayank, and Anand V. Swamy. 1999. "Explaining Ethnic Differentials in Credit Market Outcomes in Zimbabwe." *Economic Development and Cultural Change*, 47:3, 585-604.
Reardon, Thomas, J. Edward Taylor, Kostas Stamoulis, Peter Lanjouw, and Arsenio Balisacan. 2000. "Effects of non-farm employment on rural income inequality in developing countries: an investment perspective." *Journal of Agricultural Economics*, 51:2, 266-288.
Riley, J.G. 2002. "Weak and strong signals." *Scandinavian Journal of Economics*, 104:2, 213-36.
Roy, A.D. 1951. "Some thoughts on the distribution of earnings." *Oxford Economic Papers*, 3, 135-146.
Schultz, T.W. 1980. "Investment in Entrepreneurial Activity." *Scandinavian Journal of Economics*, 82:4, 437-448
Spence, A.M. 1973. *Market Signaling: Information Transfer in Hiring and Related Processes*, Cambridge, MA: Harvard University Press.
Van Der Sluis, J., C.M. Van Praag and W. Vijverberg. 2003. "Entrepreneurship Selection and Performance: A Meta-analysis of the Impact of Education in Industrialized Countries." Working Paper.
Van Praag, C. M. and J. S. Cramer. 2001. "The roots of entrepreneurship and labour demand: Individual ability and low risk aversion." *Economica*, 68:269, 45-62.
Vijverberg, Wim P.M. 1992. "Measuring income from family enterprises with household surveys." *Small Business Economics*, 4, 287-305.
Wolpin, K. 1977. "Education and screening." *American Economic Review*, Vol. 67, No. 5, 949-958.
World Bank. 2003. *World Develpoment Indicators*, http://www.worldbank.org/data/dataquery.html, accessed on 21 August 2003.

Studies included in the database.

Alvarez, Roberto, and Gustavo Crespi. 2003. "Determinants of technical efficiency in small firms." *Small Business Economics*, 20, 233-244

Appleton, Simon, Arne Bigsten, and Damiano Kulundu Manda. 1999. "Educational expansion and economic decline: returns to education in Kenya, 1978-1995." Working Paper WPS/99-6, Centre for the Study of African Economies, University of Oxford.

Barr, Abigail M. 1995. "The missing factor: entrepreneurial networks, enterprises and economic growth in Ghana." WPS 95-11, Centre for the Study of African Economies, Oxford University.

Barr, Abigail M., and Truman Packard. 2002. "Revealed preference and self insurance: can we learn from the self-employed in Chile?" World Bank Discussion Paper 3365, The World Bank.

Blau, David. 1985. "Self-employment and self-selection in developing country labor markets." *Southern Economic Journal*, 52:2, 351-363,

Blau, David. 1986. "Self-employment, earnings, and mobility in Peninsular Malaysia." *World Development*, 14:7, 839-854.

Burki, Abid A. 1998. "Measuring production efficiency of small firms in Pakistan." *World Development*, 26:1, 155-169.

Chaves, Rodrigo A., Susana Sanchez, Saul Schor, and Emil Tesliuc. 2001. "Financial markets, credit constraints, and investment in rural Romania." World Bank Technical Paper 499, The World Bank.

Christofides, Louis N., and Panos Pashardes. 2002. "Self/paid-employment, public/private sector selection, and wage differentials." *Labour Economics*, 9, 737-762.

Co, Catherine Y., Ira N. Gang, and Myeong-Su Yun. 2002. "Self-employment and wage earning in Hungary." Discussion Paper, Rutgers University.

Cohen, Barney, and William House. 1996. "Labor market choices, earnings, and informal networks in Khartoum, Sudan." *Economic Development and Cultural Change*, 589-618.

Copestake, James and Sonia Bhalotra and Susan Johnson. 2001. "Assessing the impact of microcredit: a Zambian case study." *Journal of Development Studies*, 37:4, 81-100.

Corral, Leonardo. 2001. "Rural nonfarm incomes in Nicaragua." *World Development*, 29:3, 427-442.

Cortes, Mariluz, Albert Berry, Ashfaq Ishaq. 1987. *Success in Small and Medium-Scale Enterprises*. New York: Oxford University Press.

Coulombe, Harold, and Andrew McKay. 1996. "Modeling determinants of poverty in Mauritania." *World Development*, 24:6, 1015-1031.

Cunningham, Wendy V., and William F. Maloney. 2001. "Heterogeneity among Mexico's micro-enterprises: an application of factor and cluster analysis." *Economic Development and Cultural Change*, 131-156.

Daniels, Lisa, and Donald C. Mead. 1998. "The contribution of small enterprises to household and national income in Kenya." *Economic Development and Cultural Change*, 45-71.

Dasgupta, Sukti. 2003. "Structural and behavioural characteristics of informal service employment: evidence from a survey in New Delhi." *Journal of Development Studies*, 39, 351-80.

De Janvry, Alain and Elisabeth Sadoulet. 2000. "Rural poverty in Latin America: determinants and exit paths." *Food Policy*, 25, 389-409.

Earle, John S., and Zuzana Sakova. 1999. "Entrepreneurship from Scratch: Lessons on the Entry Decision into Self-Employment from Transition Economies." Discussion Paper 79, IZA.

Earle, John S., and Zuzana Sakova. 2000. "Business start-ups or disguised unemployment? Evidence on the character of self-employment from transition economies." *Labour Economics*, 7, 575-601.

Elbers, Chris, and Peter Lanjouw. 2001. "Intersectoral transfer, growth, and inequality in rural Ecuador." *World Development*, 29:3, 481-496.

Escobal, Javier. 2001. "The determinants of nonfarm income diversification in rural Peru." *World Development*, 29:3, 497-508.

Funkhouser, Edward. 1992. "Migration from Nicaragua: some recent evidence." *World Development*, 20:8, 1209-1218.
Funkhouser, Edward. 1997. "Labor market adjustment to political conflict: changes in the labor market in El Salvador during the 1980s." *Journal of Development Economics*, 52, 31-64.
Giugale, M, S. El-Diwany and S. Everhart. 2000. "Informality, size and regulation: theory and an application to Egypt." *Small Business Economics*, 14, 95-106.
Glick, Peter, and David E. Sahn. 1997. "Gender and education impacts on employment and earnings in West Africa: evidence from Guinea." *Economic Development and Cultural Change*, 793-823.
Glick, Peter, and David E. Sahn. 1998. "Health and productivity in a heterogeneous urban labour market." *Applied Economics*, 30, 203-216.
Goedhuys, Micheline and Leo Sleuwaegen. 2000. "Entrepreneurship and growth of entrepreneurial firms in Cote d'Ivoire." *Journal of Development Studies*, 36:3, 122-145.
Gokcekus, Omer and Kwabena Anyane-Ntow and T. T. Richmond. 2001. "Human capital and efficiency: the role of education and experience in micro-enterprises of Ghana's wood-products industry." *Journal of Economic Development*, 26:1, 103-113.
Henderson, James. 1983. "Earnings functions for the self-employed: Comment." *Journal of Development Economics*, 13, 97-102.
Honig, Benson. 1996. "Education and self-employment in Jamaica." *Comparative Education Review*, 40:2, 177-193.
Honig, Benson. 1998. "What determines success? Examining the human, financial, and social capital of Jamaican microentrepreneurs." *Journal of Business Venturing*, 13, 371-394.
Honig, Benson. 2001. "Human capital and structural upheaval: a study of manufacturing firms in the West Bank." *Journal of Business Venturing*, 16, 575-594.
Khandker, Shahidur. 1987. "Labor market participation of married women in Bangladesh." *Review of Economics and Statistics*, 69:3, 536-541.
Lanjouw, Peter. 1999. "Rural nonagricultural employment and poverty in Ecuador." *Economic Development and Cultural Change*, 48:1, 91-122.
Lanjouw, Peter. 2001. "Nonfarm employment and poverty in rural El Salvador." *World Development*, 29:3, 529-547.
Lanjouw, Peter, Jaime Quizon, and Robert Sparrow. 2001. "Non-agricultural earnings in peri-urban areas of Tanzania: evidence from household survey data." *Food Policy*, 26, 385-403.
Lanot, Gauthier, and Christophe Muller. 1997. "Dualistic sector choice and female labour supply: evidence from formal and informal sectors in Cameroon." Working Paper WPS/97-9, Centre for the Study of African Economies, University of Oxford.
Lerner, Miri, and Sigal Haber. 2001. "Performance factors of small tourism ventures: the interface of tourism entrepreneurship and the environment." *Journal of Business Venturing*, 16, 77-100.
Lerner, Miri, Candida Brush, and Robert Hisrich. 1997. "Israeli women entrepreneurs: an examination of factors affecting performance." *Journal of Business Venturing*, 12, 315-339.
Little, Ian M.D., Dipak Mazumdar, John M. Page Jr. 1987. *Small Manufacturing Enterprises: A Comparative Analysis of India and Other Economies*. New York: Oxford University Press.
Magnac, T. 1991. "Segmented or competitive labor markets." *Econometrica*, 59:1, 165-187.
Maloney, William F. 1999. "Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico." *World Bank Economic Review*, 13:2, 275-302.
Martin, John P., and John M. Page Jr. 1983. "The impact of subsidies on X-Efficiency in LDC industry: theory and an empirical test." *Review of Economics and Statistics*, 65:4, 608-617.
McPherson, Michael A. 1996. "Growth of micro and small enterprises in southern Africa." *Journal of Development Economics*, 48, 253-277.
Melmed-Sanjak, Jolyne, and Carlos E. Santiago. 1996. "The household and employment in small-scale nonfarm enterprises." *World Development*, 24:4, 749-764.
Mengistae, Taye. 2001. "Indigenous ethnicity and entrepreneurial success in Africa: some evidence from Ethiopia." Policy Research Working Paper 2534, The World Bank.

Mesnard, Alice, and Martin Ravallion. 2001 "Is inequality bad for business? A nonlinear microeconomic model of wealth effects on self-employment." Policy Research Working Paper 2527, The World Bank.

Mills, Bradford, and David E. Sahn. 1997. "Labor market segmentation and the implications for public sector retrenchment programs." *Journal of Comparative Economics*, 25, 485-402.

Moock, Peter, Philip Musgrove, and Morton Stelcner. 1989. "Education and earnings in Peru's informal nonfarm family enterprises." WPS 236, The World Bank.

Nafziger, Wayne and Dek Terrell. 1996. "Entrepreneurial human capital and the long-run survival of firms in India." *World Development*, 24:4, 689-696.

Nielsen, Helena Skyt and Niels Westergard-Nielsen. 2001. "Returns to schooling in less developed countries: new evidence from Zambia." *Economic Development and Cultural Change*, 49:2, 365-394.

Nziramasanga, Mudziviri and Minsoo Lee. 2001. "Duration of self-employment in developing countries: evidence from small enterprises in Zimbabwe." *Small Business Economics*, 17, 239-253.

Nziramasanga, Mudziviri and Minsoo Lee. 2002. "On the duration of self-employment: the impact of macroeconomic conditions." *Journal of Development Studies*, 39:1, 46-73.

Page, John M., Jr. 1984. "Firm size and technical efficiency: applications of production frontiers to Indian survey data." *Journal of Development Economics*, 16:1/2, 129-152.

Paulson, Anna, and Robert Townsend. 2001. "Entrepreneurship and financial constraints in Thailand." Working Paper, University of Chicago.

Pradhan, Menno, and Arthur van Soest. 1997. "Household labor supply in urban areas of Bolivia." *Review of Economics and Statistics*, 300-310.

Ramachandran, Vijaya, and Manju Kedia Shah. 1999. "Minority entrepreneurs and firm performance in Sub-Saharan Africa." *Journal of Development Studies*, 36:2, 71-87.

Robert, Peter, and Erzsebet Bukodi. 2000. "Who are the entrepreneurs and where do they come from? Transition to self-employment before, under and after communism in Hungary." *International Review of Sociology*, 10:1, 147-171.

Rona-Tas, Akos. 1994. "The first shall be last? Entrepreneurship and communist cadres in the transition from socialism." *American Journal of Sociology*, 100:1, 40-69.

Ruben, Ruerd, and Marrit van den Berg. 2001. "Nonfarm employment and poverty alleviation of rural farm households in Honduras." *World Development*, 29:3, 549-560.

Saavedra, Jaime and Alberto Chong. 1999. "Structural reform, institutions and earnings: evidence from the formal and informal sectors in urban Peru." *Journal of Development Studies*, 35:4, 95-116.

Shavit, Yossi, and Ephraim Yuchtman-Yaar. 2001. "Ethnicity, education, and other determinants of self-employment in Israel." *International Journal of Sociology*, 31:1, 59-91.

Singh, Surendra, Ruthie Reynolds, and Safdar Muhammad. 2001. "A gender-based performance analysis of micro and small enterprises in Java, Indonesia." *Journal of Small Business Management*, 39:2, 174-182.

Smith, Paula, and Michael R. Metzger. 1998. "The return to education: street vendors in Mexico." *World Development*, 26:2, 289-296.

Strassmann, W. Paul. 1987. "Home-based enterprises in cities in developing countries." *Economic Development and Cultural Change*, 36:1, 121-144.

Teilhet-Waldorf, Saral, and William H. Waldorf. 1983. "Earnings of self-employed in an informal sector: a case study of Bangkok." *Economic Development and Cultural Change*, 587-607.

Telles, Edward E. 1993. "Urban labor market segmentation and income in Brazil" *Economic Development and Cultural Change*, 231-247.

Tiefenthaler, Jill. 1994. "A multisector model of female labor force participation: empirical evidence from Cebu Island, Philippines." *Economic Development and Cultural Change*, 719-742.

Tran, Quoc Trung. 2000. *Determinants of Income from Rural Non-Farm Business Activities in Vietnam.* Master's Thesis, National Economics University, Hanoi/Institute for Social Studies, The Hague.

Verme, Paolo. 2000. "The choice of the working sector in transition: income and non-income determinants of sector participation in Kazakhstan." *Economics of Transition*, 8:3, 691-731.

Vijverberg, Wim. 1986. "Consistent estimates of the wage equation when individuals choose among income-earning activities." *Southern Economic Journal*, 52:4, 1028-1042.
Vijverberg, Wim. 1991. "Profits from self-employment: the case of Cote d'Ivoire." *World Development*, 19:6, 683-696.
Vijverberg, Wim. 1993. "Educational investments and returns for women and men in Cote d'Ivoire." *Journal of Human Resources*, 28:4, 933-974.
Vijverberg, Wim. 1995. "Returns to schooling in non-farm self-employment: an econometric case study of Ghana." *World Development*, 23:7, 1215-1227.
Vijverberg, Wim. 1998. "Nonfarm household enterprises in Vietnam." In David Dollar, P. Glewwe, and J. Litvack, eds., *Household Welfare and Vietnam's Transition*, Washington D.C.: The World Bank, 137-177.
Vijverberg, Wim. 1999. "The impact of schooling and cognitive skills on income from non-farm self-employment." In P. Glewwe, ed., *The Economics of School Quality Investments in Developing Countries*, London: MacMillan Press, 206-252.
Vijverberg, Wim, and Jonathan Haughton. 2002. "Household Enterprises in Vietnam: Survival, Growth, and Living Standards." Discussion Paper WPS 2773, The World Bank.
Winters, Paul, Benjamin Davis, and Leonardo Corral. 2002. "Assets, activities and income generation in rural Mexico: factoring in social and public capital." *Agricultural Economics*, 27, 139-156.
Woldenhanna, T., and A. Oskam. 2001. "Income diversification and entry barriers: evidence from the Tigray region of Northern Ethiopia." *Food Policy*, 26, 351-365.
Wu, Xiaogang. 2002. "Embracing the market: entry into self-employment in transitional China, 1978-1996." William Davidson Working Paper 512, University of Michigan.
Yang, Dennis Tao and Mark Yuying An. 2002. "Human capital, entrepreneurship, and farm household earnings." *Journal of Development Economics*, 68, 65-68.
Yunez-Naude, Antonio and Edward Taylor. 2001. "The determinants of non-farm activities and incomes of rural households in Mexico, with emphasis on education." *World Development*, 29:3, 561-572.
Zhao, Yaohui. 1999. "Labor migration and earnings differences: the case of rural China." *Economic Development and Cultural Change*, 767-782.

Figure 1: Distribution of Sample Sizes across Studies, by Study Type

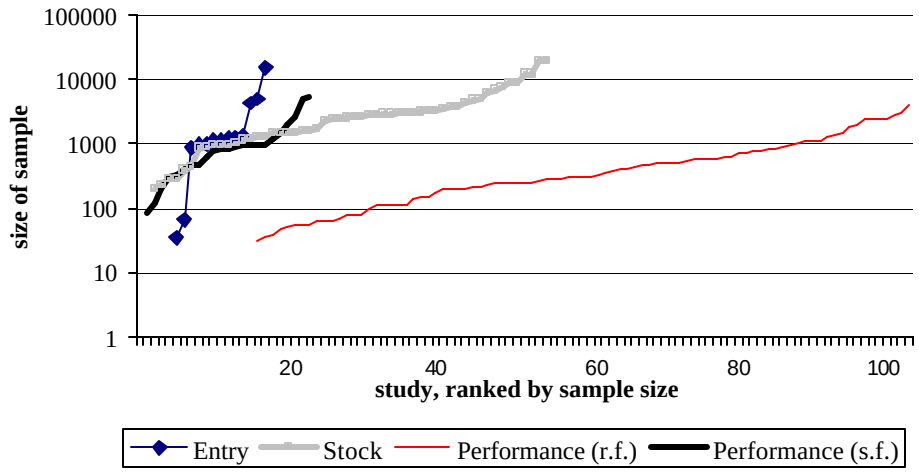

Table 1: Key Variables

A: Entry/Stock: Type of Comparisons Made

Type of Comparison	Logit/Probit	Multinomial logit	Maximum likelihood	Total
Any other form of employment	11		2	13
Other employment + nonemployed	11		1	12
Wage workers + nonemployed		2		2
Wage employment	7	32	3	42
Farming	2	9		11
Nonemployed	2	14	1	17
Unemployed		2		2
No entry	5			5
Migrant work		1		1
Contract/piece rate work		2		2
Total	38	62	7	107

B: Performance: Variation across Studies

		Linear	Logarithmic	Total
Earnings/Income/Profit:	(i) Hourly	1	34	35
	(ii) Monthly	6	19	25
	(iii) Annual	3	3	8[a]
	(iv) Unspecified	1	1	2
Inputs/Size				21
Technical efficiency index		9	1	10
Duration/Survival				19
Other				9
Total		20	58	129

[a] Two studies examined annual performance measures but not in a linear or logarithmic manner.

C: Education: Variation across Studies

	Entry	Stock	Performance	Total
Years of schooling				
Entered linearly	9	20	43	72
Entered in quadratic form		9	15	24
Entered in spline form			9	9
Dummy variables				
Lower Elementary	1	7	7	15
Upper Elementary	1	7	7	15
Elementary	1	19	18	38
Lower Secondary	7	9	11	27
Upper Secondary	7	9	11	27
Secondary	2	15	29	46
Postsecondary	3	15	19	37
College Graduate		9	9	18
Postgraduate		3	2	5
Master's Degree			3	3
PhD Degree			2	2
Other human capital variables				
Dummy: Literate		6	4	10
Training	1	1	20	22
Apprenticeship	1	8	6	15

Table 2: Characteristics of Studies

A: Field of publication (row percentage)

	Journal category of published studies						Number of published studies	Number of working papers
	GE	L/E	Dev	SmB/E	M/S	Total		
Entry	10.0	0.0	20.0	0.0	70.0	100.0	10	8
Stock	16.0	18.0	64.0	0.0	2.0	100.0	50	6
Performance, reduced form	10.0	5.0	56.3	25.0	3.8	100.0	80	25
Performance, structural	23.8	9.5	66.7	0.0	0.0	100.0	21	3
Overall	13.7	9.3	57.8	12.4	6.8	100.0	161	42

Note: GE = General economics; L/E = Labor and education; Dev = Development; SmB/E = Small business and entrepreneurship; M/S = Management and sociology

B: Year of publication (column percentage)

	1994 and before	Between 1995 and 2000	2001 and after	Total	Number of observations
Entry	10.0	10.0	6.4	8.7	18
Stock	24.0	30.0	27.0	27.6	56
Performance, reduced form	64.0	43.3	54.0	51.7	105
Performance, structural	1.0	16.7	12.7	11.8	24
Total	100.0	100.0	100.0	100.0	203

C: Countries Studied

	Sub-Saharan Africa	North Africa + Middle East	South Asia	East + Southeast Asia	Latin America + Caribbean	Countries in Transition	Number of observations
Entry		1		3	7	7	18
Stock	19	2	1	7	22	5	56
Performance, reduced form	45	8	9	10	31	2	105
Performance, structural	8			5	6	5	24
Total	72	11	10	25	65	20	200

D. Gender of the Samples (row percentage)

	Female only	Male only	Both male and female	Households	Total	Number of observations
Entry	5.6	61.1	16.7	16.7	100	18
Stock	25.0	25.0	37.5	12.5	100	56
Performance, reduced form	11.8	23.5	62.7	2.0	100	102
Performance, structural	20.8	16.7	50.0	12.5	100	24
Total	16.0	26.5	50.0	7.5	100	200

Table 3: Impact of Education on Performance

	significant negative (t < -1.96)	insignificant negative	insignificant positive	significant positive (t > 1.96)	Total number of observations
Years of schooling					
All years, entered linearly	1	6	14	19	40
Primary years		2	4	3	9
All secondary years			2	1	3
Lower secondary years		2	2	2	6
Upper secondary years			2	4	6
Years of schooling squared		3	12		15
Dummy variables					
Lower primary	1	1	2	1	5
Upper primary			2	3	5
All primary		3	10	4	17
Lower secondary		2	2	3	7
Upper secondary			1	6	7
All secondary	1	5	15	8	29
Post-secondary		3	5	8	16
College	1		4	4	9
Other human capital variables					
Training		7	7	5	19
Apprenticeships		3	2	1	6

Table 4: Meta-Analysis of the Effect of Years of Schooling on Performance

	Type of performance measure		
	log income	any income	any performance
Description of Subsample			
number of studies	49	52	69
Average value of *b*	0.055		
Standard deviation of *b*	0.064		
Significant negative (t < -1.96)		0	1
Insignificant negative		10	21
Insignificant positive		18	29
Significant positive (t > 1.96)		24	28

	Dependent variable in meta-analytical regression analysis					
	b		t^*		t^*	
Base model	coefficient	t-value	coefficient	t-value	coefficient	t-value
Intercept	0.012	0.58	…	…	…	…
Proportion female	0.040	1.96	0.787	1.93	0.769	2.04
Proportion urban	0.037	1.58	1.118	2.43	0.675	1.80
$N^{0.5}$	…	…	0.038	2.58	0.030	2.59
Added variables (added in groups)						
A Include inputs?	0.003	0.15	-0.499	1.23	-0.011	0.03
Include sector?	-0.030	1.24	0.362	0.83	0.185	0.53
B Year of sample	-0.002	1.05	-0.006	0.18	0.024	0.88
C Structural model?	0.017	0.73	0.856	1.70	0.643	1.54
D Published in journal?	-0.042	1.40	-0.836	1.23	-0.385	0.64
Impact factor	0.018	1.17	0.513	1.43	0.359	1.13
E Standard error of *b*	0.340	1.72	…	…	…	…
F Earnings model	…	…	…	…	0.419	1.19
G Sub-Saharan Africa	0.038	0.98	0.786	0.91	0.777	1.30
North Africa/Middle East	0.013	0.17	1.157	0.78	1.026	0.81
South Asia	0.010	0.14	0.887	0.73	-0.054	0.06
South-East Asia [a]	0.000	…	0.000	…	0.000	…
Latin America	0.049	1.33	0.941	1.14	0.973	1.78
Eastern Europe	0.022	0.51	1.309	1.42	1.234	1.72
H Proportion illiterate	-0.005	0.07	-0.452	0.39	-0.542	0.61
I Agriculture/GDP	-0.111	1.07	-0.980	0.52	-1.369	1.12
J Industry/GDP	-0.175	0.69	3.751	0.78	0.735	0.28
K Services/GDP	0.106	0.90	0.582	0.26	1.867	1.23
L Income per capita*10^5	-0.241	0.63	0.373	0.05	4.470	0.71
M Gross investment/GDP	-0.213	1.14	-2.405	0.64	-2.460	0.86
Estimation method	OLS		Ordered probit		Ordered probit	

Note: [a] Omitted category among the regional variables in group G.

Table 5: Meta-Analysis of the Effect of Schooling Dummy Variables on Performance

	Primary, combined			Secondary, combined			All post-secondary, combined		
Performance measure:	log income	any income	any performance	log income	any income	any performance	log income	any income	any performance
Description of Subsample									
number of studies	15	20	27	20	27	43	11	13	25
Average value of *b*	0.174			0.294			0.874		
Standard deviation of *b*	0.232			0.254			0.487		
Significant negative (t < -1.96)		1	1		0	1			1
Insignificant negative		2	4		5	7			3
Insignificant positive		11	14		12	18		5	9
Significant positive (t > 1.96)		6	8		10	17		8	12

		Dependent variable in meta-analytical regression analysis															
		b		t^*		t^*		*b*		t^*		t^*		*b*	t^*	t^*	
Base model		coeff.	t-val	coeff	t-val	coeff	t-val	coeff	t-val	coeff	t-val	coeff	t-val			coeff	t-val
	Intercept	0.372	2.87	…	…	…	…	0.333	2.40	…	…	…	…	…	…	…	…
	Proportion female	0.018	0.14	0.534	0.83	0.152	0.27	-0.137	1.04	-0.838	1.60	-0.601	1.33	…	…	0.327	0.53
	Proportion urban	-0.269	1.82	-0.953	1.43	-0.495	0.93	0.033	0.22	-0.082	0.15	-0.256	0.55	…	…	-0.566	0.82
	$N^{0.5}$	…	…	0.049	1.72	0.033	1.71	…	…	0.001	0.06	0.006	0.49	…	…	-0.008	0.38
Added variables (added in groups)[b]																	
A	Include inputs?	…	…	0.504	0.36	-1.190	1.16	…	…	-0.111	0.14	-0.460	-0.83	…	…	-2.126	2.51
	Include sector?	…	…	-1.786	1.51	0.087	0.14	…	…	-0.225	0.35	0.166	0.38	…	…	-0.962	1.31
B	Year of sample	0.004	0.33	-0.033	0.76	0.000	0.86	0.016	1.64	0.046	1.20	-0.001	1.47	…	…	0.000	0.75
C	Structural mo del?	0.067	0.55	-0.453	0.75	-0.142	0.30	0.060	0.43	0.628	1.23	0.088	0.22	…	…	1.231	1.63
D	Published in journal?	0.282	0.92	-1.046	0.69	-0.846	0.62	…	…	…	…	…	…	…	…	…	…
	Impact factor	-0.246	0.88	0.024	0.03	0.443	0.61	-0.354	1.52	0.368	0.55	0.175	0.37	…	…	0.654	1.04
E	Standard error of *b*	-0.422	0.27	…	…	…	…	1.762	2.90	…	…	…	…	…	…	…	…
F	Earnings model	…	…	…	…	-0.530	0.81	…	…	…	…	0.020	0.04	…	…	2.283	2.96
H	Proportion illiterate	0.656	1.05	5.652	1.64	3.908	2.16	1.325	2.42	0.928	0.54	0.626	0.54	…	…	7.976	2.52
I	Agriculture/GDP	0.651	1.19	5.131	1.77	3.639	1.60	1.534	2.61	2.460	1.05	2.526	1.33	…	…	7.852	2.18
J	Industry/GDP	-0.334	0.57	-0.293	0.12	-1.641	0.79	-1.217	1.80	-2.756	1.07	-1.881	0.89	…	…	-4.334	1.01
K	Services/GDP	-2.179	1.65	-12.782	2.40	-3.605	1.17	-1.288	1.13	-0.341	0.10	-1.791	0.72	…	…	-11.328	2.35
L	Income per capita*10^3	-0.053	1.26	-0.172	0.86	-0.173	1.06	-0.061	1.24	-0.119	0.59	-0.100	0.66	…	…	-0.098	1.36
M	Gross investment/GDP	-0.485	0.75	-0.203	0.07	-1.305	0.48	-0.931	1.32	-1.999	0.67	-0.224	0.08	…	…	-0.356	0.09
Estimation method		OLS		Ordered probit		Ordered probit		OLS		Ordered probit		Ordered probit		OLS	Ordered probit	Ordered probit	

Note: [a] The group of regional variables is omitted for lack of variation.

Table 6: Education and Entrepreneurship Choice: Descriptive Analysis

		Impact on probability of non-farm self-employment		Significance levels			
Education variable	Number of observations	Mean	Standard deviation	Significant negative	Insignificant negative	Insignificant positive	Significant positive
A. Relative to any other form of employment							
years	10	-0.013	0.036	6		4	
primary	3	-0.055	0.047	2	1		
secondary	6	-0.052	0.043	2	2	2	
post-sec	3	-0.374	0.155	3			
B. Relative to any other form of employment or non-employment							
years	2	0.003	0.002			1	1
primary	12	0.023	0.039		5	4	3
secondary	10	0.053	0.068		2	5	3
post-sec	8	0.013	0.057		4	4	
C. Relative to wage employment							
years	24	-0.068	0.193	11	7	2	4
primary	19	-0.209	0.383	8	7	2	2
secondary	24	-0.350	0.520	14	4	2	4
post-sec	16	-0.683	0.880	12	1	1	2
D. Relative to farming							
years	7	0.081	0.128			4	3
primary	4	0.113	0.173		1		3
secondary	7	0.312	0.320			1	6
post-sec	1	0.224	…				1
E. Relative to non-employment							
years	11	-0.008	0.018	1	6	4	
primary	6	-0.059	0.177	4	1	1	
secondary	6	-0.183	0.187	1	5		
post-sec	3	-0.397	0.600		2		1
F. Relative to no entry into non-farm self-employment							
secondary	10	0.118	0.165	2	2	6	

Table 7: Education and Entrepreneurship Choice: Meta Analysis of the Choice between Wage Employment and Non-Farm Self-Employment[a]

		Years (n=24)		Primary (n=19)		Secondary (n=24)		Post-secondary (n=16)	
		coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value
Base model									
	Intercept	0.124	1.26	0.349	1.12	0.494	1.87	0.841	0.69
	Proportion female	-0.283	2.51	-0.418	2.11	-0.692	2.82	-1.031	2.15
	Proportion urban	-0.082	0.90	-0.422	1.36	-0.703	2.73	-1.180	0.96
Added variables									
A	Year of sample	0.008	1.34	-0.047	0.95	-0.012	0.23	0.007	0.06
B	Structural model?	0.097	0.53	-0.084	0.22	-0.129	0.29	-0.038	0.04
C	Study of entry?	0.048	0.47	…		…		…	
D	Published in journal?	-0.058	0.44	…		…		…	
	Impact factor	-0.016	0.17	0.420	0.75	0.679	1.05	1.620	0.95
E	Standard error of *b*	0.520	0.13	…		…		…	
F	Proportion illiterate	-0.368	2.58	-0.215	0.75	0.075	0.16	2.153	0.82
G	Agriculture/GDP	-1.244	2.39	0.604	0.59	1.404	1.38	2.916	1.34
H	Industry/GDP	1.470	2.29	-1.717	1.81	-2.456	2.56	-4.482	2.38
I	Services/GDP	0.497	0.62	6.777	3.56	7.630	2.84	13.017	2.51
J	Income per capita*10^3	0.055	2.26	-0.004	0.12	-0.017	0.42	-0.342	1.39
K	Gross investment/GDP	0.918	1.77	2.094	1.46	1.968	1.25	2.624	0.74

Note: [a] The dependent variable is the parameter estimate (*b*) of the specified education variable in the model that explains the choice between wage employment (as the base) and non-farm self-employment. The meta analytical regression models are estimated by means of OLS. The number of observations in each regression model is indicated in parentheses at the head of each column.